Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

Compensated Compactness, Separately convex Functions and Interpolatory Estimates between Riesz Transforms and Haar Projections

by

Jihoon Lee, Paul F. X. Müller, and Stefan Müller

Preprint no.: 7 2008

Compensated Compactness, Separately convex Functions and Interpolatory Estimates between Riesz Transforms and Haar Projections

Jihoon Lee[*], Paul F. X . Müller [†]and Stefan Müller

Dezember 6 th, 2007.

Contents

1 The main Results 2
1.1 Interpolatory Estimates 2
1.2 Lower semi-continuity and compensated compactness 4

2 Multiscale Analysis of directional Haar Projections 11

3 Tooling up 14

4 Basic Dyadic Operations 21
4.1 Projections and Ring Domains 21
4.2 Rearrangement Operators 23

5 The Proof of Theorem 2.1. 27
5.1 Estimates for $T_{\ell,m}$, $\ell \geq 0$, $m < -\ell$. 29
5.2 Estimates for $T_{\ell,m}$, $\ell \geq 0$, $m > 0$. 31
5.3 Estimates for $T_{\ell,m}$, $\ell \geq 0$, $-\ell \leq m \leq 0$. 33
5.4 Estimates for $T_\ell^{(\varepsilon)} R_{i_0}^{-1}$, $\ell \geq 0$. 35

6 The Proof of Theorem 2.2. 39

7 Sharpness of the exponents in Theorem 1.1. 40
7.1 The building blocks $s \otimes d$. 41
7.2 Upper estimate for $\|f_\epsilon\|_p$ and $\|R_1(f_\epsilon)\|_p$. 42
7.3 Lower bound for $\|P(f_\epsilon)\|_p$, $p \geq 2$. 46
7.4 The proof of theorem 7.1 48

[*]J.L. supported by Korean Research Foundation Gr. Nr. KRF-2006-311-C00007.
[†]P.F.X.M. is supported by the Austrian Science foundation (FWF) Pr.Nr. P15907-N08.

1 The main Results

In this work we prove sharp interpolatory estimates that exhibit a new link between Riesz transforms and directional projections of the Haar system in $\mathbb{R}^n$. To a given direction $\varepsilon \in \{0,1\}^n, \varepsilon \neq (0,\dots,0)$, we let $P^{(\varepsilon)}$ be the orthogonal projection onto the span of those Haar functions that oscillate along the coordinates $\{i : \varepsilon_i = 1\}$. When $\varepsilon_{i_0} = 1$ the identity operator and the Riesz transform R_{i_0} provide a logarithmically convex estimate for the L^p norm of $P^{(\varepsilon)}$, see Theorem 1.1. Apart from its intrinsic interest Theorem 1.1 has direct applications to variational integrals, the theory of compensated compactness, Young measures, and to the relation between rank one and quasi convex functions. In particular we exploit our Theorem 1.1 in the course of proving a conjecture of L. Tartar on semi-continuity of separately convex integrands; see Theorem 1.5.

1.1 Interpolatory Estimates

We first recall the definitions of the Haar system in $\mathbb{R}^n$, indexed and supported on dyadic cubes, its associated directional Haar projections and the usual Riesz transforms; thereafter we state the main theorem of this paper.

Let $\mathcal{D}$ denote the collection of dyadic intervals in the real line. Thus $I \in \mathcal{D}$ if there exists $i \in \mathbb{Z}$ and $k \in \mathbb{Z}$ so that $I = [i2^k, (i+1)2^k[$. Define the Haar function over the unit interval as

$$h_{[0,1[} = 1_{[0,1/2[} - 1_{[1/2,1[}.$$

The L^∞ normalized Haar system $\{h_I : I \in \mathcal{D}\}$ is obtained from $h_{[0,1[}$ by rescaling. Let $I \in \mathcal{D}$, let l_I denote the left endpoint of I, thus $l_I = \inf I$. Then put

$$h_I(x) = h_{[0,1[}\left(\frac{x - l_I}{|I|}\right), \qquad x \in \mathbb{R}.$$

Thus defined, the Haar system $\{h_I : I \in \mathcal{D}\}$ is a complete orthogonal system in $L^2(\mathbb{R})$. Next we recall its n dimensional analog. Let $I_1, \dots, I_n$ be dyadic intervals so that $|I_i| = |I_j|$, where $1 \leq i, j \leq n$. Define the dyadic cube $Q \subset \mathbb{R}^n$,

$$Q = I_1 \times \cdots \times I_n.$$

Let $\mathcal{S}$ denote the collection of all dyadic cubes in $\mathbb{R}^n$. To define the associated Haar system consider first $\mathcal{A} = \{\varepsilon \in \{0,1\}^n : \varepsilon \neq (0,\dots,0)\}$. For $Q = I_1 \times \cdots \times I_n \in \mathcal{S}$ and $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n) \in \mathcal{A}$ let

$$h_Q^{(\varepsilon)}(x) = \prod_{i=1}^n h_{I_i}^{\varepsilon_i}(x_i), \qquad x = (x_1, \dots, x_n). \tag{1.1}$$

We call $\{h_Q^{(\varepsilon)} : Q \in \mathcal{S}, \varepsilon \in \mathcal{A}\}$ the Haar system in $\mathbb{R}^n$. It is a complete orthogonal system in $L^2(\mathbb{R}^n)$. Hence for $u \in L^2(\mathbb{R}^n)$,

$$u = \sum_{\varepsilon \in \mathcal{A},\, Q \in \mathcal{S}} \langle u, h_Q^{(\varepsilon)} \rangle h_Q^{(\varepsilon)} |Q|^{-1}, \tag{1.2}$$

where the series on the right hand side converges unconditionally in $L^2(\mathbb{R}^n)$. For $\varepsilon \in \mathcal{A}$ define the associated directional projection on $L^2(\mathbb{R}^n)$ by

$$P^{(\varepsilon)}(u) = \sum_{Q \in \mathcal{S}} \langle u, h_Q^{(\varepsilon)} \rangle h_Q^{(\varepsilon)} |Q|^{-1}, \qquad u \in L^2(\mathbb{R}^n).$$

The operators $P^{(\varepsilon)}, \varepsilon \in \mathcal{A}$, project onto orthogonal subspaces of $L^2(\mathbb{R}^n)$ so that

$$u = \sum_{\varepsilon \in \mathcal{A}} P^{(\varepsilon)}(u) \quad \text{and} \quad \|u\|_2^2 = \sum_{\varepsilon \in \mathcal{A}} \|P^{(\varepsilon)}(u)\|_2^2. \tag{1.3}$$

Let $\mathcal{F}$ denote the Fourier transformation on $\mathbb{R}^n$ given as

$$\mathcal{F}(u)(\xi) = \int_{\mathbb{R}^n} e^{-i\langle x,\xi\rangle} u(x)dx, \quad \xi \in \mathbb{R}^n, \quad x \in \mathbb{R}^n.$$

The Riesz transform R_i $(1 \leq i \leq n)$ is a Fourier multiplier defined by

$$R_i(u)(x) = -\sqrt{-1}\mathcal{F}^{-1}\left(\frac{\xi_i}{|\xi|}\mathcal{F}(u)(\xi)\right)(x) \quad \text{where} \quad \xi = (\xi_1, \ldots, \xi_n).$$

The analytic backbone of this paper is the following theorem showing that the norm in $L^p(\mathbb{R}^n)$ of $P^{(\varepsilon)}(u)$ is dominated through a logarithmically convex estimate by $R_{i_0}(u)$, provided that a carefully analyzed relation holds between i_0 (appearing in the Riesz transform) and ε defining the directional projections $P^{(\varepsilon)}$.

Theorem 1.1 *Let $1 < p < \infty$ and $1/p + 1/q = 1$. For $1 \leq i_0 \leq n$ define*

$$\mathcal{A}_{i_0} = \{\varepsilon \in \mathcal{A} : \varepsilon = (\varepsilon_1, \ldots \varepsilon_n) \quad \textit{and} \quad \varepsilon_{i_0} = 1\}.$$

Let $u \in L^p(\mathbb{R}^n)$. If $\varepsilon \in \mathcal{A}_{i_0}$ then $P^{(\varepsilon)}$ and R_{i_0} are related by interpolatory estimates in $L^p(\mathbb{R}^n)$,

$$||P^{(\varepsilon)}(u)||_p \leq C_p \|u\|_p^{1/2} \|R_{i_0}(u)\|_p^{1/2} \qquad \textit{if} \quad p \geq 2,$$

and

$$||P^{(\varepsilon)}(u)||_p \leq C_p \|u\|_p^{1/p} \|R_{i_0}(u)\|_p^{1/q} \qquad \textit{if} \quad p \leq 2.$$

The exponents $(1/2, 1/2)$ for $p \geq 2$ and $(1/p, 1/q)$ for $p \leq 2$ appearing in Theorem 1.1 are sharp. We show in Section 7 that for $\eta > 0$, $1 < p < \infty$ and $N >> 1$ there exists $u = u_{\eta,p,N} \in L^p$ so that

$$||P^{(\varepsilon)}(u)||_p \geq N \|u\|_p^{1/2-\eta} \|R_{i_0}(u)\|_p^{1/2+\eta} \qquad \text{if} \quad p \geq 2,$$

and

$$||P^{(\varepsilon)}(u)||_p \geq N \|u\|_p^{1/p-\eta} \|R_{i_0}(u)\|_p^{1/q+\eta} \qquad \text{if} \quad p \leq 2.$$

A first consequence of Theorem 1.1. In the next subsection we will show how Theorem 1.1 is used in problems originating in the theory of compensated compactness. To this end we formulate here a concise inequality that follows from the above interpolatory estimates, and record its immediate consequences. See (1.5)–(1.7).

Let $1 \leq j \leq n$. Let $e_j \in \mathcal{A}$ denote the unit vector in $\mathbb{R}^n$ pointing in the positive direction of the $j-th$ coordinate axis, $e_j = (0, \ldots, 1, \ldots, 0)$, where 1 appears in the $j-th$ entry. By (1.3)

$$u - P^{(e_j)}(u) = \sum_{\varepsilon \in \mathcal{A} \setminus \{e_j\}} P^{(\varepsilon)}(u).$$

The above identity and the estimates of Theorem 1.1 combined yield the inequality

$$||u - P^{(e_j)}(u)||_p \leq C_{p,n} \|u\|_p^{1/2} \left[\sum_{\substack{1 \leq i \leq n \\ i \neq j}} \|R_i(u)\|_p\right]^{1/2}, \qquad p \geq 2. \tag{1.4}$$

On $L^p(\mathbb{R}^n, \mathbb{R}^n)$ define the vector valued projection P by

$$P(v) = \left(P^{(e_1)}(v_1), \ldots, P^{(e_n)}(v_n)\right),$$

where $v : \mathbb{R}^n \to \mathbb{R}^n$, $v = (v_1, \ldots, v_n)$. Applying (1.4) to each component of v yields

$$\|v - P(v)\|_p \leq C_{p,n}\|v\|^{1/2} \cdot \left(\sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} \|R_i(v_j)\|_p\right)^{1/2} \tag{1.5}$$

Assume now that $(v_{r,1}, \ldots, v_{r,n})$ is a sequence in $L^p(\mathbb{R}^n, \mathbb{R}^n)$ so that

$$\lim_{r\to\infty} \|R_i(v_{r,j})\|_p = 0 \quad \text{for} \quad 1 \leq i \leq n, i \neq j. \tag{1.6}$$

The assumption (1.6) and the estimate (1.5) imply that

$$\lim_{r\to\infty} \|(v_{r,1}, \ldots, v_{r,n}) - P\left((v_{r,1}, \ldots, v_{r,n})\right)\|_p = 0. \tag{1.7}$$

Being able to draw the conclusion (1.7) from the hypothesis (1.6) provided the main impetus for proving Theorem 1.1.

1.2 Lower semi-continuity and compensated compactness

Here we provide a frame of reference for the problems considered in this paper. We review briefly some of the ideas of the theory of compensated compactness which has been developed by F. Murat and L. Tartar [12, 14, 16, 17].

Weak lower-semicontinuity and differential constraints. Fix a system of first-order, linear differential operators $\mathcal{A}$. It is given by matrices $A^{(i)} \in \mathbb{R}^{p\times d}$, $i \leq n$, so that

$$\mathcal{A}(v) = \sum_{i=1}^{n} A^{(i)} \partial_i(v),$$

where $v : \mathbb{R}^n \to \mathbb{R}^d$ and ∂_i denotes the partial differentiation with respect to the $i-$th coordinate. To $\mathcal{A}$ we associate the cone $\Lambda \subseteq \mathbb{R}^d$ of "dangerous" amplitudes. It consists of those $a \in \mathbb{R}^d$ for which there is a vector of frequencies $\xi \in \mathbb{R}^n, \xi \neq 0$, so that for any smooth $h : \mathbb{R} \to \mathbb{R}$ the function

$$w(x) = ah(\langle \xi, x\rangle),$$

satisfies

$$\mathcal{A}(w) = 0.$$

Thus, to $a \in \Lambda$ there exists a non-zero $\xi \in \mathbb{R}^n$, so that $\mathcal{A}(w_m) = 0$ for the increasingly oscillatory sequence

$$w_m(x) = a\sin(m\langle \xi, x\rangle), \quad m \in \mathbb{N}.$$

Since $\xi \neq 0$ there is $i_0 \leq n$ so that the sequence of partial derivatives $\partial_{i_0} w_m$ is unbounded while $\mathcal{A}(w_m) = 0$. In other words, the linear differential constraint $\mathcal{A}(w) = 0$ does not imply any control on the partial derivative ∂_{i_0}. Expressed formally, the cone of "dangerous" amplitudes is given as

$$\Lambda = \left\{ a \in \mathbb{R}^d : \exists \xi \in \mathbb{R}^n \setminus \{0\} \quad \text{such that} \quad \sum_{i=1}^{n} \xi_i A^{(i)}(a) = 0 \right\}.$$

The methods of compensated compactness allow one to exploit a given set of information on the differential constraints $\mathcal{A}(v)$ (respectively on Λ) to analyze the limiting behaviour of non-linear integrands acting on v under weak conmvergence Consider a sequence of functions $v_r : \mathbb{R}^n \to \mathbb{R}^d$ so that

$$v_r \rightharpoonup v \quad \text{weakly in} \quad L^p(\mathbb{R}^n, \mathbb{R}^d), \tag{1.8}$$

and

$$\mathcal{A}(v_r) \quad \text{precompact in} \quad W^{-1,p}(\mathbb{R}^n, \mathbb{R}^d). \tag{1.9}$$

The following comments are included to clarify the relation between the hypotheses (1.8) and (1.9).

1. Had we imposed, instead of (1.8), that $v_r \to v$ strongly in $L^p(\mathbb{R}^n, \mathbb{R}^d)$, then (1.9) would hold automatically.
2. More subtle aspects of the interplay between (1.8) and (1.9) are depending on the structure of $\mathcal{A}$ or Λ. For instance, in the special case when $\mathcal{A}(v)$ controls all partial derivatives of v, we use Sobolev's *compact* embedding theorem to see that (1.9), implies that $v_r \to v$ strongly in $L^p(\mathbb{R}^n, \mathbb{R}^d)$. This case occurs when $\Lambda = \{0\}$,
3. The generic (and most interesting) case arises when $\mathcal{A}(v)$ fails to control some of the partial derivatives of v. This occurs when $\Lambda \neq \{0\}$.

In the generic case one goal of the theory is to isolate sharp conditions on a given $f : \mathbb{R}^d \to \mathbb{R}$ that *compensate* for the lack of compactness provided by $\mathcal{A}$, and ensure that (1.8) and (1.9) imply

$$\liminf_{r \to \infty} \int_{\mathbb{R}^n} f(v_r(x))\varphi(x)dx \geq \int_{\mathbb{R}^n} f(v(x))\varphi(x)dx, \qquad \varphi \in C_o^+(\mathbb{R}^n). \tag{1.10}$$

Here (and below) $C_o^+(\mathbb{R}^n)$ denotes the set of non-negative compactly supported continous functions on $\mathbb{R}^n$. Note that up to growth conditions on f and up to passing to subsequences of v_r, the condition (1.10) states that

$$\text{weak limit} \quad f(v_r) \geq f(v).$$

In summary, based on knowledge of $\mathcal{A}$ or Λ one goal of the theory of compensated compactness aims at describing and classifying those non-linearities $f : \mathbb{R}^d \to \mathbb{R}$ for which (1.8) and (1.9) imply (1.10).

Classical results on compensated compactness. We assume now that (1.8) and (1.9) hold and that the differential operator $\mathcal{A}$ satisfies the so called constant rank hypothesis; for its definition see below. The classical results of compensated compactness, as developed by F. Murat and L. Tartar [12, 14, 16, 17] assert that a general non-linearity f satisfies (1.10) precisely when it is $\mathcal{A}-$quasi-convex. Furthermore, in the special case of a quadratic integrand $f(a) = \langle Ma, a \rangle$ the constant rank hypothesis is not needed and the conclusion (1.10) is equivalent to $\Lambda-$convexity of $f(a) = \langle Ma, a \rangle$. We state now explictely the characterizations mentioned above, and recall the notions of $\Lambda-$convexity, $\mathcal{A}-$quasi-convexity, and the constant rank hypothesis on $\mathcal{A}$.

A function $f : \mathbb{R}^d \to \mathbb{R}$ is $\Lambda-$ convex if

$$f(\lambda a + (1-\lambda)b) \leq \lambda f(a) + (1-\lambda) f(b), \quad a - b \in \Lambda, \ 0 < \lambda < 1.$$

The following result is due to F. Murat [12], [13] and L. Tartar [17].

Proposition 1.2 *If for every sequence $v_r : \mathbb{R}^n \to \mathbb{R}^d$, the hypotheses* (1.8) *and* (1.9) *imply* (1.10)*, then $f : \mathbb{R}^d \to \mathbb{R}$ is $\Lambda-$convex.*

Thus $\Lambda-$convexity is a necessary condition on f for (1.8) and (1.9) to imply (1.10). If, moreover f is quadratic,

$$f(a) = \langle Ma, a \rangle, \quad M \in \mathbb{R}^{d\times d}, a \in \mathbb{R}^d,$$

then $\Lambda-$convexity is already sufficient. This is the content of the following result by L. Tartar [17].

Theorem 1.3 *Assume that f is quadratic and $\Lambda-$convex. Then, for every sequence $v_r : \mathbb{R}^n \to \mathbb{R}^d$,* (1.8) *and* (1.9) *imply* (1.10).

We next review the results beyond the case of quadratic integrands. They involve the notion of $\mathcal{A}-$quasi-convexity and the constant rank hypothesis. We define $f : \mathbb{R}^d \to \mathbb{R}$ to be $\mathcal{A}-$quasi-convex if

$$\int_{[0,1]^n} f(a + u(x))dx \geq f(a), \tag{1.11}$$

for each smooth and $[0,1]^n$ periodic $u : \mathbb{R}^n \to \mathbb{R}^d$, that satisfies $\int_{[0,1]^n} u = 0$ and $\mathcal{A}(u) = 0$. Note that (1.11) asks for Jensen's inequality to hold under the decisive restriction that $\mathcal{A}(w) = 0$. It was proved essentially by C.B. Morrey [8] that $\mathcal{A}-$quasi-convexity implies $\Lambda-$convexity (see [3]). The linear differential operator $\mathcal{A}$ satisfies the constant rank hypothesis if there exists $r \leq n$ so that

$$rk(A(\xi)) = r, \quad \xi \in \mathbb{S}^{n-1},$$

where

$$A(\xi) = \sum_{i=1}^{n} \xi_i A^{(i)}.$$

The next theorem provides a full characterization of those integrands f for which (1.8) and (1.9) imply (1.10).

Theorem 1.4 ([14]) *Let $0 \leq f(a) \leq C(1+|a|^p)$ and assume that $\mathcal{A}$ satisfies the constant rank hypothesis. Then $f : \mathbb{R}^d \to \mathbb{R}$ is $\mathcal{A}-$ quasi-convex if and only if* (1.8) *and* (1.9) *imply* (1.10).

A crucial component in the proof of Theorem 1.4 links the constant rank hypothesis and $\mathcal{A}-$quasi-convexity as follows:

1. Let $v : \mathbb{R}^n \to \mathbb{R}^d$ be $[0,1]^n$ periodic and of mean zero in $[0,1]^n$. Under the constant rank hypothesis, there exists a decomposition of v as

   $$v = u + w,$$

   where

   $$\mathcal{A}(u) = 0 \quad \text{and} \quad \|w\|_{L^p([0,1]^n)} \leq C \|\mathcal{A}(v)\|_{W^{-1,p}([0,1]^n)}.$$

   The decomposition can be expressed in terms of an explicit Fourier multiplier, for which standard L^p estimates are available, provided that the constant rank hypothesis holds.

2. Let now $v_r \in L^p([0,1]^n, \mathbb{R}^d)$ be a sequence of $[0,1]^n$ periodic, mean zero functions so that $\mathcal{A}(v_r) \to 0$ in $W^{-1,p}$. Then, by the foregoing remark, we may split v_r as $v_r = u_r + w_r$ so that

$$\mathcal{A}(u_r) = 0 \quad \text{and} \quad \|w_r\|_p \to 0. \tag{1.12}$$

3. Assume moreover that f is $\mathcal{A}-$quasi-convex. The decomposition

$$v_r = u_r + w_r \tag{1.13}$$

with the properties (1.12) satisfies then

$$\int_{[0,1]^n} f(a + u_r(x))dx \geq f(a), \quad \text{and} \quad \|w_r\|_p \to 0. \tag{1.14}$$

Separately convex integrands. Wide ranging applications illustrate the power of Theorem 1.4, yet there are important linear differential constraints $\mathcal{A}$, for which the constant rank hypothesis does not hold and the classical proof does not apply. Among the earliest examples considered is the following $\mathcal{A}_0$, defined as

$$(\mathcal{A}_0(v))_{i,j} = \begin{cases} \partial_i v_j & i \neq j; \\ 0 & i = j, \end{cases}$$

where $v : \mathbb{R}^n \to \mathbb{R}^n$. Observe that for $v = (v_1, \ldots, v_n)$ the condition $\mathcal{A}_0(v) = 0$ holds precisely when $v_i : \mathbb{R}^n \to \mathbb{R}$ is actually a function of the variable x_i alone, that is $v_i(x) = v_i(x_i)$. By a direct calculation, the cone of dangerous amplitudes associated to $\mathcal{A}_0$ is given as

$$\Lambda_0 = \bigcup_{i=1}^n \mathbb{R}e_i,$$

where $\{e_i\}$ denotes the unit vectors in $\mathbb{R}^n$. It follows that the Λ_0-convex functions are just separately convex functions on $\mathbb{R}^n$.

For the operator $\mathcal{A}_0$ the constant rank hypothesis, does not hold, since $\ker A_0(\xi) = 0$ for $\xi \in \{e_1, \ldots, e_n\}$ and $\ker A_0(e_i) = \mathbb{R}e_i$, $i \leq n$. As a result the classical theory of compensated compactness for non quadratic functionals does not apply to the operator $\mathcal{A}_0$. Nevertheless it is an important consequence of the interpolatory estimates in Theorem 1.1 that separately convex functions yield weakly semi-continuous integrands on sequences $v_r : \mathbb{R}^n \to \mathbb{R}^n$ for which $\mathcal{A}_0(v_r)$ is precompact in $W^{-1,p}(\mathbb{R}^n, \mathbb{R}^d)$. The following theorem verifies a conjecture formulated by L.Tartar [19].

Theorem 1.5 *Let $1 < p < \infty$. Assume that $f : \mathbb{R}^n \to \mathbb{R}$ is Λ_0- convex and satisfy $0 \leq f(a) \leq C(1 + |a|^p)$. Let $v_r : \mathbb{R}^n \to \mathbb{R}^n$ satisfy*

$$v_r \rightharpoonup v \quad \text{weakly in} \quad L^p(\mathbb{R}^n, \mathbb{R}^n), \tag{1.15}$$

and

$$\mathcal{A}_0(v_r) \quad \text{precompact in} \quad W^{-1,p}(\mathbb{R}^n, \mathbb{R}^n). \tag{1.16}$$

Then,

$$\liminf_{r\to\infty} \int_{\mathbb{R}^n} f(v_r(x))\varphi(x)dx \geq \int_{\mathbb{R}^n} f(v(x))\varphi(x)dx, \qquad \varphi \in C_o^+(\mathbb{R}^n). \tag{1.17}$$

As discussed in [10] this result implies that gradient Young measures supported on diagonal entries are laminates, and this in turn gives an interesting relation between rank-one convexity and quasi-convexity on subspaces with few rank-one directions.

In the approach of the present paper we fully exploit the methods introduced in [10]. We base the proof of Theorem 1.5 on the decomposition given by the directional Haar projection

$$v = P(v) + \{v - P(v)\},$$

invoke the interpolatory estimates of Theorem 1.1, and use the fact that Λ_0-convexity yields Jensen's inequality on the range of P :

1. By inequality (1.5), the norm of $\{v - P(v)\}$ in L^p is controlled by the norm of $\mathcal{A}_0(v)$ in $W^{-1,p}$.

2. The operator $\mathcal{A}_0$ does not exert any control over $P(v)$. It is Λ_0-convexity that compensates for that. Indeed when f is separately convex we have the following form of Jensen's inequality
$$f\left(\int_{[0,1]^n} P(v)dx\right) \leq \int_{[0,1]^n} f(P(v))dx. \tag{1.18}$$
By rescaling of (1.18) we get
$$f(E_M(P(v))) \leq E_M(f(P(v))), \quad v \in L^p(\mathbb{R}^n, \mathbb{R}^n), \quad M \in \mathbb{Z}, \tag{1.19}$$
where E_M denotes the conditional expectation operatpor given as
$$E_M(g)(x) = \sum_{\{R \in \mathcal{S}: |R| = 2^{-Mn}\}} \int_{\mathbb{R}^n} g(y)\frac{dy}{|R|} 1_R(x), \quad g \in L^p(\mathbb{R}^n).$$
We verify (1.18) below. The proof is based on the observation that Haar functions are exactly localized, three-valued martingale differences.

3. Assume that f is separately convex and that $v_r \in L^p([0,1]^n, \mathbb{R}^n)$ is a sequence of $[0,1]^n$ periodic, mean zero functions so that $\mathcal{A}_0(v_r) \to 0$ in $W^{-1,p}$. With $u_r = P(v_r)$ and $w_r = \{v_r - P(v_r)\}$, the decomposition
$$v_r = u_r + w_r \tag{1.20}$$
satisfies the central properties
$$\int_{[0,1]^n} f(a + u_r(x))dx \geq f(a), \quad \text{and} \quad \|w_r\|_p \to 0. \tag{1.21}$$
The splitting (1.20) with the property (1.21) is parallel to the classical decomposition (1.13) and (1.14) based on Fourier multipliers and the constant rank hypothesis.

Jensen's inequality on the range of P. We prove (1.18) by induction over the levels of the Haar system. Fix e_j, the unit vector in $\mathbb{R}^n$ pointing along the $j-$th coordinate axis and a dyadic cube $Q = I_1 \times \cdots \times I_n$. The restriction of $h_Q^{(e_j)}$ to the cube Q is a function of x_j alone, indeed
$$h_Q^{(e_j)}(x) = h_{I_j}(x_j), \quad x \in Q.$$

Hence for $a = (a_1, \dots, a_n)$ and $c = (c_1, \dots, c_n)$ we have the identity

$$\begin{aligned}&\int_Q f(a_1 + c_1 h_Q^{(e_1)}(x), \; \dots \;, a_n + c_n h_Q^{(e_n)}(x))dx \\ &= \int_Q f(a_1 + c_1 h_{I_1}(x_1), \; \dots \;, a_n + c_n h_{I_n}(x_n))dx.\end{aligned} \tag{1.22}$$

Using (1.22) and applying Jensen's inequality to each of the variables $x_1, \dots, x_n$ of the separately convex integrand f gives

$$\int_Q f(a_1 + c_1 h_Q^{(e_1)}(x), \; \dots \;, a_n + c_n h_Q^{(e_n)}(x))dx \geq |Q| f(a). \tag{1.23}$$

Next we fix $v = (v_1, \dots, v_n) \in L^p(\mathbb{R}^n, \mathbb{R}^n)$ and assume that v_j is finite linear combination of Haar functions and not constant over the unit cube. Define

$$A_{k,j} = \sum_{\{Q \in S : |Q| = 2^{-kn}\}} c_{Q,j} h_Q^{(e_j)}, \quad c_{Q,j} = \langle v_j, h_Q^{(e_j)} \rangle |Q|^{-1}.$$

Choose $M \in \mathbb{N}$ and put

$$S_{M,j} = \sum_{k=-\infty}^{M} A_{k,j}.$$

By our assumption on v_j the sum defining $S_{M,j}$ is actually finite, and there exists M_0 with $M_0 \geq 0$ so that

$$S_{M_0,j} = P^{(e_j)}(v_j), \quad 1 \leq j \leq n.$$

Choose now $M \leq M_0$. Fix a dyadic cube Q contained in $[0,1]^n$ with $|Q| = 2^{-Mn}$. Note that $S_{M-1,j}$ is constant on Q, and put $a_j = S_{M-1,j}(y)$ where $y \in Q$ is chosen arbitrarily. Furthermore,

$$A_{M,j}(x) = c_{Q,j} h_Q^{(e_j)}(x), \quad x \in Q.$$

Then, using $S_{M,j} = S_{M-1,j} + A_{M,j}$ and (1.23) we obtain

$$\begin{aligned}\int_Q f\left(S_{M,1}(x), \dots, S_{M,n}(x)\right) dx &= \int_Q f\left(a_1 + c_{Q,1} h_Q^{(e_1)}(x), \dots, a_n + c_{Q,n} h_Q^{(e_n)}(x)\right) dx \\ &\geq |Q| f\left(S_{M-1,1}(y), \dots, S_{M-1,n}(y)\right).\end{aligned} \tag{1.24}$$

It follows from (1.24) by taking the sum over $Q \subset [0,1]^n$ with $|Q| = 2^{-Mn}$, that

$$\int_{[0,1]^n} f\left(S_{M,1}(x), \dots, S_{M,n}(x)\right) dx \geq \int_{[0,1]^n} f\left(S_{M-1,1}(y), \dots, S_{M-1,n}(y)\right) dy.$$

We next replace M by $M-1$ and repeat. Starting the process with $M = M_0$ and stopping at $M = 1$ yields the claimed inequality

$$\int_{[0,1]^n} f\left(S_{M_0,1}(x), \dots, S_{M_0,n}(x)\right) dx \geq f\left(\int_{[0,1]^n} P(v)\right).$$

■

Proof of Theorem 1.5 : Choose $v \in L^p(\mathbb{R}^n, \mathbb{R}^n)$ and a sequence $v_r \in L^p(\mathbb{R}^n, \mathbb{R}^n)$ so that (1.15) and (1.16) hold. Let $C_0^+((0,1)^n)$ denote the continuous, non-negative and compactly supported functions on the open unit cube $(0,1)^n$. We first show the conclusion (1.17) under the additional restriction that

$$v_{|(0,1)^n} = const, \qquad \text{and} \qquad \varphi \in C_0^+((0,1)^n). \tag{1.25}$$

Clearly we may then assume that $v_{|(0,1)^n} = 0$, since otherwise we replace f by $f(\cdot + c)$. Next we choose a smooth function $\alpha \in C_0^+((0,1)^n)$ so that $\alpha(x) = 1$ for $x \in \operatorname{supp} \varphi$. By considering the sequence (αv_r) instead of (v_r) we may further assume that

$$v_r \rightharpoonup 0 \text{ weakly in } L^p \text{ and } \mathcal{A}_0(v_r) \to 0 \text{ in } W^{-1,p}. \tag{1.26}$$

By (1.26) we obtain for $v_r = (v_{r,1}, \ldots, v_{r,n})$ that

$$\lim_{r\to\infty} \|R_i(v_{r,i})\|_{L^p(\mathbb{R}^n)} = 0, \quad i \neq j.$$

Hence by (1.7),

$$\lim_{r\to\infty} \|v_r - P(v_r)\|_{L^p(\mathbb{R}^n,\mathbb{R}^n)} = 0. \tag{1.27}$$

Since f is separately convex and satisfies $f(t) \leq C(1 + |t|^p)$ we get

$$|f(s) - f(t)| \leq C(1 + |s| + |t|)^{p-1}|s - t|. \tag{1.28}$$

Using (1.28) and $1/p + 1/q = 1$ gives

$$\begin{aligned}\int_{\mathbb{R}^n} f(v_r)\varphi dx &= \int_{\mathbb{R}^n} f(P(v_r))\varphi dx + \int_{\mathbb{R}^n} \left(f(v_r) - f(P(v_r)\right)\varphi dx \\ &\geq \int_{\mathbb{R}^n} f(P(v_r))\varphi dx - C\|1 + |v_r| + |P(v_r)|\|_p^{p/q}\|v_r - P(v_r)\|_p.\end{aligned} \tag{1.29}$$

Next fix M and rewrite by adding and subtracting the conditional expectation operator E_M,

$$\int_{\mathbb{R}^n} f(P(v_r))\varphi dx = \int_{\mathbb{R}^n} f(P(v_r))E_M(\varphi)dx + \int_{\mathbb{R}^n} f(P(v_r))(\varphi - E_M(\varphi))dx. \tag{1.30}$$

Clearly the conditional expectation E_M satisfies

$$\int_{\mathbb{R}^n} f(P(v_r))E_M(\varphi)dx = \int_{\mathbb{R}^n} E_M(f(P(v_r)))E_M(\varphi)dx.$$

Now we may invoke (1.19), Jensen's inequality on the range of P. This gives,

$$\int_{\mathbb{R}^n} E_M\left(f(P(v_r))\right)E_M(\varphi)dx \geq \int_{\mathbb{R}^n} f\left(E_M(P(v_r)\right)E_M(\varphi)dx$$

Hence adding and subtracting $f(0)$ to the leading term in the right hand side of (1.30) gives

$$\int_{\mathbb{R}^n} f(P(v_r))E_M(\varphi)dx \geq \int_{\mathbb{R}^n} f(0)E_M(\varphi)dx + \int_{\mathbb{R}^n} \left(f\left(E_M(P(v_r)\right) - f(0)\right)E_M(\varphi)dx. \tag{1.31}$$

It remains to specify how the above estimates are to be combined: Given $\epsilon > 0$ choose M large enough so that

$$|\varphi - E_M\varphi| \le \epsilon.$$

Next, depending on M, and ϵ select $r_0 \in \mathbb{N}$ so that for $r \ge r_0$,

$$|E_M(P(v_r))| \le \epsilon \quad \text{and} \quad \|v_r - P(v_r)\|_p \le \epsilon.$$

Combining now (1.28) – (1.31) with our choice of M and r we get

$$\int_{\mathbb{R}^n} f(v_r)\varphi dx \ge \int_{\mathbb{R}^n} f(0)\varphi dx - C\epsilon.$$

It remains to show how to remove the additional restriction (1.25). In view of the Lipschitz condition (1.28) it suffices to prove the theorem for those weak-limits v that are contained in a suitable dense set D where dense refers to the L^p_{loc} topology. We take

$$D = \{v \in L^p(\mathbb{R}^n, \mathbb{R}^n) : v \quad \text{is a finite sum of Haar functions}\}.$$

Let $v \in D$. Since the estimate (1.17) is invariant under dilations $x \to \lambda x$ it suffices to consider the case

$$v(x) = \sum_{k\in\mathbb{Z}^n} b_k 1_{k+(0,1)^n}(x), \tag{1.32}$$

and only finitely many of the b_k are different from zero.

Let $\eta \in C_0^+((0,1)^n)$ and extend η to a $(0,1)^n$ periodic continous function on $\mathbb{R}^n$. Since we proved (1.17) already under the restriction (1.25) we obtain for functions v satisfying (1.32) and $\varphi \in C_0^+(\mathbb{R}^n)$ that

$$\liminf_{r\to\infty} \int_{\mathbb{R}^n} f(v_r(x))(\varphi\cdot\eta)(x)dx \ge \int_{\mathbb{R}^n} f(v(x))(\varphi\cdot\eta)(x)dx. \tag{1.33}$$

Finally we remove η from the estimate (1.33). To this end let $\eta_k \in C_0^+((0,1)^n)$ be a sequence that converges pointwise to $1_{[0,1]^n}$ and extend each η_k periodically. Then for each k by (1.33)

$$\begin{aligned}\liminf_{r\to\infty} \int_{\mathbb{R}^n} f(v_r(x))\varphi(x)dx &\ge \liminf_{r\to\infty} \int_{\mathbb{R}^n} f(v_r(x))(\varphi\cdot\eta_k)(x)dx \\ &\ge \int_{\mathbb{R}^n} f(v(x))(\varphi\cdot\eta_k)(x)dx.\end{aligned} \tag{1.34}$$

Apply now the monotone convergence theorem to conclude that (1.17) holds true.

■

2 Multiscale Analysis of directional Haar Projections

In this section we outline the proof of Theorem 1.1. We start by performing a multiscale analysis of $P^{(\varepsilon)}$ with the purpose of successively resolving the discontinuities of the Haar system. We expand $P^{(\varepsilon)}$ in a series of operators, where each summand corresponds to a dyadic length scale. Thereafter we state the estimates of Theorem 2.1 and Theorem 2.2 that quantify the interplay between the resolving operators and the inverse of the Riesz transform R_{i_0}. Finally we show how the assertions of Theorem 1.1 follow.

Recall that $\mathcal{A} = \{\varepsilon \in \{0,1\}^n : \varepsilon \neq (0,\dots,0)\}$. We decompose the projection $P^{(\varepsilon)}, \varepsilon \in \mathcal{A}$, using a smooth compactly supported approximation of unity. To this end we choose $b \in C^\infty(\mathbb{R})$, supported in $[-1,1]$, so that for $t \in \mathbb{R}$,

$$b(t) = b(-t),\ 0 \leq b(t) \leq 4,\ \mathrm{Lip}(b) \leq 8,\ \text{and}\ \int_{-1}^{+1} b(t)dt = 1.$$

Let

$$d(x) = b(x_1)\cdot \dots \cdot b(x_n) - 2^n b(2x_1)\cdot \dots \cdot b(2x_n), \quad x = (x_1,\dots,x_n).$$

Since b was chosen to be even around 0, we have $\int_{-1}^{+1} tb(t)dt = 0$ hence also

$$\int_{\mathbb{R}} d(x_1,\dots,x_i,\dots,x_n)x_i dx_i = 0, \qquad (1 \leq i \leq n). \tag{2.1}$$

Let Δ_ℓ, $\ell \in \mathbb{Z}$ be the self adjoint operator defined by convolution as

$$\Delta_\ell(u) = u * d_\ell, \quad \text{where} \quad d_\ell(x) = d(2^\ell x)2^{n\ell}. \tag{2.2}$$

For $u \in L^p(\mathbb{R}^n)$ we get $u = \sum_{\ell=-\infty}^{\infty} \Delta_\ell(u)$. Convergence holds almost everywhere and in $L^p(\mathbb{R}^n)$. Recall that $\mathcal{S}$ denotes the collection of all dyadic cubes in $\mathbb{R}^n$. Let $j \in \mathbb{Z}$ and put

$$\mathcal{S}_j = \{Q \in \mathcal{S} : |Q| = 2^{-nj}\}. \tag{2.3}$$

Let $\ell \in \mathbb{Z}$, $\varepsilon \in \mathcal{A}$, define $T_\ell^{(\varepsilon)}$ as

$$T_\ell^{(\varepsilon)}(u) = \sum_{j=-\infty}^{\infty} \sum_{Q \in \mathcal{S}_j} \langle u, \Delta_{j+\ell}(h_Q^{(\varepsilon)})\rangle h_Q^{(\varepsilon)} |Q|^{-1}.$$

Since the operators $\Delta_{j+\ell}$ are self adjoint,

$$P^{(\varepsilon)}(u) = \sum_{\ell=-\infty}^{\infty} T_\ell^{(\varepsilon)}(u).$$

Let $1 \leq i_0 \leq n$. Recall that $\mathcal{A}_{i_0} = \{\varepsilon \in \mathcal{A} : \varepsilon = (\varepsilon_1,\dots,\varepsilon_n) \quad \text{and} \quad \varepsilon_{i_0} = 1\}$. Let $\epsilon \in \mathcal{A}_{i_0}$. In Section 3 we verify that

$$T_\ell^{(\varepsilon)} R_{i_0}^{-1} = T_\ell^{(\varepsilon)} R_{i_0} + \sum_{\substack{i=1 \\ i \neq i_0}}^{n} T_\ell^{(\varepsilon)} \mathbb{E}_{i_0} \partial_i R_i,$$

where R_i denotes the $i-$th Riesz transform, ∂_i denotes the differentiation with respect to the x_i variable and $\mathbb{E}_{i_0}$ the integration with respect to the $x_{i_0} - th$ coordinate,

$$\mathbb{E}_{i_0}(f)(x) = \int_{-\infty}^{x_{i_0}} f(x_1,\dots,s,\dots,x_n)ds, \qquad x = (x_1,\dots,x_n).$$

The following two theorems record the norm estimates for the operators $T_\ell^{(\varepsilon)}$ and $T_\ell^{(\varepsilon)} R_{i_0}^{-1}$ by which we obtain the upper bounds for $P^{(\varepsilon)}(u)$ stated in Theorem 1.1.

Theorem 2.1 *Let $1 < p < \infty$ and $1/p + 1/q = 1$ and $\ell \geq 0$. For $\varepsilon \in \mathcal{A}$ the operator $T_\ell^{(\varepsilon)}$ satisfies the norm estimates,*

$$\|T_\ell^{(\varepsilon)}\|_p \leq \begin{cases} C_p 2^{-\ell/2} & \text{if} \quad p \geq 2; \\ C_p 2^{-\ell/q} & \text{if} \quad p \leq 2. \end{cases} \tag{2.4}$$

Let $1 \leq i_0 \leq n$, and $\varepsilon \in \mathcal{A}_{i_0}$ then

$$\|T_\ell^{(\varepsilon)} R_{i_0}^{-1}\|_p \leq \begin{cases} C_p 2^{+\ell/2} & \text{if} \quad p \geq 2; \\ C_p 2^{+\ell/p} & \text{if} \quad p \leq 2. \end{cases} \tag{2.5}$$

Theorem 2.2 *Let $1 < p < \infty$. Let $\ell \leq 0$. Then for $\varepsilon \in \mathcal{A}$ the operator $T_\ell^{(\varepsilon)}$ satisfies the norm estimates,*

$$\|T_\ell^{(\varepsilon)}\|_p \leq \begin{cases} C_p 2^{-|\ell|/p} & \text{if} \quad p \geq 2; \\ C_p 2^{-|\ell|} & \text{if} \quad p \leq 2. \end{cases} \tag{2.6}$$

If moreover $1 \leq i_0 \leq n$, and $\varepsilon \in \mathcal{A}_{i_0}$, then

$$\|T_\ell^{(\varepsilon)} R_{i_0}^{-1}\|_p \leq \begin{cases} C_p 2^{-|\ell|/p} & \text{if} \quad p \geq 2; \\ C_p 2^{-|\ell|} & \text{if} \quad p \leq 2. \end{cases} \tag{2.7}$$

We show how Theorem 2.1 and Theorem 2.2 yield the proof of Theorem 1.1.

Proof of Theorem 1.1. Let $1 \leq i_0 \leq n$. Define $M \in \mathbb{N}$ by the relation

$$2^{M-1} \leq \frac{||u||_p ||R_{i_0}||_p}{||R_{i_0}(u)||_p} \leq 2^M. \tag{2.8}$$

Consider first $p \geq 2$. Let $\varepsilon \in \mathcal{A}_{i_0}$. Theorem 2.1 and Theorem 2.2 imply that

$$\sum_{\ell=M}^{\infty} ||T_\ell^{(\varepsilon)}||_p \leq C_p 2^{-M/2} \quad \text{and} \quad \sum_{\ell=-\infty}^{M-1} ||T_\ell^{(\varepsilon)} R_{i_0}^{-1}||_p \leq C_p 2^{M/2}.$$

Since $P^{(\varepsilon)}(u) = \sum_{\ell=-\infty}^{\infty} T_\ell^{(\varepsilon)}(u)$ triangle inequality gives that

$$\begin{aligned} \|P^{(\varepsilon)}(u)\|_p &\leq \sum_{\ell=M}^{\infty} ||T_\ell^{(\varepsilon)}\|_p \|u||_p + \sum_{\ell=-\infty}^{M-1} ||T_\ell^{(\varepsilon)} R_{i_0}^{-1}||_p \, ||R_{i_0}(u)||_p \\ &\leq C_p 2^{-M/2} ||u||_p + C_p 2^{M/2} ||R_{i_0}(u)||_p. \end{aligned} \tag{2.9}$$

Inserting the value of M specified in (2.8) gives

$$C_p 2^{-M/2} ||u||_p + C_p 2^{M/2} ||R_{i_0}(u)||_p \leq C_p ||u||_p^{1/2} ||R_{i_0}(u)||_p^{1-1/2}.$$

Assume next that $p \leq 2$. Let q be the Hölder conjugate index to p so that $1/p + 1/q = 1$. By Theorem 2.1 and Theorem 2.2, for $\varepsilon \in \mathcal{A}_{i_0}$,

$$\sum_{\ell=M}^{\infty} ||T_\ell^{(\varepsilon)}||_p \leq C_p 2^{-M/q} \quad \text{and} \quad \sum_{\ell=-\infty}^{M-1} ||T_\ell^{(\varepsilon)} R_{i_0}^{-1}||_p \leq C_p 2^{M/p}.$$

Triangle inequality applied to $P^{(\varepsilon)}(u) = \sum_{\ell=-\infty}^{\infty} T_\ell^{(\varepsilon)} u$ gives

$$\begin{aligned} \|P^{(\varepsilon)}(u)\|_p &\le \sum_{\ell=M}^{\infty} ||T_\ell^{(\varepsilon)}\|_p \|u||_p + \sum_{\ell=-\infty}^{M-1} ||T_\ell^{(\varepsilon)} R_{i_0}^{-1}||_p \, ||R_{i_0}(u)||_p \\ &\le C_p 2^{-M/q} ||u||_p + C_p 2^{M/p} ||R_{i_0}(u)||_p. \end{aligned} \tag{2.10}$$

With M defined as in (2.8) above we obtain

$$C_p 2^{-M/q} ||u||_p + C_p 2^{M/p} ||R_{i_0} u||_p \le C_p ||u||_p^{1/p} ||R_{i_0} u||_p^{1-1/p}.$$

■

3 Tooling up

In this section we prepare the tools provided by the Calderon Zygmund School of Harmonic Analysis. They simplify our tasks and save the reader time and effort. We exploit the Haar system indexed by (and supported on) dyadic cubes, its unconditionality in $L^p (1 < p < \infty)$, projections onto block bases of the Haar system, the connection of singular integral operators to wavelet systems, and interpolation theorems for operators on dyadic H^1 and dyadic BMO.

The Haar system in $\mathbb{R}^n$. We base this review on the work of T. Figiel [4] and Z. Ciesielski [2]. Denote by $\mathcal{D}$ the collection of all dyadic interval in the real line $\mathbb{R}$, and let $\{h_I : I \in \mathcal{D}\}$ be the associated L^∞ normalized Haar system. It forms a complete orthogonal system in $L^2(\mathbb{R})$. Analogs of the Haar system in the multi-dimensional case were developed by Z. Ciesielski in [2]. For our purposes the mere tensor products of the one dimensional Haar system is not quite sufficient. Instead we employ the Haar system supported on dyadic cubes.

Recall that $\mathcal{S}$ denotes the collection of dyadic cubes in $\mathbb{R}^n$. and that $\mathcal{A} = \{\varepsilon \in \{0,1\}^n : \varepsilon \ne (0,\dots.0)\}$. The system

$$\{h_Q^{(\varepsilon)} : Q \in \mathcal{S}, \varepsilon \in \mathcal{A}\}$$

is a complete orthogonal system in $L^2(\mathbb{R}^n)$ with $\|h_Q^{(\varepsilon)}\|_2^2 = |Q|$. It is also an unconditional basis in $L^p(\mathbb{R}^n)$ $(1 < p < \infty)$. Given $f \in L^p(\mathbb{R}^n)$ define its dyadic square function $\mathbb{S}(f)$ as

$$\mathbb{S}^2(f) = \sum_{\varepsilon\in\mathcal{A},\, Q\in\mathcal{S}} \langle f, h_Q^{(\varepsilon)}\rangle^2 1_Q |Q|^{-2} \tag{3.1}$$

The norm of $f \in L^p(\mathbb{R}^n)$ and that of its square function $\mathbb{S}(f)$ are related by the estimate

$$C_p^{-1} \|f\|_{L^p(\mathbb{R}^n)} \le \|\mathbb{S}(f)\|_{L^p(\mathbb{R}^n)} \le C_p \|f\|_{L^p(\mathbb{R}^n)}, \tag{3.2}$$

where $C_p \le Cp^2/(p-1)$. Repeatedly we exploit the unconditionality of the Haar system in the following form. Let $\{c_Q^{(\varepsilon)} : Q \in \mathcal{S}, \varepsilon \in \mathcal{A}\}$ be a bounded set of coefficients and $f \in L^p(\mathbb{R}^n)$. Then

$$g = \sum_{\varepsilon\in\mathcal{A},\, Q\in\mathcal{S}} c_Q^{(\varepsilon)} \langle f, h_Q^{(\varepsilon)}\rangle h_Q^{(\varepsilon)} |Q|^{-1},$$

satisfies the square function estimate $\mathbb{S}(g) \le \left(\sup |c_Q^{(\varepsilon)}|\right) \mathbb{S}(f)$, hence by (3.2)

$$\|g\|_{L^p(\mathbb{R}^n)} \le C_p \left(\sup |c_Q^{(\varepsilon)}|\right) \cdot \|f\|_{L^p(\mathbb{R}^n)}. \tag{3.3}$$

Wavelet systems. We refer to Y. Meyer and R. Coifman [7] for the unconditionality of the wavelet systems and the fact that they are equivalent to the Haar system. Recall that $\mathcal{S}$ denotes the collection of dyadic cubes in $\mathbb{R}^n$. We say that

$$\{\psi_Q^{(\varepsilon)} : Q \in \mathcal{S}, \varepsilon \in \mathcal{A}\}$$

is a wavelet system if $\{\psi_Q^{(\varepsilon)}/\sqrt{|Q|} : Q \in \mathcal{S}, \varepsilon \in \mathcal{A}\}$ is an orthonormal basis in $L^2(\mathbb{R}^n)$ satisfying $\int \psi_Q^{(\varepsilon)} = 0$ and there exists $C > 0$ so that for $Q \in \mathcal{S}$, and $\varepsilon \in \mathcal{A}$ the following structure condition holds,

$$\operatorname{supp} \psi_Q^{(\varepsilon)} \subseteq C \cdot Q, \qquad |\psi_Q^{(\varepsilon)}| \leq C, \qquad \operatorname{Lip}(\psi_Q^{(\varepsilon)}) \leq C \operatorname{diam}(Q)^{-1}. \tag{3.4}$$

The wavelet system $\{\psi_Q^{(\varepsilon)} : Q \in \mathcal{S}, \varepsilon \in \mathcal{A}\}$ is an unconditional basis in $L^p(\mathbb{R}^n)$ $(1 < p < \infty)$ and equivalent to the Haar system $\{h_Q^{(\varepsilon)} : Q \in \mathcal{S}, \varepsilon \in \mathcal{A}\}$: Indeed there exists $C_p \leq Cp^2/(p-1)$, so that for any choice of finite sums,

$$f = \sum_{\varepsilon \in \mathcal{A},\, Q \in \mathcal{S}} a_Q^{(\varepsilon)} h_Q^{(\varepsilon)} \quad \text{and} \quad g = \sum_{\varepsilon \in \mathcal{A},\, Q \in \mathcal{S}} a_Q^{(\varepsilon)} \psi_Q^{(\varepsilon)},$$

the following norm estimates hold,

$$C_p^{-1} \|f\|_{L^p(\mathbb{R}^n)} \leq \|g\|_{L^p(\mathbb{R}^n)} \leq C_p \|f\|_{L^p(\mathbb{R}^n)} . \tag{3.5}$$

Notational convention. Given a dyadic cube $Q \in \mathcal{S}$ we write h_Q as shorthand for any of the functions

$$h_Q^{(\varepsilon)},\ \varepsilon \in \mathcal{A}. \tag{3.6}$$

If a statement in this paper involves h_Q where $Q \in \mathcal{S}$ then that statement is meant to hold true with h_Q replaced by any of the functions $h_Q^{(\varepsilon)}, \varepsilon \in \mathcal{A}$.

Square function estimates and integral operators. In this (and the following) paragraph we isolate a class of integral operators for which boundedness in $L^p(\mathbb{R}^n)$ $(1 < p < \infty)$ can be obtained directly from the unconditionality of the Haar system. (Naturally we discuss those operators here because they will appear in later sections.) Let $\{c_Q, Q \in \mathcal{S}\}$ be a set of bounded coefficients where (for convenience) only finitely many of them are $\neq 0$. Let $u \in L^p(\mathbb{R}^n)$. Then

$$K(u)(x) = \int_{\mathbb{R}^n} k(x,y)u(y)dy \quad \text{with kernel} \quad k(x,y) = \sum_{Q \in \mathcal{S}} c_Q h_Q(x) h_Q(y) |Q|^{-1}, \tag{3.7}$$

satisfies the square function estimate $\mathbb{S}(K(u)) \leq \left(\sup |c_Q|\right) \mathbb{S}(u)$. Hence by (3.3),

$$\|K(u)\|_{L^p(\mathbb{R}^n)} \leq C_p \left(\sup |c_Q|\right) \cdot \|u\|_{L^p(\mathbb{R}^n)}, \tag{3.8}$$

where $C_p \leq C \max\{p^2, p/(p-1)\}$.

Projections onto block bases. Our reference to projections onto block bases of the Haar system is [6] by P. W. Jones. Let $\mathcal{B}$ be a collection of dyadic cubes. For $Q \in \mathcal{B}$ let $\mathcal{U}(Q)$ denote a collection of pairwise disjoint dyadic cubes. We assume that the collections $\mathcal{U}(Q)$ are disjoint as Q ranges over the cubes in $\mathcal{B}$. More precisely we assume the following conditions throughout:

$$\text{If } W \in \mathcal{U}(Q),\ W' \in \mathcal{U}(Q'), \text{ and } Q \neq Q' \text{ then } W \neq W'. \tag{3.9}$$

$$\text{If } W, W' \in \mathcal{U}(Q) \text{ and } W \neq W' \text{ then } W \cap W' = \emptyset. \tag{3.10}$$

Consider the block bases

$$d_Q = \sum_{W \in \mathcal{U}(Q)} h_W, \quad Q \in \mathcal{B}.$$

Given scalars c_Q we are interested in the operator

$$K_1(u) = \sum_{Q \in \mathcal{B}} c_Q \langle u, h_Q \rangle d_Q |Q|^{-1} \tag{3.11}$$

that maps $\sum_{Q \in \mathcal{B}} a_Q h_Q$ to $\sum_{Q \in \mathcal{B}} a_Q c_Q d_Q$. Similarly, given a wavelet system $\{\psi_K\}$ as above and scalars b_W we consider the block bases

$$\tilde{\psi}_Q = \sum_{W \in \mathcal{U}(Q)} b_W \psi_W$$

and the operator

$$K_2(u) = \sum_{Q \in \mathcal{B}} c_Q \langle u, h_Q \rangle \tilde{\psi}_Q |Q|^{-1}.$$

We shall see below that K_2 can be controlled by K_1. To estimate $K_1(u)$ it is sometimes convenient to use a different collection of cubes as follows. Let $U(Q) = \bigcup_{Q \in \mathcal{U}(Q)} W$ denote the pointset covered by the collection $\mathcal{U}(Q)$. Suppose that there exist dyadic cubes $E_1(Q), \ldots, E_k(Q)$, where k may depend on Q, so that

$$U(Q) \subseteq E_1(Q) \cup \cdots \cup E_k(Q).$$

Assume that the collections $\{E_1(Q), \ldots, E_k(Q)\}$ are disjoint as Q ranges over the cubes in $\mathcal{B}$. Let

$$g_Q = \sum_{i=1}^{k} h_{E_i(Q)}, \quad Q \in \mathcal{B}, \tag{3.12}$$

put $\gamma = \sup |c_Q|$, and define the integral operator

$$K_0(u) = \gamma \sum_{Q \in \mathcal{B}} \langle u, h_Q \rangle g_Q |Q|^{-1}.$$

Our construction gives the square function estimate

$$\mathbb{S}(K_1(u)) \leq \mathbb{S}(K_0(u)),$$

hence $\|K_1(u)\|_p \leq C_p \|K_0(u)\|_p$. Consequently, $L^p - L^q$ duality gives the norm estimate

$$\|K_1^*\|_p \leq C_p \|K_0^*\|_p. \tag{3.13}$$

Note that the transposed operators K_1^* and K_0^* are given as,

$$K_1^*(u) = \sum_{Q \in \mathcal{B}} c_Q \langle u, d_Q \rangle h_Q |Q|^{-1} \quad \text{and} \quad K_0^*(u) = \gamma \sum_{Q \in \mathcal{B}} \langle u, g_Q \rangle h_Q |Q|^{-1}.$$

Exchanging Haar functions and wavelets. The equivalence of the wavelet system to the Haar basis allows us to write down further examples of L^p bounded integral operators. We use again the notational convention to write ψ_Q denoting any of the wavelet functions $\psi_Q^{(\varepsilon)}, \varepsilon \in \mathcal{A}$.

Assume that $\mathcal{U}(Q)$, $Q \in \mathcal{B}$ satisfies (3.9) and (3.10). Let $b_W, W \in \mathcal{U}(Q)$ be scalars, and assume that $|b_W| \leq B$. Recall that

$$K_2(u) = \sum_{Q\in\mathcal{B}} c_Q \langle u, h_Q\rangle \tilde{\psi}_Q |Q|^{-1}, \qquad \tilde{\psi}_Q = \sum_{W\in\mathcal{U}(Q)} b_W \psi_W,$$

and that K_1 was defined in (3.11). Since K_2 can be viewed as the composition of K_1 with the map $h_W \to b_W\psi_W$ it follows from (3.3) and (3.5) that

$$\|K_2(u)\|_{L^p(\mathbb{R}^n)} \leq C_p B \cdot \|K_1(u)\|_{L^p(\mathbb{R}^n)}. \tag{3.14}$$

Duality gives estimates for the transposed operator as,

$$\|K_2^*\|_p \leq C_p B \|K_1^*\|_p, \tag{3.15}$$

where

$$K_2^*(u) = \sum_{Q\in\mathcal{B}} c_Q \langle u, \tilde{\psi}_Q\rangle h_Q |Q|^{-1} \quad \text{and} \quad K_1^*(u) = \sum_{Q\in\mathcal{B}} c_Q \langle u, d_Q\rangle h_Q |Q|^{-1}. \tag{3.16}$$

Calderon Zygmund kernels. We use the book by Y. Meyer and R. Coifman [7] as our source for singular integral operators and their relation to wavelet systems. Let $\{k_Q : Q \in \mathcal{S}\}$ be a family of functions satisfying $\int k_Q = 0$ and these standard estimates: There exists $C > 0$ so that for $Q \in \mathcal{S}$,

$$\operatorname{supp} k_Q \subseteq C \cdot Q, \qquad |k_Q| \leq 1, \qquad \operatorname{Lip}(k_Q) \leq C \operatorname{diam}(Q)^{-1}. \tag{3.17}$$

Let $\{c_Q : Q \in \mathcal{S}\}$ be a bounded sequence of scalars. Assume for simplicity that only finitely many of the c_Q are different from zero. Then

$$k_3(x,y) = \sum c_Q \psi_Q(x) k_Q(y) |Q|^{-1},$$

defines a standard Calderon-Zygmund kernel (see [7]) so that

$$K_3(u)(x) = \int k_3(x,y) u(y) dy$$

satisfies the norm estimate

$$\|K_3(u)\|_p \leq C_p \sup |c_Q| \cdot \|u\|_p.$$

By (3.5), the operator

$$K_4(u)(x) = \int k_4(x,y) u(y) dy \quad \text{with kernel} \quad k_4(x,y) = \sum c_Q h_Q(x) k_Q(y) |Q|^{-1},$$

satisfies

$$\|K_4(u)\|_p \leq C_p \sup |c_Q| \cdot \|u\|_p. \tag{3.18}$$

We will apply (3.18) in the following specialized situation. Let W be a dyadic cube and let V be a cube in $\mathbb{R}^n$ (not necessarily dyadic) so that

$$V \supseteq C_1 \cdot W, \qquad |V| \le C_2|W|. \tag{3.19}$$

Let $Q \subseteq W$ be a dyadic cube. Since $\int k_Q = 0$ and supp $k_Q \subseteq V$ we have

$$\langle u, k_Q\rangle = \langle 1_V(u - m_V(u)), k_Q\rangle,$$

where $m_V(u) = \int_V u/|V|$. This yields the identity

$$\sum_{Q\subseteq W} \langle u, k_Q\rangle h_Q|Q|^{-1} = \sum_{Q\subseteq W} \langle 1_V(u - m_V(u)), k_Q\rangle h_Q|Q|^{-1}.$$

To the kernel $\sum_{Q\subseteq W} h_Q(x)k_Q(y)|Q|^{-1}$ we apply the estimate (3.18) with $p = 2$. Since the Haar system is orthogonal we obtain

$$\begin{aligned}\sum_{Q\subseteq W} \langle u, k_Q\rangle^2|Q|^{-1} &= \|\sum_{Q\subseteq W} \langle u, k_Q\rangle h_Q|Q|^{-1}\|_2^2 \\ &= \|\sum_{Q\subseteq W} \langle 1_V(u - m_V(u))h_Q\rangle|Q|^{-1}\|_2^2 \\ &\le \|1_V(u - m_V(u))\|_2^2.\end{aligned} \tag{3.20}$$

With (3.20) we obtain BMO estimates for operators with Calderon Zygmund kernels as above.

The Riesz Transforms. We review basic facts about Riesz transforms and base the discussion on chapter III of [15] by E. M. Stein. Let $\mathcal{F}$ denote the Fourier transformation on $\mathbb{R}^n$. The Riesz transform R_i is a Fourier multiplier defined by

$$\mathcal{F}(R_i(u))(\xi) = -\sqrt{-1}\frac{\xi_i}{|\xi|}\mathcal{F}(u)(\xi) \quad \text{where} \quad 1 \le i \le n, \quad \xi = (\xi_1, \dots, \xi_n). \tag{3.21}$$

Riesz transforms, satisfy the estimates $\|R_j u\|_p \le C_p\|u\|_p$ $(1 < p < \infty)$, hence define bounded linear operators on the reflexive $L^p(\mathbb{R}^n)$ spaces. The defining relation (3.21) yields a convenient formula for the inverse of R_i, again by Fouriermultipliers. Consider for simplicity $i = 1$. Let u be a smooth and compactly supported test function such that $\mathcal{F}^{-1}(|\xi|/\xi_1\mathcal{F}(u)(\xi))$ is well defined. Then compute $\mathcal{F}(R_1^{-1}(u))(\xi)$ as

$$\begin{aligned}\mathcal{F}(R_1^{-1}(u))(\xi) &= -\sqrt{-1}\mathcal{F}(u)(\xi)\frac{|\xi|}{\xi_1} = \mathcal{F}(u)(\xi)\frac{-\sqrt{-1}}{\xi_1}\sum_{i=1}^n \frac{\xi_i^2}{|\xi|} \\ &= -\sqrt{-1}\mathcal{F}(u)(\xi)\left[\frac{\xi_1}{|\xi|} + \sum_{i=2}^n \frac{\xi_i}{\xi_1}\cdot\frac{\xi_i}{|\xi|}\right].\end{aligned}$$

Taking the inverse Fourier transform yields

$$R_1^{-1} = R_1 + \sum_{i=2}^n \mathbb{E}_1\partial_i R_i, \tag{3.22}$$

where $\mathbb{E}_1(f)(x_1,\dots,x_n)=\int_{-\infty}^{x_1} f(s,x_2,\dots,x_n)ds$ and ∂_i denotes the partial differentiation with respect to the $i-th$ coordinate.

Next fix $1\le i_0\le n$ and $\varepsilon\in\mathcal{A}_{i_0}$. After permuting the coordinates the above calculation gives the formula for $R_{i_0}^{-1}$ as follows

$$R_{i_0}^{-1}=R_{i_0}+\sum_{\substack{i=1\\ i\neq i_0}}^{n}\mathbb{E}_{i_0}\partial_i R_i. \tag{3.23}$$

Dyadic BMO, H^1_d **and Interpolation.** We use [1] by C. Bennett and R. Sharply as basic reference to interpolation theorems. Recall first the definition of dyadic BMO. Let $f\in L^2(\mathbb{R}^n)$ with Haar expansion given by (1.2) We say that f belongs to dyadic BMO and write $f\in \mathrm{BMO}_d$ if the norm defined by (3.24) is finite

$$\|f\|^2_{\mathrm{BMO}_d}=\left|\int f\right|^2+\sup_{Q\in\mathcal{S}}\frac{1}{|Q|}\sum_{\varepsilon\in\mathcal{A}}\sum_{W\subseteq Q}\left\langle f,h_W^{(\varepsilon)}\right\rangle^2|W|^{-1}. \tag{3.24}$$

Given a dyadic cube Q the system

$$\{1_Q\}\cup\{h_W^{(\varepsilon)}:W\in\mathcal{S},W\subseteq Q,\varepsilon\in\mathcal{A}\}$$

is a complete orthogonal system in the Hilbert space $L^2(Q,dt)$. This yields the identity

$$1_Q(f-m_Q(f))=\sum_{\varepsilon\in\mathcal{A}}\sum_{W\subseteq Q}\langle f,h_W^{(\varepsilon)}\rangle h_W^{(\varepsilon)}|W|^{-1},$$

where $m_Q(f)=(\int_Q f)/|Q|$. Hence the BMO_d norm of f can be rewritten as

$$\|f\|^2_{\mathrm{BMO}_d}=\left|\int f\right|^2+\sup_{Q}\int_Q|f(t)-m_Q(f)|^2\frac{dt}{|Q|}. \tag{3.25}$$

Given $f\in\mathrm{BMO}_d$ with $\int f=0$. Let $\mathcal{G}=\{W\in\mathcal{S}:\exists\varepsilon\ \langle f,h_W^{(\varepsilon)}\rangle\neq 0\}$. It is well known that in order to evaluate the BMO_d norm of f it suffices to consider the cubes in $\mathcal{G}$. Put

$$A_0\ =\sup_{Q\in\mathcal{G}}\frac{1}{|Q|}\sum_{\varepsilon\in\mathcal{A}}\sum_{W\subseteq Q}\langle f,h_W^{(\varepsilon)}\rangle^2|W|^{-1}.$$

We claim that

$$A_0=\|f\|^2_{\mathrm{BMO}_d}. \tag{3.26}$$

It suffices to observe that $A_0\ge\|f\|^2_{\mathrm{BMO}_d}$, since $A_0\le\|f\|^2_{\mathrm{BMO}_d}$, by definition. To this end we fix a dyadic cube $K\in\mathcal{S}$ so that $K\notin\mathcal{G}$. Let $\mathcal{M}\subseteq\mathcal{G}$ denote the collection of maximal cubes of $\mathcal{G}$ that are contained in K. (Maximality is with respect to inclusion.) Thus $\mathcal{M}$ consists of pairwise disjoint dyadic cubes,

$$\sum_{Q\in\mathcal{M}}|Q|\le|K|,$$

and,

$$\sum_{\varepsilon\in\mathcal{A}}\sum_{W\subseteq K}\langle f,h_W^{(\varepsilon)}\rangle^2|W|^{-1}=\sum_{Q\in\mathcal{M}}\sum_{\varepsilon\in\mathcal{A}}\sum_{W\subseteq Q}\langle f,h_W^{(\varepsilon)}\rangle^2|W|^{-1}.$$

Since $\mathcal{M} \subseteq \mathcal{G}$, for $Q \in \mathcal{M}$,

$$\sum_{\varepsilon \in \mathcal{A}} \sum_{W \subseteq Q} \langle f, h_W^{(\varepsilon)} \rangle^2 |W|^{-1} \leq A_0 |Q|.$$

Consequently we have the following estimates

$$\begin{aligned} \sum_{\varepsilon \in \mathcal{A}} \sum_{W \subseteq K} \langle f, h_W^{(\varepsilon)} \rangle^2 |W|^{-1} &= A_0 \sum_{Q \in \mathcal{M}} |Q| \\ &= A_0 |K|. \end{aligned}$$

Taking the supremum over all such K implies that $A_0 \geq \|f\|^2_{\mathrm{BMO}_d}$.

We review the definition of dyadic H^1, its relation to the scale of L^p spaces and to BMO_d. Let K be a dyadic cube in $\mathbb{R}^n$. We say that $a : \mathbb{R}^n \to \mathbb{R}$ is a dyadic atom if

$$\|a\|_{L^2(\mathbb{R}^n)} \leq |K|^{-1/2}, \quad \text{supp } a \subseteq K, \quad \text{and} \quad \int a = 0. \tag{3.27}$$

By definition a function $f \in L^1(\mathbb{R}^n)$ belongs to dyadic H^1 if there exists a sequence of dyadic atoms $\{a_i\}$ and a sequence of scalars $\{\lambda_i\}$ so that

$$f = \sum \lambda_i a_i \quad \text{and} \quad \sum |\lambda_i| < \infty. \tag{3.28}$$

We denote

$$\|f\|_{H^1_d} = \inf\{\sum |\lambda_i|\} \tag{3.29}$$

where the infimum is extended over all representations (3.28). For the resulting space of functions we write H^1_d. Recall also that the dual Banach space to H^1_d is identifiable with BMO_d.

Interpolation of operators links the spaces H^1_d, BMO_d on the one hand and the scale of L^p spaces on the other hand. Assume that T is a bounded operator on H^1_d and on L^2. Let A_1 denote the the norm of T on H^1_d and let A_2 denote the norm of T on L^2. Then for $1 < p < 2$ and $\theta = 2 - 2/p$

$$\|T\|_p \leq C A_1^{1-\theta} A_2^{\theta}.$$

If on the other hand the operator T is bounded on BMO_d with norm equal to A_∞ then for $2 < p < \infty$ and $\theta = 2/p$

$$\|T\|_p \leq C A_\infty^{1-\theta} A_2^{\theta}.$$

In addition to dyadic BMO at one point of the proof we employ the continuous analog of BMO_d. Let $f \in L^2(\mathbb{R}^n)$. Let $W \subseteq \mathbb{R}^n$ be a cube (not necessarily dyadic). Write

$$m_W(f) = \int_W f(t) \frac{dt}{|W|}.$$

We say that $f \in \mathrm{BMO}(\mathbb{R}^n)$ if

$$\|f\|^2_{\mathrm{BMO}(\mathbb{R}^n)} = \left| \int f \right|^2 + \sup_W \int_W |f(t) - m_W(f)|^2 \frac{dt}{|W|} < \infty,$$

where the supremum is extended over all cubes $W \subseteq \mathbb{R}^n$ (not just dyadic ones). Clearly for a given function $\|f\|_{\mathrm{BMO}(\mathbb{R}^n)} \geq \|f\|_{\mathrm{BMO}_d}$. In Section 4 we use $\mathrm{BMO}(\mathbb{R}^n)$ and interpolation as follows. Let $T : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ and $T : \mathrm{BMO}_d \to \mathrm{BMO}(\mathbb{R}^n)$ be bounded. Let A_2 be the operator norm of T on $L^2(\mathbb{R}^n)$ and put

$$A_\infty = \|T : \mathrm{BMO}(\mathbb{R}^n) \to \mathrm{BMO}_d\|.$$

Then for $1 < p < \infty$ and $\theta = 2/p$,

$$\|T\|_p \leq C A_\infty^{1-\theta} A_2^{\theta}.$$

4 Basic Dyadic Operations

The norm estimates for the operators $T_\ell^{(\varepsilon)}$ reflect boundedness of two basic dyadic operations. These are rearrangement operators of the Haar basis and averaging projections onto block bases of the Haar system. In this section we isolate the basic dyadic models and prove estimates in the spaces H^1, L^2 and BMO. In later sections the boundedness properties of $T_\ell^{(\varepsilon)}, \ell \le 0$, are reduced to the case of rearrangement operators. The estimates for $T_\ell^{(\varepsilon)}, \ell \ge 0$, are harder and involve rearrangements as well as orthogonal projections onto certain ring domains, surrounding the discontinuity set of Haar functions.

4.1 Projections and Ring Domains

The following definitions enter in the construction of the orthogonal projection (4.5). Recall the set of directions $\mathcal{A} = \{\varepsilon \in \{0,1\}^n : \varepsilon \ne (0,\dots.0)\}$. Let $\mathcal{B}$ be a collection of dyadic cubes. For $Q \in \mathcal{B}$ and $\varepsilon \in \mathcal{A}$ let $D^{(\varepsilon)}(Q)$ denote the set of discontinuities of the Haar function $h_Q^{(\varepsilon)}$. Fix $\lambda \in \mathbb{N}$ and define

$$D_\lambda^{(\varepsilon)}(Q) = \{x \in \mathbb{R}^n : \ \operatorname{dist}(x, D^{(\varepsilon)}(Q)) \le C2^{-\lambda} \operatorname{diam}(Q)\}.$$

Thus $D_\lambda^{(\varepsilon)}(Q)$ is the set of points that have distance $\le C2^{-\lambda}\operatorname{diam}(Q)$ to the set of discontinuities of $h_Q^{(\varepsilon)}$. Let $k(Q) \le C2^{\lambda(n-1)}$ and let $E_1(Q), \dots, E_{k(Q)}(Q)$ be the collection of all dyadic cubes satisfying

$$\operatorname{diam}(E_k(Q)) = 2^{-\lambda}\operatorname{diam}(Q), \qquad E_k(Q) \cap D_\lambda^{(\varepsilon)}(Q) \ne \emptyset. \tag{4.1}$$

We *assume* throughout this chapter that $\mathcal{B}$ is such that the collections $\{E_1(Q), \dots, E_{k(Q)}(Q)\}$ are pairwise disjoint as Q ranges over $\mathcal{B}$.

Thus we defined a covering of $D_\lambda^{(\varepsilon)}(Q)$ with dyadic cubes $\{E_1(Q), \dots, E_{k(Q)}(Q)\}$ satisfying these conditions:

1. There holds the measure estimate
$$|E_1(Q) \cup \dots \cup E_{k(Q)}(Q)| \le C2^{-\lambda}|Q|. \tag{4.2}$$

2. Let $Q, Q_0 \in \mathcal{B}$, $k \le k(Q)$ and $k_0 \le k(Q_0)$.
$$\text{If } E_k(Q) \subset E_{k_0}(Q_0) \text{ then } Q \subset Q_0. \tag{4.3}$$

3. Let $Q, Q_0 \in \mathcal{B}$, $k \le k(Q)$, $k_0 \le k(Q_0)$ and $Q \subset Q_0$.
$$\text{If } E_k(Q) \cap E_{k_0}(Q_0) \ne \emptyset \text{ then } E_k(Q) \subset E_{k_0}(Q_0). \tag{4.4}$$

Note that our hypothesis (4.2)–(4.4) are modeled after Jones's compatibility condition in [6]. With $\mathcal{U}(Q) = \{E_1(Q), \dots, E_{k(Q)}(Q)\}$ we define the block bases as $g_Q = \sum_{E \in \mathcal{U}(Q)} h_E$. The associated projection operator is given by the equation

$$S(u) = \sum_{Q \in \mathcal{B}} \langle u, h_Q \rangle g_Q |Q|^{-1}. \tag{4.5}$$

Recall that h_Q is shorthand for any of the Haar functions $h_Q^{(\varepsilon)}$, where $\varepsilon \in \mathcal{A}$. Moreover, if a statement in this paper involves h_Q then that statement is meant to hold true with h_Q replaced by any of the functions $h_Q^{(\varepsilon)}$.

The norm estimates for the operator S are recorded in the next theorem. For its use in the later sections of this paper the relation between the spaces, on which the operator acts, and the dependence of the operator norm on the value of λ becomes crucial.

Theorem 4.1 *There exists $C_0 = C_0(C, n)$ so that the orthogonal projection given by (4.5) satisfies these estimates*

$$\|S\|_{H^1_d} \le C_0 2^{-\lambda/2}, \qquad \|S\|_2 \le C_0 2^{-\lambda/2}, \qquad \text{and} \qquad \|S\|_{BMO_d} \le C_0.$$

Proof. The proof splits canonically into three parts. The first part treats L^2, the second part H^1_d, and the last part the BMO_d estimate of the operator S.

Part 1. We start with L^2. Since $|E_1(Q) \cup \cdots \cup E_{k(Q)}(Q)| \le C_n 2^{-\lambda}|Q|$, we have $\|g_Q\|_2^2 \le C_n 2^{-\lambda}|Q|$. As we assume that the collections $\{E_1(Q), \dots, E_{k(Q)}(Q)\}$ are pairwise disjoint as Q ranges over $\mathcal{B}$, the induced block bases $\{g_Q : Q \in \mathcal{B}\}$ are orthogonal. Hence

$$\begin{aligned} \|S(u)\|_2^2 &= \sum_{Q\in\mathcal{B}} \langle u, h_Q\rangle^2 \|g_Q\|_2^2 |Q|^{-2} \\ &\le C 2^{-\lambda} \|u\|_2^2. \end{aligned} \tag{4.6}$$

Part 2. The H^1_d estimate. Let a be a dyadic atom supported on a dyadic cube K so that $\|a\|_2^2 \le |K|^{-1}$. If $\langle a, h_Q\rangle \ne 0$, then $Q \subseteq K$ and $\operatorname{supp} g_K \subseteq C \cdot K$. Hence

$$\operatorname{supp} S(a) \subseteq C \cdot K.$$

The L^2 estimate (4.6) gives $\|S(a)\|_2^2 \le C_n 2^{-\lambda}|K|^{-1}$. As $\operatorname{supp} S(a) \subseteq C \cdot K$, we obtain the H^1_d estimate, $\|S(a)\|_{H^1_d} \le 2^{-\lambda/2} C$.

Part 3. The BMO_d estimate. Define

$$\mathcal{G} = \bigcup_{Q\in\mathcal{B}} \{E_1(Q), \dots, E_{k(Q)}(Q)\}.$$

Given $u \in \mathrm{BMO}_d$, by (3.26), it is sufficient to test the BMO_d norm of $S(u)$ using only the cubes $K \in \mathcal{G}$. Indeed,

$$\|S(u)\|_{\mathrm{BMO}_d}^2 = \sup_{K\in\mathcal{G}} \frac{1}{|K|} \int_K |S(u) - \frac{1}{|K|}\int_K S(u)|^2.$$

Let $K \in \mathcal{G}$. Note that, $\frac{1}{|K|}\int_K |S(u) - \frac{1}{|K|}\int_K S(u)|^2$ coincides with

$$\sum_{Q\in\mathcal{B}} \left\langle u, \frac{h_Q}{|Q|}\right\rangle^2 \sum_{\{k : E_k(Q)\subseteq K\}} |E_k(Q)|. \tag{4.7}$$

Choose $Q_0 \in \mathcal{B}, k_0 \le k(Q_0)$ so that $K = E_{k_0}(Q_0)$. By (4.3), if $Q \in \mathcal{B}$ and $E_k(Q) \subseteq E_{k_0}(Q_0)$, then $Q \subseteq Q_0$ and if moreover $E_k(Q) \cap E_{k_0}(Q_0) \ne \emptyset$ then, by (4.4), $E_k(Q) \subseteq E_{k_0}(Q_0)$. Hence if $Q \subseteq Q_0$ then

$$\sum_{\{k : E_k(Q)\subseteq K\}} |E_k(Q)| = \int_K g_Q^2,$$

and (4.7) equals,

$$\sum_{Q\in\mathcal{B},\,Q\subseteq Q_0}\left\langle u,\frac{h_Q}{|Q|}\right\rangle^2\int_K g_Q^2. \tag{4.8}$$

To get estimates for (4.8) consider $s\in\mathbb{N}\cup\{0\}$ such that $s\le\lambda$. Split the (effective) index set in (4.8) into

$$\mathcal{H}_s=\left\{Q\in\mathcal{B}:\ Q\subseteq Q_0,\ \operatorname{diam}(Q)=2^{-s}\operatorname{diam}(Q_0),\ \int_K g_Q^2\neq 0\right\},\quad s\le\lambda,$$

and

$$\mathcal{H}_\infty=\left\{Q\in\mathcal{B}:\ Q\subseteq Q_0,\ \operatorname{diam}(Q)<2^{-\lambda}\operatorname{diam}(Q_0),\ \int_K g_Q^2\neq 0\right\}.$$

First estimate the contribution to (4.8) coming from $\mathcal{H}_\infty$. If $Q\in\mathcal{H}_\infty$ then by (4.2), $\int_K g_Q^2\le C2^{-\lambda}|Q|$. Since clearly the pointset covered by $\mathcal{H}_\infty$ is contained in $C\cdot K$, we get

$$\begin{aligned}\sum_{Q\in\mathcal{H}_\infty}\left\langle u,\frac{h_Q}{|Q|}\right\rangle^2\int_K d_Q^2 &\le C2^{-\lambda}\sum_{Q\in\mathcal{H}_\infty}\langle u,h_Q\rangle^2|Q|^{-1}\\ &\le C2^{-\lambda}\|u\|^2_{\mathrm{BMO}_d}|K|.\end{aligned} \tag{4.9}$$

Next turn to the $\mathcal{H}_s$, $s\le\lambda$. The analysis is parallel to the previous case. The cardinality of $\mathcal{H}_s$ is bounded by C_n with C_n independent of s or λ. For $Q\in\mathcal{H}_s$ we get $\int_K g_Q^2\le C2^{-s}|K|$. Hence

$$\sum_{Q\in\mathcal{H}_s}\left\langle u,\frac{h_Q}{|Q|}\right\rangle^2\int_K g_Q^2\le C2^{-s}\|u\|^2_{\mathrm{BMO}_d}|K|.$$

Taking the sum over $0\le s\le\lambda$, gives

$$\sum_{s=0}^{\lambda}\sum_{Q\in\mathcal{H}_s}\left\langle u,\frac{h_Q}{|Q|}\right\rangle^2\int_K g_Q^2\le C\|u\|^2_{\mathrm{BMO}_d}|K|. \tag{4.10}$$

Adding (4.9) and (4.10) gives the BMO_d estimate $\|S(u)\|_{\mathrm{BMO}_d}\le C\|u\|_{\mathrm{BMO}_d}$.

4.2 Rearrangement Operators

We next turn to defining the rearrangement operator S given by (4.12) below. Let $\lambda\in\mathbb{N}$ and let $Q\in\mathcal{S}$ be a dyadic cube. The $\lambda-th$ dyadic predecessor of Q, denoted $Q^{(\lambda)}$, is given by the relation

$$Q^{(\lambda)}\in\mathcal{S},\quad |Q^{(\lambda)}|=2^{n\lambda}|Q|,\quad Q\subset Q^{(\lambda)}.$$

Let $\tau:\mathcal{S}\to\mathcal{S}$ be the map that associates to each $Q\in\mathcal{S}$ its $\lambda-th$ dyadic predecessor. Thus

$$\tau(Q)=Q^{(\lambda)},\quad Q\in\mathcal{S}.$$

Clearly $\tau:\mathcal{S}\to\mathcal{S}$ is not injective. We canonically split $\mathcal{S}=\mathcal{Q}_1\cup\cdots\cup\mathcal{Q}_{2^{n\lambda}}$ such that the restriction of τ to each of the collections $\mathcal{Q}_k$, is injective: Given $Q\in\mathcal{S}$, form

$$\mathcal{U}(Q)=\left\{W\in\mathcal{S}:W^{(\lambda)}=Q\right\}.$$

Thus $\mathcal{U}(Q)$ is a covering of Q and contains exactly $2^{n\lambda}$ pairwise disjoint dyadic cubes. We enumerate them, rather arbitrarily, as $W_1(Q), \dots, W_{2^{n\lambda}}(Q)$. For $1 \le k \le 2^{n\lambda}$, define

$$\mathcal{Q}_k = \{W_k(Q) : Q \in \mathcal{S}\}\,.$$

Note that $\tau : \mathcal{Q}_k \to \mathcal{S}$ is a bijection, and

$$\tau(W_k(Q)) = Q, \qquad W_k(Q) \in \mathcal{Q}_k, \quad Q \in \mathcal{S}.$$

Let $1 \le k \le 2^{n\lambda}$. Let $\{\varphi_Q^{(k)} : Q \in \mathcal{S}\}$ be a family of functions for which $\int \varphi_Q^{(k)} = 0$ and which satisfy the following structural conditions: There exists $C > 0$ so that for each $Q \in \mathcal{S}$

$$\operatorname{supp}\,\varphi_Q^{(k)} \subseteq C \cdot Q, \qquad |\varphi_Q^{(k)}| \le C, \qquad \operatorname{Lip}(\varphi_Q^{(k)}) \le C\,\operatorname{diam}(Q)^{-1}. \tag{4.11}$$

We emphasize that the actual function $\varphi_Q^{(k)}$ may depend on k, by contrast the structural conditions (4.11) are independent of the value of k. Define the operator S by the equation

$$S(g) = \sum_{k=1}^{2^{n\lambda}} \sum_{Q \in \mathcal{Q}_k} \left\langle g, \varphi_{\tau(Q)}^{(k)} \right\rangle h_Q |Q|^{-1}. \tag{4.12}$$

The action of S is best understood by viewing it as the transposition of the rearrangement operator defined by τ followed by a Calderon Zygmund Integral. The next theorem records the operator norm of S, particularly its joint $(n, \lambda)-$dependence, on the spaces H^1_d, L^2 and BMO_d.

Theorem 4.2 *The operator S defined by (4.12) is bounded on the spaces H^1_d, L^2 and from BMO($\mathbb{R}^n$) to BMO$_d$. The norm estimates depend on the value of $\lambda \in \mathbb{N}$ and the dimension of the ambient space $\mathbb{R}^n$ as follows:*

$$\|S\|_2 \le C_0 2^{n\lambda}, \qquad \|S\|_{H^1_d} \le C_0 2^{n\lambda}, \qquad \|S : \, BMO(\mathbb{R}^n) \to BMO_d\,\| \le C_0 \lambda^{1/2} 2^{n\lambda}. \tag{4.13}$$

Proof. The three parts of the proof correspond to the three operator estimates in (4.13). The first part treats L^2, the second part H^1_d and the third part BMO_d.

Part 1. We start with L^2. Let $u \in L^2$. Then

$$\|S(u)\|_2^2 = \sum_{k=1}^{2^{n\lambda}} \sum_{Q \in \mathcal{Q}_k} \left\langle u, \varphi_{\tau(Q)}^{(k)} \right\rangle^2 |Q|^{-1}.$$

Let $1 \le k \le 2^{n\lambda}$. Since $\tau : \mathcal{Q}_k \to \mathcal{S}$ is bijective, the standard conditions (4.11) and the L^2 estimates for Calderon Zygmund operators (3.18) yield,

$$\sum_{Q \in \mathcal{Q}_k} \left\langle u, \varphi_{\tau(Q)}^{(k)} \right\rangle^2 |\tau(Q)|^{-1} \le C\|u\|_2^2. \tag{4.14}$$

Recall that $|\tau(Q)| = 2^{n\lambda}|Q|$. On the left hand side of (4.14) replace $|\tau(Q)|^{-1}$ by $2^{-n\lambda}|Q|^{-1}$ then take the sum over $1 \le k \le 2^{n\lambda}$. This gives

$$\sum_{k=1}^{2^{n\lambda}} \sum_{Q \in \mathcal{Q}_k} \left\langle u, \varphi_{\tau(Q)}^{(k)} \right\rangle^2 |Q|^{-1} \le C 2^{2n\lambda} \|u\|_2^2.$$

Hence $\|S\|_2 \le C_0 2^{n\lambda}$, as claimed.

Part 2. The H^1_d estimate. Let a be a dyadic atom supported on a dyadic cube K. Define

$$\mathcal{H} = \left\{ Q \in \mathcal{S} : \ \operatorname{diam}(\tau(Q)) \geq \ \operatorname{diam}(K), \ \langle a, \varphi^{(k)}_{\tau(Q)} \rangle \neq 0 \right\}.$$

Then put $S(a) = b_1 + b_2$ where

$$b_1 = \sum_{Q \in \mathcal{H}} \langle S(a), h_Q \rangle \, h_Q |Q|^{-1},$$

and $b_2 = S(a) - b_1$. We treat separately the norm of b_1 and b_2. First we estimate $\|b_1\|_{H^1_d}$. Fix $s \in \mathbb{N} \cup \{0\}$ and put

$$\mathcal{H}_s = \{ Q \in \mathcal{H} : \ \operatorname{diam}(\tau(Q)) = 2^s \operatorname{diam}(K) \}.$$

Let $Q \in \mathcal{Q}_k \cap \mathcal{H}_s$ and let $q \in Q$. As $\int a = 0$ we obtain

$$\begin{aligned} \left| \left\langle a, \varphi^{(k)}_{\tau(Q)} \right\rangle \right| &= \left| \left\langle a, \varphi^{(k)}_{\tau(Q)} - \varphi^{(k)}_{\tau(Q)}(q) \right\rangle \right| \\ &\leq C \|a\|_{L^1} \operatorname{diam}(Q) \operatorname{Lip}(\varphi^{(k)}_{\tau(Q)}) \end{aligned}$$

By the structural conditions (4.11), $Q \in \mathcal{Q}_k \cap \mathcal{H}_s$ implies $\operatorname{Lip}(\varphi^{(k)}_{\tau(Q)}) \leq C 2^{-s} \operatorname{diam}(K)^{-1}$. Hence $|\langle a, \varphi^{(k)}_{\tau(Q)} \rangle| \leq C 2^{-s}$. Note that the cardinality of $\mathcal{Q}_k \cap \mathcal{H}_s$ is bounded by an absolute constant C. Hence,

$$\sum_{s=0}^{\infty} \sum_{k=1}^{2^{n\lambda}} \sum_{Q \in \mathcal{Q}_k \cap \mathcal{H}_s} |\langle a, \varphi^{(k)}_{\tau(Q)} \rangle| \leq C 2^{n\lambda}. \tag{4.15}$$

Since $h_Q/|Q|$ is of norm one in H^1_d, the triangle inequality and (4.15) give $\|b_1\|_{H^1_d} \leq C 2^{n\lambda}$. It remains to consider $\|b_2\|_{H^1_d}$. Here the estimates are a direct consequence of the operator L^2 norm of S. First

$$\begin{aligned} \|b_2\|_2^2 &\leq \|S(a)\|_2 \\ &\leq C 2^{2n\lambda} \|a\|_2^2 \\ &\leq C 2^{2n\lambda} |K|. \end{aligned}$$

Second, a moments reflection shows that the Haar support of b_2 is contained in $C \cdot K$. Let

$$\mathcal{M} = \{ W \in \mathcal{S} : W \cap \operatorname{supp} b_2 \neq 0, \ |W| = |K| \}$$

Clearly the union of the cubes in $\mathcal{M}$ covers supp b_2. The cardinality of $\mathcal{M}$ is bounded by a constant C_n, and $\int_W b_2 = 0$ for $W \in \mathcal{M}$. Hence the functions

$$C^{-1} 2^{-n\lambda} 1_W b_2, \quad W \in \mathcal{M},$$

are dyadic atoms, and $\|b_2\|_{H^1_d} \leq C 2^{n\lambda}$. Since $\|S(a)\|_{H^1_d} \leq \|b_1\|_{H^1_d} + \|b_2\|_{H^1_d}$ it follows that $\|S\|_{H^1_d} \leq C_0 2^{n\lambda}$.

Part 3. Let $u \in \mathrm{BMO}(\mathbb{R}^n)$. We obtain the BMO_d estimate for $S(u)$ by verifying that for every dyadic cube W,

$$\sum_{k=1}^{2^{n\lambda}} \sum_{Q \in \mathcal{Q}_k, Q \subseteq W} \langle u, \varphi^{(k)}_{\tau(Q)} \rangle^2 |Q|^{-1} \leq C|W| \cdot \lambda \cdot 2^{2n\lambda} \cdot \|u\|^2_{\mathrm{BMO}(\mathbb{R}^n)}. \tag{4.16}$$

To this end fix a dyadic cube W. Split $\{Q \in \mathcal{S}, Q \subseteq W\} = \mathcal{G} \cup \mathcal{H}$, where

$$\mathcal{H} = \{Q \in \mathcal{S} : \ Q \subseteq W, \ \operatorname{diam}(Q) \geq \ \operatorname{diam}(W) 2^{-\lambda}\} \quad \text{and} \quad \mathcal{G} = \{Q \in \mathcal{S}, Q \subseteq W\} \setminus \mathcal{H}.$$

Fix $1 \leq k \leq 2^{n\lambda}$, put $\mathcal{G}_k = \mathcal{G} \cap \mathcal{Q}_k$ and observe that

$$\bigcup_{Q \in \mathcal{G}_k} \tau(Q) \subseteq W.$$

Recall further that $\tau : \mathcal{G}_k \to \mathcal{S}$ is injective. Hence the standard conditions (4.11), the Calderon-Zygmund estimate (3.20), and (3.19) yield

$$\sum_{Q \in \mathcal{G}_k} \langle u, \varphi^{(k)}_{\tau(Q)} \rangle^2 |\tau(Q)|^{-1} \leq C|W| \cdot \|u\|^2_{\mathrm{BMO}(\mathbb{R}^n)}. \tag{4.17}$$

Next replace $|\tau(Q)|^{-1}$ by $2^{-n\lambda}|Q|^{-1}$, then take the sum of (4.17) over $1 \leq k \leq 2^{n\lambda}$. We obtain that

$$\sum_{k=1}^{2^{n\lambda}} \sum_{Q \in \mathcal{G}_k} \langle u, \varphi^{(k)}_{\tau(Q)} \rangle^2 |Q|^{-1} \leq C|W| \cdot 2^{2n\lambda} \cdot \|u\|^2_{\mathrm{BMO}(\mathbb{R}^n)}.$$

We turn to estimating the contribution to (4.16) coming from $\mathcal{H}$. Let $0 \leq s \leq \lambda$. Write

$$\mathcal{H}_s = \{Q \in \mathcal{H} : \ \operatorname{diam}(Q) = 2^{-s} \operatorname{diam}(W)\}.$$

The cardinality of $\mathcal{H}_s$ equals 2^{ns}. It is useful to observe that, since $s \leq \lambda$, there exists exactly one dyadic cube K_s so that

$$\tau(Q) = K_s, \quad \text{for all} \quad Q \in \mathcal{H}_s.$$

Hence the following identity holds

$$\sum_{k=1}^{2^{n\lambda}} \sum_{Q \in \mathcal{H}_s \cap \mathcal{Q}_k} \langle u, \varphi^{(k)}_{\tau(Q)} \rangle^2 |Q|^{-1} = \langle u, \varphi^{(k)}_{K_s} \rangle^2 \left[\sum_{Q \in \mathcal{H}_s} |Q|^{-1} \right]. \tag{4.18}$$

Each $Q \in \mathcal{H}_s$ satisfies $|Q| = |W| 2^{-ns}$. As $\mathcal{H}_s$ has cardinality equal to 2^{ns}, it follows that

$$\sum_{Q \in \mathcal{H}_s} |Q|^{-1} = 2^{2ns} |W|^{-1}.$$

By definition $|K_s| = 2^{-ns+n\lambda}|W|$. Squaring and regrouping gives

$$2^{2ns}|W|^{-1} = 2^{2n\lambda}|K_s|^{-2}|W|.$$

Hence the right hand side of (4.18) equals

$$2^{2n\lambda}|W| \left\langle u, \varphi^{(k)}_{K_s} \right\rangle^2 |K_s|^{-2}. \tag{4.19}$$

By (4.11), $\|\varphi^{(k)}_{K_s}\|_2 \le |K_s|^{1/2}$. Let B_s be a cube in $\mathbb{R}^n$ so that supp $(\varphi^{(k)}_{K_s}) \subseteq B_s$ and diam$(B_s) \le C$ diam(K_s). Let $m_{B_s}(u) = \frac{1}{|B_s|}\int_{B_s} u(x)dx$. As $\int \varphi^{(k)}_{K_s} = 0$ we get

$$\begin{aligned} |\langle u, \varphi^{(k)}_{K_s}\rangle| &= |\langle u - m_{B_s}(u), \varphi^{(k)}_{K_s}\rangle| \\ &\le C\|1_{B_s} \cdot (u - m_{B_s}(u))\|_2 |K_s|^{1/2} \\ &\le C|K_s| \cdot \|u\|_{\mathrm{BMO}(\mathbb{R}^n)}. \end{aligned} \tag{4.20}$$

Inserting (4.20) into (4.19) gives that the latter is bounded by

$$C2^{2n\lambda}|W| \cdot \|u\|^2_{\mathrm{BMO}(\mathbb{R}^n)}. \tag{4.21}$$

Thus we showed that the left hand side of (4.18) equals (4.19) which in turn is bounded by (4.21). Hence

$$\sum_{k=1}^{2^{n\lambda}} \sum_{Q\in\mathcal{H}_s\cap\mathcal{Q}_k} \langle u, \varphi^{(k)}_{\tau(Q)}\rangle^2 |Q|^{-1} \le C2^{2n\lambda}|W| \cdot \|u\|^2_{\mathrm{BMO}(\mathbb{R}^n)}. \tag{4.22}$$

Finally in (4.22) we take the sum over $0 \le s \le \lambda$ and obtain (4.16)

■

5 The Proof of Theorem 2.1.

In this section we prove Theorem 2.1. The sub-sections 5.1 – 5.3 are devoted to the estimates for the operator $T^{(\varepsilon)}_\ell$, $\ell \ge 0$. In sub-section 5.4 we discuss the reduction of the estimates for $T^{(\varepsilon)}_\ell R^{-1}_{i_0}$, $\varepsilon \in \mathcal{A}_{i_0}$, to those of $T^{(\varepsilon)}_\ell$. Recall that

$$\mathcal{A}_{i_0} = \{\varepsilon \in \mathcal{A} : \varepsilon = (\varepsilon_1, \dots \varepsilon_n) \quad \text{and} \quad \varepsilon_{i_0} = 1\}.$$

Let $\varepsilon \in \mathcal{A}_{i_0}$. Let $\ell \ge 0$. Recall that for $j \in \mathbb{Z}$ we let $\mathcal{S}_j$ be the collection of all dyadic cubes in $\mathbb{R}^n$ with measure equal to 2^{-nj}. Let $Q \in \mathcal{S}_j$ and define

$$f^{(\varepsilon)}_{Q,\ell} = \Delta_{j+\ell}(h^{(\varepsilon)}_Q). \tag{5.1}$$

With the abbreviation (5.1) we have

$$T^{(\varepsilon)}_\ell(f) = \sum_{Q\in\mathcal{S}} \langle f, f^{(\varepsilon)}_{Q,\ell}\rangle h^{(\varepsilon)}_Q |Q|^{-1}. \tag{5.2}$$

The functions $f^{(\varepsilon)}_{Q,\ell}$ have vanishing mean and satisfy the basic estimates

$$\text{supp } f^{(\varepsilon)}_{Q,\ell} \subseteq D^{(\varepsilon)}_\ell(Q), \quad |f^{(\varepsilon)}_{Q,\ell}| \le C, \quad \mathrm{Lip}(f^{(\varepsilon)}_{Q,\ell}) \le C2^\ell(\text{ diam}(Q))^{-1}, \tag{5.3}$$

where $D_\ell^{(\varepsilon)}(Q)$ is the set of points that have distance $\le C2^{-\ell}$ diam(Q) to the set of discontinuities of $h_Q^{(\varepsilon)}$. *Based only* on the expansion (5.2) and the scale invariant conditions (5.3) we prove in the following subsections that $T_\ell^{(\varepsilon)}$, $\ell \ge 0$ satisfies the norm estimates

$$\|T_\ell^{(\varepsilon)}\|_p \le \begin{cases} C_p 2^{-\ell/2} & \text{for} \quad p \ge 2; \\ C_p 2^{-\ell/q} & \text{for} \quad p \le 2. \end{cases} \tag{5.4}$$

To this end we decompose the operator $T_\ell^{(\varepsilon)}$, $\ell \ge 0$ into a series of operators $T_{\ell,m}, m \in \mathbb{Z}$ using a wavelet system $\{\psi_K^{(\alpha)} : K \in \mathcal{S}, \alpha \in \mathcal{A}\}$ so that $\{\psi_K^{(\alpha)}/\sqrt{|K|}\}$ is an orthonormal basis in $L^2(\mathbb{R}^n)$, satisfying $\int \psi_K^{(\alpha)} = 0$ and the structure conditions,

$$\text{supp } \psi_K^{(\alpha)} \subseteq C \cdot K, \qquad |\psi_K^{(\alpha)}| \le C, \qquad \text{Lip}(\psi_K^{(\alpha)}) \le C \text{ diam}(K)^{-1}.$$

To simplify expressions below we suppress the superindeces (α) and, with a slight abuse of notation, in place of $\{\psi_K^{(\alpha)}\}$ we write just $\{\psi_K\}$. Then expanding a function f along the wavelet basis we get

$$f = \sum_{K \in \mathcal{S}} \left\langle f, \frac{\psi_K}{|K|} \right\rangle \psi_K.$$

Fix $m \in \mathbb{Z}$ and define $T_{\ell,m}$ by the equation

$$T_{\ell,m}(f) = \sum_{j=-\infty}^{\infty} \sum_{Q \in \mathcal{S}_j} \sum_{K \in \mathcal{S}_{j+\ell+m}} \left\langle f, \frac{\psi_K}{|K|} \right\rangle \langle \Delta_{j+\ell}(h_Q^{(\varepsilon)}), \psi_K \rangle h_Q^{(\varepsilon)} |Q|^{-1}. \tag{5.5}$$

Then

$$T_\ell^{(\varepsilon)}(f) = \sum_{m=-\infty}^{\infty} T_{\ell,m}(f). \tag{5.6}$$

In this section we prove that

$$\sum_{m=-\infty}^{-\ell-1} ||T_{\ell,m}||_p \le C_p 2^{-\ell}, \quad \text{and} \quad \sum_{m=-\ell}^{\infty} \|T_{\ell,m}\|_p \le \begin{cases} C_p 2^{-\ell/2} & \text{for} \quad p \ge 2; \\ C_p 2^{-\ell/q} & \text{for} \quad p \le 2. \end{cases} \tag{5.7}$$

The bounds of (5.7) imply the norm estimates for $T_\ell^{(\varepsilon)}$, $\ell \ge 0$ as stated in (5.4).

There are three relevant length scales in the series (5.5).

1. The scale 2^{-j}. This is the sidelength of $Q \in \mathcal{S}_j$, the cube under consideration.
2. The scale $2^{-(j+\ell)}$. This is the scale of $\Delta_{j+\ell}(h_Q^{(\varepsilon)})$. More precisely, since $\Delta_{j+\ell}$ is given by a convolution kernel of zero mean, the function $\Delta_{j+\ell}(h_Q^{(\varepsilon)})$ is supported in a strip of width proportional to $2^{-(j+\ell)}$ around the discontinuity set of $h_Q^{(\varepsilon)}$.
3. The scale $2^{-(j+\ell+m)}$. This is the scale of the test functions ψ_K, $K \in \mathcal{S}_{j+\ell+m}$.

The estimate (5.7) follows from Proposition 5.1, Proposition 5.2 and Proposition 5.3 below which deal with the regimes

1. $2^{-(j+\ell+m)} > 2^{-j}$,

2. $2^{-(j+\ell+m)} < 2^{-(j+\ell)}$,

3. $2^{-(j+\ell+m)} \in [2^{-(j+\ell)}, 2^{-j}]$,

respectively. Accordingly we treat separately the following three cases, $m > 0$, $0 \geq m \geq -\ell$, and $m < -\ell$.

5.1 Estimates for $T_{\ell,m}$, $\ell \geq 0$, $m < -\ell$.

In the case when $m < -\ell$ and $\ell \geq 0$ we have $2^{-(j+\ell+m)} > 2^{-j}$. Thus the length scale of the test function ψ_K is larger than the scale of $h_Q^{(\varepsilon)}$ when $Q \in \mathcal{S}_j$.

We obtain in Proposition 5.1 the estimates for $T_{\ell,m}$ from those of the rearrangement operators treated in the previous section, and from the fact that the wavelet bases in $L^p(1 < p < \infty)$ are equivalent to the Haar basis. The fruitful idea of exploiting rearrangements of the Haar system in the analysis of singular integral operators originates in T. Figiel's work [4]. (See also [9] for an exposition of T. Figiel's approach.)

Proposition 5.1 *Let $1 < p < \infty$ and $1/p + 1/q = 1$. For $\ell \geq 0$, and $m < -\ell$ the operator*

$$T_{\ell,m}(f) = \sum_{j=-\infty}^{\infty} \sum_{Q \in \mathcal{S}_j} \sum_{K \in \mathcal{S}_{j+\ell+m}} \left\langle f, \frac{\psi_K}{|K|} \right\rangle \langle \Delta_{j+\ell}(h_Q^{(\varepsilon)}), \psi_K \rangle h_Q^{(\varepsilon)} |Q|^{-1}$$

satisfies the norm estimate

$$\|T_{\ell,m}\|_p \leq \begin{cases} C_p 2^m \sqrt{-m-\ell} & \text{for} \quad p \geq 2; \\ C_p 2^m & \text{for} \quad p \leq 2. \end{cases} \tag{5.8}$$

and consequently

$$\sum_{m=-\infty}^{-\ell-1} ||T_{\ell,m}||_p \leq C_p 2^{-\ell}.$$

Proof. Fix $\ell \geq 0$ and $-\infty < m < -\ell$. Let $j \in \mathbb{Z}$ and fix a dyadic cube $Q \in \mathcal{S}_j$. Then form the collection of dyadic cubes

$$\mathcal{U}_{\ell,m}(Q) = \{K \in \mathcal{S}_{j+\ell+m} : \langle \psi_K, \Delta_{j+\ell}(h_Q^{(\varepsilon)}) \rangle \neq 0\}.$$

Clearly for $T_{\ell,m}(f)$ holds the identity

$$T_{\ell,m}(f) = \sum_{j=-\infty}^{\infty} \sum_{Q \in \mathcal{S}_j} \sum_{K \in \mathcal{U}_{\ell,m}(Q)} \left\langle f, \frac{\psi_K}{|K|} \right\rangle \langle \Delta_{j+\ell}(h_Q^{(\varepsilon)}), \psi_K \rangle h_Q^{(\varepsilon)} |Q|^{-1}. \tag{5.9}$$

Observe that for $-\infty < m < -\ell$ the cardinality of the collection $\mathcal{U}_{\ell,m}(Q)$ is uniformly bounded. Next for $K \in \mathcal{U}_{\ell,m}(Q)$ we prove that

$$|\langle \Delta_{j+\ell}(h_Q^{(\varepsilon)}), \psi_K \rangle| \leq C 2^m |Q|. \tag{5.10}$$

Since

$$\int_{\mathbb{R}^n} |\Delta_{j+\ell}(h_Q^{(\varepsilon)})| dx \leq C 2^{-\ell} |Q|,$$

and since $\Delta_{j+\ell}(h_Q^{(\varepsilon)})$ has vanishing mean, we get for $q \in Q$

$$
\begin{aligned}
|\langle \Delta_{j+\ell}(h_Q^{(\varepsilon)}), \psi_K \rangle| &= |\langle \Delta_{j+\ell}(h_Q^{(\varepsilon)}), (\psi_K - \psi_K(q)) \rangle| \\
&\leq C \mathrm{Lip}(\psi_K) \operatorname{diam}(Q) \int_{\mathbb{R}^n} |\Delta_{j+\ell}(h_Q^{(\varepsilon)})| dx \\
&\leq C \frac{\operatorname{diam}(Q)}{\operatorname{diam}(K)} 2^{-\ell} |Q|.
\end{aligned}
$$

Next recall that $Q \in \mathcal{S}_j$ and $K \in \mathcal{S}_{j+\ell+m}$. Hence $\operatorname{diam}(Q) = \sqrt{n} 2^{-j}$ and $\operatorname{diam}(K) = \sqrt{n} 2^{-j-m-\ell}$. Inserting these values gives (5.10).

By (3.8), in combination with (5.9) and (5.10) we obtain that

$$
\|T_{\ell,m}(f)\|_p \leq C_p 2^m \left\| \sum_{Q \in \mathcal{S}} \sum_{K \in \mathcal{U}_{\ell,m}(Q)} \left\langle f, \frac{\psi_K}{|K|} \right\rangle h_Q \right\|_p . \tag{5.11}
$$

Recall $K \in \mathcal{U}_{\ell,m}(Q)$ satisfies $|K| = |Q| 2^{n(-\ell-m)}$. Hence $|K|^{-1} |Q| 2^m = 2^{(n+1)m+n\ell}$. Thus the right hand side of (5.11) is bounded by

$$
C_p 2^{(n+1)m+n\ell} \left\| \sum_{Q \in \mathcal{S}} \sum_{K \in \mathcal{U}_{\ell,m}(Q)} \langle f, \psi_K \rangle h_Q |Q|^{-1} \right\|_p . \tag{5.12}
$$

Given $Q \in \mathcal{S}$ let $K_s(Q)$ be a cube in $\mathcal{U}_{\ell,m}(Q)$. As there exist at most $C = C_n$ cubes in $\mathcal{U}_{\ell,m}(Q)$, the expression in (5.12) is bounded by

$$
C_p 2^{(n+1)m+n\ell} \max_{s \leq C} \left\| \sum_{Q \in \mathcal{S}} \langle f, \psi_{K_s(Q)} \rangle h_Q |Q|^{-1} \right\|_p . \tag{5.13}
$$

Fix $s \leq C$ so that the maximum in the right hand side is assumed. We invoke rearrangement operators to obtain good upper bounds for (5.13). Let $\tau : \mathcal{S} \to \mathcal{S}$ be the map that associates to $Q \in \mathcal{S}$ its $(-m-\ell) - th$ dyadic predecessor, denoted $Q^{(-m-\ell)}$. Thus

$$
\tau(Q) = Q^{(-m-\ell)}.
$$

In sub-section 4.2 we defined the canonical splitting of $\mathcal{S}$ as

$$
\mathcal{S} = \mathcal{Q}_1 \cup \cdots \cup \mathcal{Q}_{2^{n(-m-\ell)}},
$$

so that for each fixed $k \leq 2^{n(-m-\ell)}$, the map $\tau : \mathcal{Q}_k \to \mathcal{S}$ is a bijection. Fix now $k \leq 2^{n(-m-\ell)}$ and define the family of functions $\{\varphi_W^{(k)} : W \in \mathcal{S}\}$ by the equations

$$
\varphi_{\tau(Q)}^{(k)} = \psi_{K_s(Q)}, \qquad Q \in \mathcal{Q}_k.
$$

Let $A = 2^{n(-m-\ell)}$ and define the rearrangement operator S by

$$
S(f) = \sum_{k=1}^{A} \sum_{Q \in \mathcal{Q}_k} \left\langle f, \varphi_{\tau(Q)}^{(k)} \right\rangle h_Q |Q|^{-1}.
$$

What we have obtained so far can be summarized in one line as follows

$$\|T_{\ell,m}(f)\|_p \leq C_p 2^{(n+1)m+n\ell}\|S(f)\|_p. \tag{5.14}$$

It remains to find estimates for $\|S(f)\|_p$. To this end observe that the family of functions $\{\varphi_W^{(k)} : W \in \mathcal{S}\}$ satisfies the structural conditions (4.11): There exists $C > 0$ so that for each $W \in \mathcal{S}$

$$\text{supp } \varphi_W^{(k)} \subseteq C \cdot Q, \qquad |\varphi_W^{(k)}| \leq C, \qquad \text{Lip}(\varphi_W^{(k)}) \leq C \text{ diam}(W)^{-1}.$$

Hence Theorem 4.2 applied to the operator S, with $\lambda = -m - \ell$, gives

$$\|S\|_p \leq \begin{cases} C_p 2^{n(-m-\ell)}\sqrt{-m-\ell} & \text{for} \quad p \geq 2; \\ C_p 2^{n(-m-\ell)} & \text{for} \quad p \leq 2. \end{cases}$$

Inserting the norm estimate for S into (5.14) and simple arithmetic implies (5.8).

■

5.2 Estimates for $T_{\ell,m}$, $\ell \geq 0$, $m > 0$.

In this subsection we treat the case $m > 0$ and $\ell \geq 0$ or equivalently $2^{-(j+\ell+m)} < 2^{-(j+\ell)}$. Here the length scale of the test function ψ_K is finer than the scale of $\Delta_{j+\ell}(h_Q^{(\varepsilon)})$. We estimate the norm of $T_{\ell,m}$ by reduction to the projections onto ring domains.

Proposition 5.2 *Let $1 < p < \infty$. and $1/p + 1/q = 1$. For $m \geq 0$ and $\ell \geq 0$, the operator*

$$T_{\ell,m}(f) = \sum_{j=-\infty}^{\infty} \sum_{Q \in \mathcal{S}_j} \sum_{K \in \mathcal{S}_{j+\ell+m}} \left\langle f, \frac{\psi_K}{|K|} \right\rangle \langle \Delta_{j+\ell}(h_Q^{(\varepsilon)}), \psi_K \rangle h_Q^{(\varepsilon)} |Q|^{-1}.$$

satisfies the norm estimate

$$\|T_{\ell,m}\|_p \leq \begin{cases} C_p 2^{-m} 2^{-\ell/2} & \textit{for} \quad p \geq 2; \\ C_p 2^{-m} 2^{-\ell/q} & \textit{for} \quad p \leq 2. \end{cases} \tag{5.15}$$

Proof. We divide the proof into three parts. First we rewrite the operator by isolating the cubes $Q \in \mathcal{S}_j$ and $K \in \mathcal{S}_{j+\ell+m}$ that contribute to the series defining $T_{\ell,m}$. Second we define auxiliary operators that dominate $T_{\ell,m}$. These turn out to be projections onto ring domains as considered in sub-section 4.1. Finally we invoke norm estimates for the resulting projections onto ring domains.

Part 1. Here we rewrite $T_{\ell,m}$ by making explicit the index set $\{K \in \mathcal{S}_{j+\ell+m}\}$ that actually contributes to the series defining $T_{\ell,m}$. Fix $Q \in \mathcal{S}_j$ and define the collection of dyadic cubes

$$\mathcal{U}_{\ell,m}(Q) = \{K \in \mathcal{S}_{j+\ell+m} : \langle \Delta_{j+\ell}(h_Q^{(\varepsilon)}), \psi_K \rangle \neq 0\}.$$

Let $U_{\ell,m}(Q)$ be the pointset that is covered by the collection $\mathcal{U}_{\ell,m}(Q)$. Note that $U_{\ell,m}(Q)$ is contained in the ring domain of points that have distance $\leq C2^{-\ell-j}$ to the set of discontinuities of $h_Q^{(\varepsilon)}$. Thus $U_{\ell,m}(Q)$ can be covered by at most $C2^{(n-1)\ell}$ dyadic cubes of diameter $\sqrt{n}2^{-\ell-j}$.

We denote these cubes (that are pairwise disjoint) by $E_1, \dots, E_A$ where $A = C2^{(n-1)\ell}$. If we wish to emphasize the dependence on Q we write $E_k = E_k(Q)$. Thus

$$U_{\ell,m}(Q) \subseteq \bigcup_{k=1}^{A} E_k(Q), \qquad \operatorname{diam}(E_k(Q)) = \sqrt{n}2^{-\ell-j}, \quad A = C2^{(n-1)\ell}.$$

With $\mathcal{U}_{\ell,m}(Q)$ as index set we define the block bases of wavelet functions

$$\widetilde{\psi}_Q = \sum_{K \in \mathcal{U}_{\ell,m}(Q)} \left\langle \Delta_{j+\ell}(h_Q^{(\varepsilon)}), \psi_K \right\rangle \psi_K |K|^{-1},$$

by which we rewrite the operator $T_{\ell,m}$ as follows,

$$T_{\ell,m}(f) = \sum_{Q \in \mathcal{S}} \left\langle f, \widetilde{\psi}_Q \right\rangle h_Q^{(\varepsilon)} |Q|^{-1}. \tag{5.16}$$

Part 2. Here we exploit (5.16) and relate the representation $T_{\ell,m}$ to its dyadic counterpart, the projection onto ring domains. To this end we start by giving pointwise estimates for the function $\widetilde{\psi}_Q$. Fix $K \in \mathcal{U}_{\ell,m}(Q)$. Use that ψ_K has mean zero and that $\operatorname{diam}(K) = \sqrt{n}2^{(-j-\ell-m)}$ to obtain,

$$\begin{aligned} |\langle \Delta_{j+\ell}(h_Q^{(\varepsilon)}), \psi_K \rangle| \cdot |K|^{-1} &\leq C \operatorname{diam}(K) \operatorname{Lip}(\Delta_{j+\ell}(h_Q^{(\varepsilon)})) \\ &\leq C \operatorname{diam}(K) 2^{j+\ell} \\ &= C2^{-m}. \end{aligned} \tag{5.17}$$

Recall that

$$\operatorname{dist}(U_{\ell,m}(Q), Q) \leq C \cdot \operatorname{diam}(Q) \quad Q \in \mathcal{S}.$$

Hence there exists a universal $A_0 \in \mathbb{N}$ so that for $j \in \mathbb{Z}$ the collection $\mathcal{S}_j$ may split as

$$\mathcal{S}_j^{(1)}, \dots, \mathcal{S}_j^{(A_0)},$$

so that for $s \leq A_0$ the sets $\{U_{\ell,m}(Q) : Q \in \mathcal{S}_j^{(s)}\}$ are pairwise disjoint. Fix $s \leq A_0$ and form the collections

$$\mathcal{B}_s = \bigcup_{j \in \mathbb{Z}} \mathcal{S}_j^{(s)}.$$

As $s \leq A_0$ is fixed, the collections $\{\mathcal{U}_{\ell,m}(Q) : Q \in \mathcal{B}_s\}$ satisfy the conditions (3.9) and (3.10). Define

$$d_Q = \sum_{K \in \mathcal{U}_{\ell,m}(Q)} h_K,$$

and put

$$F_s(g) = \sum_{Q \in \mathcal{B}_s} \langle g, d_Q \rangle h_Q |Q|^{-1}.$$

By (5.17) and (3.15), (3.16),

$$\|T_{\ell,m}\|_p \leq C_p 2^{-m} \sum_{s=1}^{A_0} \|F_s\|_p.$$

Next we replace the operator F_s by a related one that is easier to analyze. To this end we define for $Q \in \mathcal{B}_s$,

$$g_Q = \sum_{k=1}^{A} h_{E_k(Q)}, \qquad A = C2^{(n-1)\ell},$$

where the collection of dyadic cubes $\{E_1(Q) \dots E_A(Q)\}$ are defined in part 1 of the proof. The block bases $\{g_Q : Q \in \mathcal{B}_s\}$ give rise to the operators G_s defined by,

$$G_s(f) = \sum_{Q \in \mathcal{B}_s} \langle f, g_Q \rangle h_Q |Q|^{-1}.$$

By (3.13), $\|F_s\|_p \le C_p \|G_s\|_p$. Hence

$$\|T_{\ell,m}\|_p \le C_p 2^{-m} \sum_{s=1}^{A_0} \|G_s\|_p. \tag{5.18}$$

Part 3. In the last part of the proof we obtain norm estimates for $T_{\ell,m}$ by recalling the bounds for the projection G_s^* obtained in Section 4. Fix $s \le A_0$, let

$$\mathcal{B} = \mathcal{B}_s \quad \text{and} \quad G = G_s.$$

The transposed operator G^* is just

$$G^*(f) = \sum_{Q \in \mathcal{B}} \langle f, h_Q \rangle \, g_Q |Q|^{-1}.$$

In part 1 of the proof, for $Q \in \mathcal{B}$, we defined the collections $\{E_1(Q), \dots, E_A(Q)\}$.They satisfy conditions (4.2)–(4.4). Hence we apply Theorem 4.1 with $S = G^*$ and $\lambda = \ell$. By duality this gives the following three norm estimates for G,

$$\|G\|_{H^1_d} \le C, \quad \|G\|_2 \le C2^{-\ell/2} \quad \text{and} \quad \|G\|_{\mathrm{BMO}_d} \le C2^{-\ell/2}. \tag{5.19}$$

By interpolation and (5.19), for $1 < p < \infty$ and $1/p + 1/q = 1$

$$\|G\|_p \le \begin{cases} C_p 2^{-\ell/2} & \text{for} \quad p \ge 2; \\ C_p 2^{-\ell/q} & \text{for} \quad p \le 2. \end{cases} \tag{5.20}$$

With (5.20) and (5.18) we deduce (5.15).

5.3 Estimates for $T_{\ell,m}$, $\ell \ge 0$, $-\ell \le m \le 0$.

Here we analyze the operators $T_{\ell,m}$, when $\ell \ge 0$, $-\ell \le m \le 0$. In this case the scale of the test functions ψ_K lies in between the scale of the cube Q and that of $\Delta_{j+\ell}(h_Q^{(\varepsilon)})$. Again we estimate $T_{\ell,m}$ by reduction to projection operators onto ring domains, following the pattern of the previous sub-section.

Proposition 5.3 *Let $1 < p < \infty$. and $1/p + 1/q = 1$. Let $\ell \ge 0$ and $-\ell \le m \le 0$ then the operator*

$$T_{\ell,m}(f) = \sum_{j=-\infty}^{\infty} \sum_{Q \in \mathcal{S}_j} \sum_{K \in \mathcal{S}_{j+\ell+m}} \left\langle f, \frac{\psi_K}{|K|} \right\rangle \langle \Delta_{j+\ell}(h_Q^{(\varepsilon)}), \psi_K \rangle h_Q^{(\varepsilon)} |Q|^{-1}$$

satisfies the norm estimate

$$\|T_{\ell,m}\|_p \le \begin{cases} C_p 2^{m/2}2^{-\ell/2} & \textit{for} \quad p \ge 2; \\ C_p 2^{m/2}2^{-\ell/q} & \textit{for} \quad p \le 2. \end{cases} \tag{5.21}$$

Proof. The proof splits canonically into three parts. First we analyze and rewrite $T_{\ell,m}$. Then we define auxiliary operators that dominate $T_{\ell,m}$, and continue with norm estimates for those operators. As above we are led to consider projections onto ring domains.

Part 1. Fix $\ell \ge 0$ and $-\ell \le m \le 0$. Let $j \in \mathbb{Z}$ and choose a dyadic cube $Q \in \mathcal{S}_j$. Then form the collection of cubes

$$\mathcal{U}_{\ell,m}(Q) = \{K \in \mathcal{S}_{j+\ell+m} : \langle \psi_K, \Delta_{j+\ell}(h_Q^{(\varepsilon)})\rangle \neq 0\}.$$

Observe that with the above definition of the collections $\mathcal{U}_{\ell,m}(Q)$ the following identity holds

$$T_{\ell,m}(f) = \sum_{j=-\infty}^{\infty} \sum_{Q\in\mathcal{S}_j} \sum_{K\in\mathcal{U}_{\ell,m}(Q)} \left\langle f, \frac{\psi_K}{|K|}\right\rangle \langle \Delta_{j+\ell}(h_Q^{(\varepsilon)}), \psi_K\rangle h_Q^{(\varepsilon)} |Q|^{-1}.$$

Part 2. Fix $Q \in \mathcal{S}_j$ and $K \in \mathcal{U}_{\ell,m}(Q)$. To find the auxiliary operators we prove first that

$$\left|\left\langle \Delta_{j+\ell}(h_Q^{(1,0)}), \psi_K\right\rangle\right| \le C2^m|K| \tag{5.22}$$

To see this make the following observation. First note that $|Q| = 2^{-nj}$ and $\operatorname{diam}(K) = \sqrt{n}2^{-j-m-\ell}$. Then observe that $\Delta_{j+\ell}(h_Q^{(\varepsilon)}$ is supported in the ring domain $D_\ell(Q)$ and estimate

$$\begin{aligned} \left|\left\langle \Delta_{j+\ell}(h_Q^{(\varepsilon)}), \psi_K\right\rangle\right| &\le C\int_K |\Delta_{j+\ell}(h_Q^{(\varepsilon)})| \\ &\le C|D_\ell(Q) \cap K| \\ &\le C2^{-\ell-j}(\operatorname{diam}(K))^{n-1} \\ &\le C2^m|K|. \end{aligned}$$

For a cube $K \in \mathcal{U}_{\ell,m}(Q)$ its distance to Q is bounded by the $C\operatorname{diam}(Q)$. Hence, there exists a universal A_0 so that for $j \in \mathbb{Z}$ the collection $\mathcal{S}_j$ can be split into

$$\mathcal{S}_j^{(1)}, \dots, \mathcal{S}_j^{(A_0)},$$

so that the sets $\{U_{\ell,m}(Q) : Q \in \mathcal{S}_j^{(s)}\}$ are pairwise disjoint. Fix $s \le A_0$ and form the collections

$$\mathcal{B}_s = \bigcup_{j\in\mathbb{Z}} \mathcal{S}_j^{(s)}.$$

Note that $\{\mathcal{U}_{\ell,m}(Q) : Q \in \mathcal{B}_s\}$ satisfies the conditions (3.9) and (3.10). Define

$$F_s(f) = 2^m \sum_{Q\in\mathcal{B}_s} \langle f, d_Q\rangle h_Q|Q|^{-1}, \qquad d_Q = \sum_{K\in\mathcal{U}_{\ell,m}(Q)} h_K.$$

The integral estimates (3.16), (3.15) and (5.22) imply

$$\|T_{\ell,m}\|_p \le C_p \sum_{s=1}^{A_0} \|F_s\|_p.$$

Part 3. It remains to estimate $\|F_s\|_p$. Notice that the collections $\mathcal{U}_{\ell,m}(Q)$, $Q \in \mathcal{B}_s$ satisfy conditions (4.2)–(4.4). Next apply Theorem 4.1 to $S = 2^{-m} F_s^*$ and $\lambda = \ell + m$. By duality this yields for F_s the norm estimates on L^2, H^1_d and BMO_d

$$\|F_s\|_2 \leq C2^{(m-\ell)/2}, \quad \|F_s\|_{H^1_d} \leq C2^m, \quad \text{and} \quad \|F_s\|_{\mathrm{BMO}_d} \leq C2^{(m-\ell)/2}. \tag{5.23}$$

By interpolation from (5.23) we get for $1 < p < \infty$ and $1/p + 1/q = 1$ that,

$$\|F_s\|_p \leq \begin{cases} C_p 2^{(m-\ell)/2} & \text{for} \quad p \geq 2; \\ C_p 2^{m/2-\ell/q} & \text{for} \quad p \leq 2. \end{cases}$$

■

5.4 Estimates for $T_\ell^{(\varepsilon)} R_{i_0}^{-1}$, $\ell \geq 0$.

We give the norm estimates for $T_\ell^{(\varepsilon)} R_{i_0}^{-1}$, $\ell \geq 0$, $\varepsilon \in \mathcal{A}_{i_0}$, and $1 \leq i_0 \leq n$. We do this *by reduction* to the estimates for the operator $T_\ell^{(\varepsilon)}$, $\ell \geq 0$. Strictly speaking we discuss the reduction to the proof given in the previous sub sections. We obtain a series representing $T_\ell^{(\varepsilon)} R_{i_0}^{-1}$, analyze the shape and form of the measures $\mathbb{E}_{i_0} \partial_i h_Q^{(\varepsilon)}$ and describe how the convolution operator $\Delta_{j+\ell}$ acts on those measures. In the following analysis we also collect the information needed for the estimates of the $T_\ell^{(\varepsilon)} R_{i_0}^{-1}$ when $\ell \leq 0$.

The representation of $T_\ell^{(\varepsilon)} R_{i_0}^{-1}$. In Theorem 2.1 and Theorem 2.2 we aim at estimates for $T_\ell^{(\varepsilon)} R_{i_0}^{-1}$ when $\varepsilon \in \mathcal{A}_{i_0}$. Hence we seek an explicit expansion for $T_\ell^{(\varepsilon)} R_{i_0}^{-1}$. By (3.23) we have

$$R_{i_0}^{-1} = R_{i_0} + \sum_{\substack{i=1 \\ i \neq i_0}}^{n} \mathbb{E}_{i_0} \partial_i R_i \quad \text{and} \quad T_\ell^{(\varepsilon)} R_{i_0}^{-1} = T_\ell^{(\varepsilon)} R_{i_0} + \sum_{\substack{i=1 \\ i \neq i_0}}^{n} T_\ell^{(\varepsilon)} \mathbb{E}_{i_0} \partial_i R_i. \tag{5.24}$$

Let $j \in \mathbb{Z}$. Recall that $\mathcal{S}_j$ denotes the family of dyadic cubes Q for which $|Q| = 2^{-nj}$. Let $Q \in \mathcal{S}_j$, $i \neq i_0$, and $\varepsilon \in \mathcal{A}_{i_0}$. Then form

$$k_Q^{(\ell,i)} = \Delta_{j+\ell} \left(\mathbb{E}_{i_0} \partial_i h_Q^{(\varepsilon)} \right). \tag{5.25}$$

Thus by (5.24)

$$T_\ell^{(\varepsilon)} R_{i_0}^{-1}(u) = T_\ell^{(\varepsilon)} R_{i_0}(u) + \sum_{Q \in \mathcal{S}} \sum_{\substack{i=1 \\ i \neq i_0}}^{n} \langle R_i(u), k_Q^{(\ell,i)} \rangle h_Q^{(\varepsilon)} |Q|^{-1}. \tag{5.26}$$

Given the representation (5.26) we further analyze the functions $\{k_Q^{(\ell,i)} : Q \in \mathcal{S}\}$. It is only at this point of our analysis that we exploit the fact that i_0 and ε are related by the condition $\varepsilon \in \mathcal{A}_{i_0}$.

The measures $\mathbb{E}_{i_0}\partial_i h_Q^{(\varepsilon)}$. We defined $k_Q^{(\ell,i)}$ by a convolution operator applied to

$$\mathbb{E}_{i_0}\partial_i h_Q^{(\varepsilon)}, \qquad i \neq i_0, \quad \varepsilon \in \mathcal{A}_{i_0},$$

where ∂_i denotes the differentiation with respect to the y_i variable and $\mathbb{E}_{i_0}$ denotes integration with respect to the $x_{i_0} - th$ coordinate,

$$\mathbb{E}_{i_0}(f)(x) = \int_{-\infty}^{x_{i_0}} f(x_1,\dots,s,\dots,x_n)ds, \qquad x = (x_1,\dots,x_n).$$

Thus, $\mathbb{E}_{i_0}\partial_i h_Q^{(\varepsilon)}$ admits a convenient factorization: Let $x = (x_1,\dots,x_n)$, then

$$\mathbb{E}_{i_0}\partial_i h_Q^{(\varepsilon)}(x) = \left[\int_{-\infty}^{x_{i_0}} h_{I_{i_0}}^{\varepsilon_{i_0}}(s)ds\right] \left[\partial_i h_{I_i}^{\varepsilon_i}(x_i)\right] \left[\prod\{h_{I_k}^{\varepsilon_k}(x_k) : k \notin \{i_0, i\}\}\right]. \tag{5.27}$$

The properties of the three factors appearing in (5.27) are as follows.

1. As $\varepsilon \in \mathcal{A}_{i_0}$, we have $\varepsilon_{i_0} = 1$, hence the first factor in (5.27)

$$x_{i_0} \to \int_{-\infty}^{x_{i_0}} h_{I_{i_0}}^{\varepsilon_{i_0}}(s)ds$$

is supported in the interval I_{i_0}. Furthermore it is bounded by $|I_{i_0}|$ and piecewise linear with nodes at $l(I_{i_0}), m(I_{i_0})$ and $r(I_{i_0})$ and slopes $+1, -1$ or 0. Here we let $l(I_{i_0})$ denote the left endpoint of I_{i_0}, and $m(I_{i_0})$, $r(I_{i_0})$ denote its midpoint, respectively its right endpoint.

2. The partial derivatives ∂_i applied to $h_Q^{(\varepsilon)}$ induces a Dirac measure, at each of the discontinuities of $h_{I_i}^{\varepsilon_i}$. The resulting formulas depend on the value of $\varepsilon_i \in \{0, 1\}$, since

$$\begin{aligned} \partial_i h_{I_i} &= \delta_{l(I_i)} - 2\delta_{m(I_i)} + \delta_{r(I_i)}, \\ \partial_i 1_{1_i} &= \delta_{l(I_i)} - \delta_{r(I_i)}. \end{aligned}$$

In either case, for $\varphi \in C^\infty(\mathbb{R})$ the above identities yield the estimate,

$$\left|\langle \partial_i h_{I_i}^{\varepsilon_i}, \varphi\rangle\right| \leq 2 \sup\left\{\frac{|\varphi(s) - \varphi(t)|}{|s-t|} \,:\, s, t, \in I\right\} |I_i|. \tag{5.28}$$

3. The third factor in (5.27) is the function

$$x \to \prod\{h_{I_k}^{\varepsilon_k}(x_k) : k \notin \{i_0, i\}\} \tag{5.29}$$

It is piecewise constant and assumes the values $\{-1, 0, +1\}$. When restricted to a dyadic cube W with $\operatorname{diam}(W) \leq \operatorname{diam}(Q)/2$ the factor (5.29) defines a constant function.

As a result of the above discussion $\mathbb{E}_{i_0}\partial_i h_Q^{(\varepsilon)}$ is a measure supported on Q so that for any continous function on $\mathbb{R}^n$,

$$|\langle \mathbb{E}_{i_0}\partial_i h_Q^{(\varepsilon)}, \varphi\rangle| \leq |Q| \cdot \|\varphi\|_\infty \quad \text{and} \quad \langle \mathbb{E}_{i_0}\partial_i h_Q^{(\varepsilon)}, 1\rangle = 0.$$

The convolution $\Delta_{j+\ell}$ acting on $\mathbb{E}_{i_0}\partial_i h_Q^{(\varepsilon)}$. Recall that in (2.2) the operator $\Delta_{j+\ell}$ is given as convolution with $d_{j+\ell}$ so that

$$\Delta_{j+\ell}\left(\mathbb{E}_{i_0}\partial_i h_Q^{(\varepsilon)}\right) = \mathbb{E}_{i_0}\partial_i h_Q^{(\varepsilon)} * d_{j+\ell},$$

with

$$\text{supp } d_{j+\ell} \subseteq [-C2^{-(j+\ell)}, C2^{-(j+\ell)}]^n, \quad |d_{j+\ell}| \leq C2^{n(j+\ell)}, \quad \text{Lip}(d_{j+\ell}) \leq C2^{(n+1)(j+\ell)}. \tag{5.30}$$

Moreover for $1 \leq i \leq n$ by (2.1)

$$\int_{\mathbb{R}} d_{j+\ell}(x-y)y_i dy_i = 0 \quad \text{and} \quad \int_{\mathbb{R}^n} d_{j+\ell}(x-y)dy = 0, \quad x \in \mathbb{R}^n. \tag{5.31}$$

We derive next for $k_Q^{(\ell,i)}$ its structural estimates concerning support, Lipschitz properties and pointwise bounds. It turns out that these depend critically on the value of $\text{sign}\,(\ell)$:

1. The case $\ell \geq 0$. For $Q \in \mathcal{S}$ and $\varepsilon \in \mathcal{A}_{i_0}$ let $D^{(\varepsilon)}(Q)$ denote the set of discontinuities of the Haar function $h_Q^{(\varepsilon)}$. Fix $\ell \in \mathbb{N}$ and define

   $$D_\ell^{(\varepsilon)}(Q) = \{x \in \mathbb{R}^n : \text{ dist}(x, D^{(\varepsilon)}(Q)) \leq C2^{-\ell} \text{ diam}(Q)\}.$$

   Thus $D_\ell^{(\varepsilon)}(Q)$ is the set of points that have distance $\leq C2^{-\ell}$ diam(Q) to the set of discontinuities of $h_Q^{(\varepsilon)}$.

   Fix $x \notin D_\ell^{(\varepsilon)}(Q)$. As we observed in the paragraphs following (5.27) there exist $A \in \{-1, 0, 1\}$ and $a \in \mathbb{R}$ so that,

   $$\mathbb{E}_{i_0} h_Q^{(\varepsilon)}(y) = A(y_{i_0} - a), \quad \text{for} \quad y \in B(x, c2^{-(j+\ell)}). \tag{5.32}$$

   Combining now (5.30) with (5.31) and we find

   $$\begin{aligned} \Delta_{j+\ell}(\mathbb{E}_{i_0} h_Q^{(\varepsilon)})(x) &= \int_{\mathbb{R}^n} d_{j+\ell}(x-y)\mathbb{E}_{i_0} h_Q^{(\varepsilon)}(y)dy \\ &= A\int_{\mathbb{R}^n} d_{j+\ell}(x-y)(y_{i_0} - a)dy \\ &= 0. \end{aligned} \tag{5.33}$$

   Since $\Delta_{j+\ell}$ is a convolution operator it commutes with differentiation, and we obtain for $x \notin D_\ell^{(\varepsilon)}(Q)$,

   $$\begin{aligned} \Delta_{j+\ell}(\mathbb{E}_{i_0}\partial_i h_Q^{(\varepsilon)})(x) &= \partial_i \Delta_{j+\ell}(\mathbb{E}_{i_0} h_Q^{(\varepsilon)})(x) \\ &= 0. \end{aligned} \tag{5.34}$$

   Combining (5.34) with (5.30) we obtain that the functions $\{k_Q^{(\ell,i)} : Q \in \mathcal{S},\, i \neq i_0,\, \ell \geq 0\}$ satisfy the structural conditions

   $$\text{supp } k_Q^{(\ell,i)} \subseteq D_\ell^{(\varepsilon)}(Q), \quad |k_Q^{(\ell,i)}| \leq C2^\ell, \quad \text{Lip}(k_Q^{(\ell,i)}) \leq C2^{2\ell}(\text{ diam}(Q))^{-1}, \tag{5.35}$$

   with $C > 0$ independent of $Q \in \mathcal{S},\ i \neq i_0,$ or $\ell \geq 0$.

2. The case $\ell \leq 0$. In this case we use (5.28) and (5.30) to see that the family $\{k_Q^{(\ell,i)} : Q \in \mathcal{S},\ i \neq i_0,\ \ell \leq 0\}$, satisfies the following conditions

$$\operatorname{supp} k_Q^{(\ell,i)} \subseteq (C2^{|\ell|})\cdot Q, \quad |k_Q^{(\ell,i)}| \leq C2^{\ell(n+1)}, \quad \operatorname{Lip}(k_Q^{(\ell,i)}) \leq C2^{\ell(n+2)}(\operatorname{diam}(Q))^{-1}, \tag{5.36}$$

were again $C > 0$ is independent of $Q \in \mathcal{S},\ i \neq i_0,$ or $\ell \leq 0$.

Proposition 5.4 *Let $1 < p < \infty$. Let $1 \leq i \neq i_0 \leq n$ and $\varepsilon \in \mathcal{A}_{i_0}$. For $\ell \geq 0$ the operator X defined by*

$$X(f) = \sum_{Q\in\mathcal{S}} \langle f, k_Q^{(\ell,i)}\rangle h_Q^{(\varepsilon)} |Q|^{-1},$$

satisfies the norm estimates

$$||X||_p \leq \begin{cases} C_p 2^{+\ell/2} & \text{if} \quad p \geq 2; \\ C_p 2^{+\ell/p} & \text{if} \quad p \leq 2. \end{cases} \tag{5.37}$$

Proof. Recall the expansion (5.2) asserting that

$$T_\ell^{(\varepsilon)}(f) = \sum_{Q\in\mathcal{S}} \langle f, f_{Q,\ell}^{(\varepsilon)}\rangle h_Q^{(\varepsilon)} |Q|^{-1},$$

where $f_{Q,\ell}^{(\varepsilon)}$ has vanishing mean and satisfies the basic estimates (5.3),

$$\operatorname{supp} f_{Q,\ell}^{(\varepsilon)} \subseteq D_\ell^{(\varepsilon)}(Q), \qquad |f_{Q,\ell}^{(\varepsilon)}| \leq C, \qquad \operatorname{Lip}(f_{Q,\ell}^{(\varepsilon)}) \leq C2^\ell(\operatorname{diam}(Q))^{-1},$$

and where $D_\ell^{(\varepsilon)}(Q)$ is the set of points that have distance $\leq C2^{-\ell}\operatorname{diam}(Q)$ to the set of discontinuities of $h_Q^{(\varepsilon)}$. Using only the scale invariant conditions (5.3) we proved that $T_\ell^{(\varepsilon)}$, $(\ell \geq 0)$ satisfies the norm estimates (5.4), that is,

$$||T_\ell^{(\varepsilon)}||_p \leq \begin{cases} C_p 2^{-\ell/2} & \text{for} \quad p \geq 2; \\ C_p 2^{-\ell/q} & \text{for} \quad p \leq 2. \end{cases}$$

Observe that by (5.35) the functions $\{2^{-\ell} k_Q^{(\ell,i)}\}$ satisfy the very same structure conditions (5.3) as $\{f_{Q,\ell}^{(\varepsilon)}\}$. Hence for the norm of the operator $2^{-\ell}X$ there hold the same upper bounds as for $T_\ell^{(\varepsilon)}$, $\ell \geq 0$. Consequently, the norm of X can be estimated as

$$||X||_p \leq \begin{cases} C_p 2^{\ell-\ell/2} & \text{if} \quad p \geq 2; \\ C_p 2^{\ell-\ell/q} & \text{if} \quad p \leq 2. \end{cases}$$

■

Proposition 5.4 in combination with (2.4) and (5.26) implies that for $\ell > 0$,

$$||T_\ell^{(\varepsilon)} R_{i_0}^{-1}||_p \leq \begin{cases} C_p 2^{\ell/2} & \text{if} \quad p \geq 2; \\ C_p 2^{\ell/p} & \text{if} \quad p \leq 2. \end{cases}$$

6 The Proof of Theorem 2.2.

In this section we prove Theorem 2.2. It turns out that for $\ell \le 0$ the norm estimates for $T_\ell^{(\varepsilon)} R_{i_0}^{-1}$ and $T_\ell^{(\varepsilon)}$ are much simpler than for $\ell \ge 0$. Indeed for $\ell < 0$ the scale of $Q \in \mathcal{S}_j$ is finer than the scale of $\Delta_{j+\ell}(h_Q^{(\varepsilon)})$ and the discontinuities of the Haar function are completely smeared out. We can therefore reduce the problem to estimates for rearrangement operators acting on Haar functions, treating $T_\ell^{(\varepsilon)} R_{i_0}^{-1}$ and $T_\ell^{(\varepsilon)}$ simultaneously by the same method.

Let u be a smooth function with vanishing mean and compact support. Let $i \ne i_0$ and $\varepsilon \in \mathcal{A}_{i_0}$. Then

$$T_\ell^{(\varepsilon)} R_{i_0}^{-1}(u) = T_\ell^{(\varepsilon)} R_{i_0}(u) + \sum_{Q \in \mathcal{S}} \sum_{\substack{i=1 \\ i \ne i_0}}^{n} \langle R_i(u), k_Q^{(\ell,i)} \rangle h_Q^{(\varepsilon)} |Q|^{-1},$$

where

$$k_Q^{(\ell,i)} = \Delta_{j+\ell}(\mathbb{E}_{i_0} \partial_i h_Q^{(\varepsilon)}), \qquad Q \in \mathcal{S}_j.$$

Since $\ell < 0$ the functions $\{k_Q^{(\ell,i)} : Q \in \mathcal{S},\, i \ne i_0,\, \ell \le 0\}$, satisfy conditions (5.36). Recall also that

$$T_\ell^{(\varepsilon)}(u) = \sum_{Q \in \mathcal{S}} \langle u, f_{Q,\ell}^{(\varepsilon)} \rangle h_Q^{(\varepsilon)} |Q|^{-1},$$

where

$$f_{Q,\ell}^{(\varepsilon)} = \Delta_{j+\ell}(h_Q^{(\varepsilon)}), \qquad Q \in \mathcal{S}_j.$$

It is easy to see that also the family $\{f_{Q,\ell}^{(\varepsilon)} : Q \in \mathcal{S}\,, \ell \le 0\}$ satisfies the same structural conditions (5.36), that is

$$\operatorname{supp} f_{Q,\ell}^{(\varepsilon)} \subseteq (C2^{|\ell|}) \cdot Q, \quad |f_{Q,\ell}^{(\varepsilon)}| \le C2^{-|\ell|(n+1)}, \quad \operatorname{Lip}(f_{Q,\ell}^{(\varepsilon)}) \le C2^{-|\ell|(n+2)}(\operatorname{diam}(Q))^{-1}. \tag{6.1}$$

Proposition 6.1 *If $\ell \le 0$ then*

$$||T_\ell^{(\varepsilon)}||_p + ||T_\ell^{(\varepsilon)} R_{i_0}^{-1}||_p \le \begin{cases} C_p 2^{-2|\ell|/p} & \text{for} \quad p \ge 2; \\ C_p 2^{-|\ell|} & \text{for} \quad p \le 2. \end{cases}$$

Proof. Let $1 \le i \ne i_0 \le n$. Let $Q \in \mathcal{S}$. Choose signs $\delta_{Q,i}, \epsilon_Q \in \{+1, 0, -1\}$ and form

$$g_{Q,\ell} = \left[\sum_{i=1,\, i \ne i_0}^{n} \delta_{Q,i} k_Q^{(\ell,i)} \right] + \epsilon_Q f_{Q,\ell}^{(\varepsilon)}. \tag{6.2}$$

We emphasize that the definition of $g_{Q,\ell}$ depends on the choice of signs $\delta_{Q,i}, \epsilon_Q \in \{+1, 0, -1\}$; nevertheless our notation suppresses this dependence. Note that by (5.36) and (6.1) the functions $\{g_{Q,\ell}\}$ are of mean zero and satisfy structure conditions, not depending on the choice of signs, namely

$$\operatorname{supp} g_{Q,\ell} \subseteq C2^{|\ell|} \cdot Q, \quad |g_{Q,\ell}| \le C2^{-(n+1)|\ell|}, \quad \operatorname{Lip}(g_{Q,\ell}) \le C2^{-(n+2)|\ell|} \operatorname{diam}(Q)^{-1}. \tag{6.3}$$

Consider the rearrangement $\tau : \mathcal{S} \to \mathcal{S}$ that maps $Q \in \mathcal{S}$ to its $|\ell| - th$ dyadic predecessor. Let $\mathcal{Q}_1, \dots, \mathcal{Q}_{2^{n|\ell|}}$ be the canonical splitting of $\mathcal{S}$ so that for fixed $k \le 2^{n|\ell|}$ the map $\tau : \mathcal{Q}_k \to \mathcal{S}$ is bijective. Fix $k \le 2^{n|\ell|}$. Determine the family $\{\varphi_W^{(k)} : W \in \mathcal{S}\}$ by the equations

$$\varphi_{\tau(Q)}^{(k)} = 2^{(n+1)|\ell|} g_{Q,\ell}, \qquad Q \in \mathcal{Q}_k. \tag{6.4}$$

Thus defined the functions $\varphi_W^{(k)}$ are of mean zero and satisfy the structural conditions

$$\operatorname{supp} \varphi_W^{(k)} \subseteq C \cdot W, \qquad |\varphi_W^{(k)}| \le C, \qquad \operatorname{Lip}(\varphi_W^{(k)}) \le C \operatorname{diam}(W)^{-1}.$$

Define the operator

$$S(u) = \sum_{k=1}^{2^{n|\ell|}} \sum_{Q \in \mathcal{Q}_k} \left\langle u, \varphi_{\tau(Q)}^{(k)} \right\rangle h_Q^{(\varepsilon)} |Q|^{-1}.$$

Apply Theorem 4.2 to S with $\lambda = |\ell|$. This yields

$$\|S\|_2 \le C_0 2^{n|\ell|}, \qquad \|S\|_{H_d^1} \le C_0 2^{n|\ell|}, \qquad \|S : \mathrm{BMO}(\mathbb{R}^n) \to \mathrm{BMO}_d\| \le C_0 |\ell|^{1/2} 2^{n|\ell|}. \tag{6.5}$$

Note that by (6.2) and (6.4) the algebraic definition of the operator S depends on the choice of signs $\delta_{Q,i}, \epsilon_Q \in \{+1, 0, -1\}$, yet by (6.5) our estimates for $\|S\|_p$ are independent thereof.

Let $g \in L^p$. Depending on g we choose $\delta_{Q,i}, \epsilon_Q \in \{+1, 0, -1\}$, hence S, so that

$$\|T_\ell^{(\varepsilon)}(g)\|_p + \|T_\ell^{(\varepsilon)} R_{i_0}^{-1}(g)\|_p \le C_p 2^{-(n+1)|\ell|} C_p \|S\|_p \|g\|_p. \tag{6.6}$$

Consequently, our upper bounds for $\|T_\ell^{(\varepsilon)}\|_p + \|T_\ell^{(\varepsilon)} R_{i_0}^{-1}\|_p$ follow from (6.5). Indeed, by interpolation and the estimate $|\ell|^{1/2} \le 2^{|\ell|/2}$, (6.5) and (6.6) imply that

$$||T_\ell^{(\varepsilon)}||_p + ||T_\ell^{(\varepsilon)} R_{i_0}^{-1}||_p \le \begin{cases} C_p 2^{-2|\ell|/p} & \text{for} \quad p \ge 2; \\ C_p 2^{-|\ell|} & \text{for} \quad p \le 2. \end{cases}$$

■

7 Sharpness of the exponents in Theorem 1.1.

In this section we construct the examples showing that the exponents $(1/2, 1/2)$ respectively $(1/p, 1/q)$ are sharp in the estimates of Theorem 1.1,

$$||P^{(\varepsilon)}(u)||_p \le C_p \|u\|_p^{1/2} \|R_{i_0}(u)\|_p^{1/2}, \quad p \ge 2, \tag{7.1}$$

and

$$||P^{(\varepsilon)}(u)||_p \le C_p \|u\|_p^{1/p} \|R_{i_0}(u)\|_p^{1/q}, \quad p \le 2, \tag{7.2}$$

where $1 \le i_0 \le n$ and $\varepsilon \in \mathcal{A}_{i_0}$.

When we say that we obtained sharp exponents in Theorem 1.1 we mean the following: Let $\eta > 0$. Since the Riesz transform is a bounded operator on $L^p (1 < p < \infty)$, replacing in (7.1) the pair of exponents $(1/2, 1/2)$ by $(1/2 - \eta, 1/2 + \eta)$ *would* lead to a statement that *implies* (7.1), hence would yield a stronger theorem. Our examples show, however, that improving the exponents in the right hand side of (7.1) is impossible. (The same holds for (7.2).) Specifically we have this theorem:

Theorem 7.1 *Let $1 \le i_0 \le n$, and $\varepsilon \in \mathcal{A}_{i_0}$. Let $1 < p < \infty$, $1/p + 1/q = 1$. and $\eta > 0$. Then*

$$\sup_{u \in L^p} \frac{||P^{(\varepsilon)}(u)||_p}{\|u\|_p^{1/2-\eta} \|R_{i_0}(u)\|_p^{1/2+\eta}} = \infty \qquad p \ge 2, \tag{7.3}$$

and

$$\sup_{u \in L^p} \frac{||P^{(\varepsilon)}(u)||_p}{\|u\|_p^{1/p-\eta} \|R_{i_0}(u)\|_p^{1/q+\eta}} = \infty, \qquad p \le 2. \tag{7.4}$$

For simplicity of notation we verify Theorem 7.1 only in the case when $n = 2$. The passage to arbitrary $n \in \mathbb{N}$ is routine and left to the reader. Moreover we carry out the proof of Theorem 7.1 with the following specification

$$n = 2, \quad i_0 = 1, \quad \varepsilon = (1, 0). \tag{7.5}$$

Throughout this section we assume (7.5) and put

$$P = P^{(1,0)}.$$

We obtain Theorem 7.1 by exhibiting a sequence of test functions for which the quotient in (7.3) respectively (7.4) tends to infinity. On each test function we prove lower L^p bounds for the action of P and upper L^p estimates for R_1. In sub-section 7.1 we define building blocks $s \otimes d$ and the test functions f_ϵ using a procedure that resembles that of adding independent copies of the basic building blocks. The proof of (7.3) requires upper estimates for $\|f_\epsilon\|_p$ and $\|R_1(f_\epsilon)\|_p$, that we prove in sub-section 7.2 and a lower estimates for $\|P(f_\epsilon)\|_p$ obtained in sub-section 7.3 .

7.1 The building blocks $s \otimes d$.

We build the examples showing sharpness of exponents on the properties of the functions $s \otimes d$ defined here. Throughout this section we fix $\epsilon > 0$.

Let A, B be Lipschitz functions on $\mathbb{R}$. Assume that

$$\text{supp } A \subseteq [0,1], \quad \int A = 0 \quad \text{and} \quad \text{supp } B \subseteq [-1,1]. \tag{7.6}$$

Given $x = (x_1, x_2)$ we define

$$s(x_1) = A(x_1), \quad d(x_2) = B(x_2/\epsilon),$$

$$s \otimes d(x) = s(x_1)d(x_2).$$

We rescale $g = s \otimes d$ to a dyadic square $Q = I \times J$ as follows. Let l_I, l_J denote the left endpoint of I respectively J. Put

$$s_I(x_1) = s(\frac{x_1 - l_I}{|I|}), \quad d_J(x_2) = d(\frac{x_2 - l_J}{|J|}),$$

and

$$g_Q(x) = s_I(x_1)d_J(x_2). \tag{7.7}$$

We next define the testing function f_ϵ that is obtained by first forming “almost independent” copies of $g = s \otimes d$ and then adding $\frac{1}{\epsilon}$ of those. Below we define a collection of dyadic squares $\mathcal{G}$ and form

$$f_\epsilon = \sum_{Q \in \mathcal{G}} g_Q. \tag{7.8}$$

To define $\mathcal{G}$ we proceed as follows. Fix $j \in \mathbb{N}$. Let $\mathcal{D}_j$ denote the collection of dyadic intervals I satisfying

$$I \subseteq [0,1] \quad \text{and} \quad |I| = 2^{-j}.$$

Let $\mathcal{L}_j \subseteq \mathcal{D}_j$ satisfy

$$I, J \in \mathcal{L}_j \quad \text{implies} \qquad \operatorname{dist}(I, J) \geq |I|, \tag{7.9}$$

and

$$\sum_{J \in \mathcal{L}_j} |J| = \frac{1}{2}. \tag{7.10}$$

To define $\mathcal{L}_j$ simply take the even numbered intervals of $\mathcal{D}_j$, counting from left to right. Next assume that $\epsilon > 0$ is power of 1/2, thus

$$\epsilon = 2^{-n_0} \quad \text{for some} \qquad n_0 \in \mathbb{N}. \tag{7.11}$$

For $1 \leq k \leq 1/\epsilon$ put

$$\mathcal{G}_k = \bigcup \{I \times J : I \in \mathcal{D}_{2kn_0},\ J \in \mathcal{L}_{2kn_0}\} \quad \text{and} \quad \mathcal{G} = \bigcup_{k=1}^{1/\epsilon} \mathcal{G}_k.$$

Observe that $|Q| = \epsilon^{4k}$ for $Q \in \mathcal{G}_k$, and by (7.10)

$$\sum_{Q \in \mathcal{G}_k} |Q| = \frac{1}{2} \quad \text{and} \quad \sum_{Q \in \mathcal{G}} |Q| = \frac{1}{2\epsilon}. \tag{7.12}$$

7.2 Upper estimate for $\|f_\epsilon\|_p$ and $\|R_1(f_\epsilon)\|_p$.

We obtain our L^p estimates of f_ϵ by proving an upper bound for its norm in the space dyadic BMO. These in turn follow from scale-invariant L^2 estimates and " almost orthogonality" of the functions

$$\sum_{Q \in \mathcal{G}_k} g_Q, \qquad k \leq \frac{1}{\epsilon}.$$

Proposition 7.2 *Let f_ϵ be defined by (7.8). The support of f_ϵ is contained in $[-1,1] \times [-1,1]$ and*

$$\|f_\epsilon\|_{BMO_d} \leq C. \tag{7.13}$$

Hence $\|f_\epsilon\|_p \leq C_p$.

Proof. Let $Q_0 \in \mathcal{G}$ and form $g = \sum_{\{Q \in \mathcal{G},\, Q \subseteq Q_0\}} g_Q$. The BMO_d inequality (7.13) is a consequence of uniform L^2 estimate

$$\|g\|^2_{L^2(\mathbb{R}^2)} \leq C|Q_0|, \tag{7.14}$$

in combination with the Lipschitz estimates,

$$\sum_{\{Q \in \mathcal{G},\, |Q| > |Q_0|\}} \|1_{Q_0}(g_Q - m_{Q_0}(g_Q))\|_{L^2(\mathbb{R}^2)} \leq C\epsilon |Q_0|^{1/2}, \tag{7.15}$$

where $m_{Q_0}(g_Q) = |Q_0|^{-1} \int_{Q_0} g_Q$. In two separate paragraphs below we will verify that (7.14) and (7.15) hold. Before that we show how these estimates yield (7.13). Let

$$\mathcal{K} = \{W \in \mathcal{S} : \exists \varepsilon \ \langle f_\epsilon, h_W^{(\varepsilon)} \rangle \neq 0\}$$

Let W be a dyadic square with $|W| \leq 1/4$, then $\int_W f_\epsilon = 0$. Hence for $W \in \mathcal{K}$, $\operatorname{diam}(W) \leq 1$. By (3.26), to estimate the BMO_d norm of f_ϵ it suffices to test the cubes of $\mathcal{K}$. Next we fix a dyadic square $W \in \mathcal{K}$. Since $\operatorname{diam}(W) \leq 1$ we may choose $k \in \mathbb{N}_0$ such that $\epsilon^{2(k+1)} \leq \operatorname{diam}(W) \leq \epsilon^{2k}$. Define a decomposition of $\mathcal{G}$ as $\mathcal{G} = \mathcal{H}_1 \cap \mathcal{H}_2 \cup \mathcal{H}_3$ where

$$\mathcal{H}_1 = \{Q \in \mathcal{G} : \operatorname{diam}(Q) = \epsilon^{2k},\ Q \cap 2 \cdot W \neq \emptyset\},$$

$$\mathcal{H}_2 = \{Q \in \mathcal{G} : \operatorname{diam}(Q) \geq \epsilon^{2(k-1)},\ Q \cap 2 \cdot W \neq \emptyset\},$$

and

$$\mathcal{H}_3 = \{Q \in \mathcal{G} : \operatorname{diam}(Q) \leq \epsilon^{2(k+1)},\ Q \cap 2 \cdot W \neq \emptyset\}$$

Accordingly let

$$g_j = \sum_{Q \in \mathcal{H}_j} g_Q, \quad j \in \{1, 2, 3\}.$$

The cardinality of $\mathcal{H}_1$ is bounded by C. Hence $\|1_W g_1\|_2 \leq C|W|^{1/2}$. With $A = |W|^{-1} \int_W g_2$, and triangle inequality (7.15) gives $\int_W |g_2 - A|^2 \leq C\epsilon^2 |W|$. The estimate (7.14) implies $\|1_W g_3\|_2 \leq C|W|^{1/2}$. To see this let $\mathcal{M}$ denote the maximal squares of $\mathcal{H}_3$. The collection $\mathcal{M} (\subseteq \mathcal{H}_3)$ consists of pairwise disjoint squares so that

$$\sum_{Q_0 \in \mathcal{M}} |Q_0| \leq C|W|.$$

Next write $G_{Q_0} = \sum_{Q \in \mathcal{H}_3, Q \subseteq Q_0} g_Q$, to obtain

$$g_3 = \sum_{Q_0 \in \mathcal{M}} G_{Q_0} \quad \text{and} \quad \|g_3\|_2^2 = \sum_{Q_0 \in \mathcal{M}} \|G_{Q_0}\|_2^2.$$

Apply (7.14) to G_{Q_0} to obtain

$$\begin{aligned} \|g_3\|_2^2 &\leq C \sum_{Q_0 \in \mathcal{M}} |Q_0| \\ &\leq C|W|. \end{aligned}$$

Finally $\|1_W g_3\|_2 \leq \|g_3\|_2 \leq C|W|^{1/2}$.

Moreover for $t \in W$ there holds the identity

$$f_\epsilon(t) = g_1(t) + g_2(t) + g_3(t).$$

Invoking the estimates for g_1, g_2, g_3 we obtain

$$\int_W |f_\epsilon - A|^2 \leq C|W|.$$

By (3.25) this estimate yields (7.13).

■

Verification of (7.14)**.** By rescaling it suffices to consider $Q_0 = [0,1] \times [0,1]$. For $Q, Q' \in \mathcal{G}$ with $|Q| = |Q'|$ and $Q \neq Q'$ we have $\langle g_Q, g_{Q'} \rangle = 0$. Hence the left hand side of (7.14) equals

$$\sum_{Q \in \mathcal{G}} \langle g_Q, g_Q \rangle + 2 \sum_{\{Q, Q' \in \mathcal{G} :\, |Q| < |Q'|\}} \langle g_Q, g_{Q'} \rangle. \tag{7.16}$$

In view of (7.16) we aim at estimates for the entries of the Gram matrix $\langle g_Q, g_{Q'} \rangle$.

We first treat the diagonal terms of the Gram matrix. A direct calculation gives $\langle g_Q, g_Q\rangle = \epsilon|Q|/4$, hence by (7.12)

$$\sum_{Q\in\mathcal{G}} \langle g_Q, g_Q\rangle \le C. \tag{7.17}$$

Next we turn to estimating the off diagonal terms. Consider $Q, Q' \in \mathcal{G}$ such that $|Q| < |Q'|$. Write $Q = I\times J$ and $Q' = I'\times J'$. Note, first if $\operatorname{dist}(Q,Q') \ge 2 \operatorname{diam}(Q')$ then $\langle g_Q, g_{Q'}\rangle = 0$. Hence it remains to consider the case $\operatorname{dist}(Q,Q') \le 2\operatorname{diam}(Q')$. Let l_I denote the left endpoint of I. The Lipschitz estimate $\operatorname{Lip}(s_{I'}) \le C|I'|^{-1}$ and that $\int |d_J(x_2)|dx_2 \le \epsilon|J|$ imply that

$$\begin{aligned}
|\langle g_Q, g_{Q'}\rangle| &= \left|\int (s_{I'}(x_1) - s_{I'}(l_I))s_I(x_1)dx_1\right| \cdot \left|\int d_{J'}(x_2)d_J(x_2)dx_2\right| \\
&\le C\frac{|I|}{|I'|}|I|\int |d_J(x_2)|dx_2 \\
&\le \epsilon C\frac{|I|}{|I'|}|Q|.
\end{aligned} \tag{7.18}$$

Since $Q = I\times J \in \mathcal{G}$ there exists $k\in\mathbb{N}$ so that $|I| = \epsilon^{2k}$. Hence for $Q' = I'\times J' \in \mathcal{G}$ with $|Q'| > |Q|$ there exists $k'\in\mathbb{N}$ with $k' \le k-1$ so that $|I'| = \epsilon^{2k'}$, and $|I|/|I'| = \epsilon^{2k-2k'}$. Note that for each $Q\in\mathcal{G}$ the cardinality of the set

$$\{Q'\in\mathcal{G} :\ |Q| < |Q'|,\ \langle g_Q, g_{Q'}\rangle \ne 0\}$$

is bounded by C_1, say. Consequently in the double sum appearing on the left hand side of (7.19), for each Q only C_1 cubes Q' give a contribution. Thus by (7.18)

$$\begin{aligned}
\sum_{\{Q,Q'\in\mathcal{G}:\, |Q|<|Q'|\}} |\langle g_Q, g_{Q'}\rangle| &\le C\epsilon^{2k+1}\sum_{k'=1}^{k-1}\epsilon^{-2k'}\sum_{Q\in\mathcal{G}}|Q| \\
&\le C\epsilon^3\sum_{Q\in\mathcal{G}}|Q|.
\end{aligned} \tag{7.19}$$

By (7.12) the last line in (7.19) is bounded by $C\epsilon^2$. Combining (7.17) and (7.19) gives (7.14).

Verification of (7.15)**.** Fix $Q, Q_0 \in \mathcal{G}$ so that $|Q_0| < |Q|$ and $\operatorname{dist}(Q,Q_0) \le C\operatorname{diam}(Q)$. Then

$$\|1_{Q_0}(g_Q - m_{Q_0}(g_Q))\|_2 \le C\operatorname{Lip}(g_Q)\operatorname{diam}(Q_0)|Q_0|^{1/2}. \tag{7.20}$$

Moreover if $Q, Q_0 \in \mathcal{G}$ so that $|Q_0| < |Q|$ and $\operatorname{dist}(Q,Q_0) \ge C\operatorname{diam}(Q)$, then

$$\|1_{Q_0}(g_Q - m_{Q_0}(g_Q))\|_2 = 0. \tag{7.21}$$

Note that $\operatorname{Lip}(g_Q) \le C(\epsilon\operatorname{diam}(Q))^{-1}$. Since $Q, Q_0\in\mathcal{G}$, with $|Q_0| < |Q|$, there exists $k, k_0 \in \mathbb{N}$, with $k \le k_0 - 1$ so that $\operatorname{diam}(Q_0) = \sqrt{2}\cdot\epsilon^{2k_0}$ and $\operatorname{diam}(Q) = \sqrt{2}\cdot\epsilon^{2k}$. The cardinality of

$$\{Q\in\mathcal{G} :\ \operatorname{diam}(Q) = \sqrt{2}\cdot\epsilon^{2k},\quad \operatorname{dist}(Q,Q_0) \le C\sqrt{2}\cdot\epsilon^{2k}\}$$

is bounded by a constant C. Hence by (7.20) and (7.21),

$$\sum_{\{Q\in\mathcal{G},\, |Q|>|Q_0|\}} \|1_{Q_0}(g_Q - m_{Q_0}(g_Q))\|_2 \le C\epsilon|Q_0|^{1/2}.$$

Thus we verified (7.15).

■

We emphasize that the above upper bound on $\|f_\epsilon\|_p$ works when the test functions $g = s \otimes d$ and its rescalings $g_Q = s_I \otimes d_J$ are defined with Lipschitz functions A, B satisfying (7.6), that is, supp $A \subseteq [0,1]$, $\int A = 0$ and supp $B \subseteq [-1,1]$. We next impose furthermore that

$$\text{A}' \text{ is Lipschitz and } \int B = 0. \tag{7.22}$$

Proposition 7.3 *Let f_ϵ be defined by* (7.8)*, assume that* (7.22) *and* (7.6) *hold. Then for* $1 < p < \infty$,

$$\|R_1(f_\epsilon)\|_p \le C_p \epsilon.$$

PROOF. The Fourier multipliers of the Riesz transforms R_1 respectivley R_2 are $\xi_1/|\xi|$ and $\xi_2/|\xi|$. Hence using (7.22) for $g_Q = s_I \otimes d_J$ we have the identity

$$R_1(g_Q) = R_2(\partial_1 \mathbb{E}_2 g_Q), \tag{7.23}$$

where ∂_1 is differentiation with respect to the variable x_1 and $\mathbb{E}_2 g_Q(x_1, x_2) = \int_{-\infty}^{x_2} g_Q(x_1, s) ds$. Define now

$$\tilde{s}(x_1) = A'(x_1), \quad \tilde{d}(x_2) = C(x_2/\epsilon), \quad C(t) = \int_{-\infty}^{t} B(s) ds.$$

Let $\tilde{s}_I, \tilde{d}_J$ be obtained from $\tilde{s}(x_1), \tilde{d}(x_2)$ by rescaling,

$$\tilde{s}_I(x_1) = \tilde{s}(\frac{x_1 - l_I}{|I|}), \quad \tilde{d}_J(x_2) = \tilde{d}(\frac{x_2 - l_J}{|J|}),$$

where l_I, l_J denote the left endpoint of I respectively J. Then with $\tilde{g}_Q = \tilde{s}_I \otimes \tilde{d}_J$ the identity (7.23) assumes the following form,

$$R_1(g_Q) = \epsilon R_2(\tilde{g}_Q). \tag{7.24}$$

By (7.22) the Lipschitz functions A', C satisfy (7.6). Hence Proposition 7.2 implies that $\tilde{f}_\epsilon = \sum_{Q \in \mathcal{G}} \tilde{g}_Q$ satisfies the L^p estimate

$$\|\tilde{f}_\epsilon\|_p \le C_p.$$

By (7.24) we have $R_1(f_\epsilon) = \epsilon R_2(\tilde{f}_\epsilon)$. Hence the L^p boundedness of the Riesz transforms yields

$$\begin{aligned} \|R_1(f_\epsilon)\|_p &\le \epsilon \|R_2(\tilde{f}_\epsilon)\|_p \\ &\le C_p \epsilon \|\tilde{f}_\epsilon\|_p \\ &\le C_p \epsilon. \end{aligned}$$

■

We remark that the proof given above containd the following estimates estimates that we will use again later. For $g = s \otimes d$ and $\tilde{g} = \tilde{s} \otimes \tilde{d}$,

$$\begin{aligned} \|R_1(g)\|_p &= \epsilon \|R_2(\tilde{g})\|_p \\ &\le \epsilon C_p \|\tilde{g}\|_p \\ &\le C_p \epsilon^{1+1/p}. \end{aligned} \tag{7.25}$$

7.3 Lower bound for $\|P(f_\epsilon)\|_p$, $p \geq 2$.

We first specialize once more the class of Lipschitz functions A, B we use to define

$$s(x_1) = A(x_1), \quad d(x_2) = B(x_2/\epsilon)$$
$$g = s \otimes d \quad \text{and} \quad f_\epsilon = \sum_{Q \in \mathcal{G}} g_Q.$$

We simply take now

$$B(x_2) = \begin{cases} \sin(\pi x_2) & x_2 \in [-1,1]; \\ 0 & x_2 \in \mathbb{R} \setminus [-1,1]. \end{cases}$$

and choose A to be smooth, so that supp $A \subseteq [0,1]$, $\int A = 0$ and

$$\int_0^1 A(x_1) h_{[0,1]}(x_1) dx_1 = \int_0^1 \sin(2\pi x_1) h_{[0,1]}(x_1) dx_1.$$

The following list of identities relates the Haar functions $\{h_Q^{(1,0)}\}$ to the test functions $\{g_Q\}$.

1. The scalar products $\langle g_Q, h_Q^{(1,0)} \rangle$ and $\langle g_Q, g_Q \rangle$ are as follows,

$$\int g_Q(x) h_Q^{(1,0)}(x) dx = \epsilon \frac{4|Q|}{\pi^2} \quad \text{and} \quad \int g_Q(x) g_Q(x) dx = \epsilon \frac{|Q|}{4}. \tag{7.26}$$

2. Let $Q' = I \times J'$, be a dyadic square where J' is the dyadic interval adjacent to J so that the right endpoint of J is the left endpoint of J'. Then

$$\begin{aligned} \int g_{Q'}(x) h_Q^{(1,0)}(x) dx &= -\int g_Q(x) h_Q^{(1,0)}(x) dx \\ &= -\epsilon \frac{4|Q|}{\pi^2}. \end{aligned} \tag{7.27}$$

3. For all choices of $Q' = I \times J'$ with $|J'| = |J|$ and $\operatorname{dist}(J, J') \geq |J|$ we have

$$\int g_{Q'}(x) h_Q^{(1,0)}(x) dx = 0. \tag{7.28}$$

4. If $Q, Q' \in \mathcal{S}$ so that $|Q'| < |Q|$ then

$$\int g_{Q'}(x) h_Q^{(1,0)}(x) = 0. \tag{7.29}$$

We consider $p \geq 2$. Since $P(f_\epsilon)$ is compactly supported, lower L^p estimates for $P(f_\epsilon)$ result from lower L^2 estimates. We obtain the latter by exploiting again the fact that $\{g_Q : Q \in \mathcal{G}\}$ is an "almost orthogonal" family of functions.

Proposition 7.4 *Let f_ϵ be defined by (7.8). The support of $P(f_\epsilon)$ is contained in $[-1,1] \times [-1,1]$ and*

$$\|P(f_\epsilon)\|_2 \geq c\epsilon^{1/2}. \tag{7.30}$$

Hence for $p \geq 2$, $\|P(f_\epsilon)\|_p \geq c\epsilon^{1/2}$.

PROOF. By Bessel's inequality,

$$\sum_{Q\in\mathcal{G}}\langle f_\epsilon, h_Q^{(1,0)}\rangle^2|Q|^{-1} \le \|P(f_\epsilon)\|_2^2. \tag{7.31}$$

Using (7.31) and (7.12) we prove below that (7.30) follows from the following lower estimate for the Haar coefficients

$$|\langle f_\epsilon, h_Q^{(1,0)}\rangle| \ge c\epsilon|Q| \quad \text{for} \quad Q\in\mathcal{G}. \tag{7.32}$$

To prove (7.32), fix a dyadic square $Q = I\times J$ with $Q\in\mathcal{G}$. Write the Haar coefficient as

$$\langle f_\epsilon, h_Q^{(1,0)}\rangle = \langle g_Q, h_Q^{(1,0)}\rangle + \sum_{Q'\in\mathcal{G}\setminus\{Q\}}\langle g_{Q'}, h_Q^{(1,0)}\rangle. \tag{7.33}$$

Recall (7.26) asserting that

$$\langle g_Q, h_Q^{(1,0)}\rangle = \epsilon 4|Q|/\pi^2.$$

Next we show that the off diagonal terms in (7.33) are negligible compared to $\langle g_Q, h_Q^{(1,0)}\rangle$. We claim,

$$\sum_{Q'\in\mathcal{G}\setminus\{Q\}}|\langle g_{Q'}, h_Q^{(1,0)}\rangle| \le C\epsilon^2|Q|. \tag{7.34}$$

The first step in the verification of the claim consists in observing that the only contribution to (7.34) comes from the index set $\{Q'\in\mathcal{G}\setminus\{Q\} : |Q'|>|Q|\}$. Indeed, if $Q'\in\mathcal{G}$, $Q'\ne Q$ and $|Q'| = |Q|$ then (7.28) in combination with (7.9) implies that $\langle g_{Q'}, h_Q^{(1,0)}\rangle = 0$. Also by (7.29) for $Q'\in\mathcal{G}$ and $|Q'|<|Q|$ we have $\langle g_{Q'}, h_Q^{(1,0)}\rangle = 0$.

Next we provide an estimate for the contribution to (7.34) coming from $\{Q'\in\mathcal{G}\setminus\{Q\} : |Q'|>|Q|\}$. Choose $k\in\mathbb{N}$ so that $|Q| = \epsilon^{4k}$ and let $k'\in\mathbb{N}$ satisfy $k'<k$. There exists at most one square $Q'\in\mathcal{G}$ satisfying

$$|Q'| = \epsilon^{4k'} \quad \text{and} \quad \langle g_{Q'}, h_Q^{(1,0)}\rangle \ne 0.$$

Next fix $Q' = I'\times J'$ with $|Q'| = \epsilon^{4k'}$ and $k'<k$. Write $Q = I\times J$ and $Q' = I'\times J'$. Let l_I denote the left endpoint of I. Recall that $\mathrm{Lip}(s_{I'}) \le C|I'|^{-1}$ and $\int |d_J(x_2)|dx_2 \le C|J|$. Hence,

$$\begin{aligned}|\langle g_{Q'}, h_Q^{(1,0)}\rangle| &= \left|\int (s_{I'}(x_1) - s_{I'}(l_I))h_I(x_1)dx_1\right| \cdot \left|\int_J d_{J'}(x_2)dx_2\right| \\ &\le C|I|\cdot|I'|^{-1}|Q| \\ &= C\epsilon^{2k-2k'}|Q|.\end{aligned} \tag{7.35}$$

By definition of $g_{Q'}$ and $h_Q^{(1,0)}$ if $|Q'|>|Q|$ and $\langle g_{Q'}, h_Q^{(1,0)}\rangle \ne 0$ then $\mathrm{dist}(Q',Q) \le C\ \mathrm{diam}(Q')$. It now follows from (7.35) that for any $Q\in\mathcal{G}$,

$$\begin{aligned}\sum_{\{Q'\in\mathcal{G},\, |Q'|>|Q|\}}|\langle g_{Q'}, h_Q^{(1,0)}\rangle| &\le C|Q|\epsilon^{2k}\sum_{k'=1}^{k-1}\epsilon^{-2k'} \\ &\le C\epsilon^2|Q|.\end{aligned} \tag{7.36}$$

Thus by (7.36) we verified the claim (7.34). Hence we have (7.32). It remains to show how the coefficient estimates (7.32) imply the norm inequality of (7.30). Using first (7.31) then (7.32) and (7.12) we obtain

$$\begin{aligned}\|P(f_\epsilon)\|_2^2 &\geq c\epsilon^2 \sum_{Q\in\mathcal{G}} |Q| \\ &\geq c\epsilon.\end{aligned}$$

∎

7.4 The proof of theorem 7.1 .

We choose Lipschitz functions A, B with specification of the previous sub-section and define testing functions $g = s_{[0,1]} \otimes d_{[0,1]}$, f_ϵ as above.

Consider first the estimate (7.4) of Theorem 7.1. Let $1 < p \leq 2$. Fix $\eta > 0$. Let $g = s_{[0,1]} \otimes d_{[0,1]}$ be defined by (7.7). Since g is bounded and supported in $[0,1] \times [-\epsilon, \epsilon]$, we have

$$\|g\|_p \leq C\epsilon^{1/p}. \tag{7.37}$$

Next observe that for the square function $\mathbb{S}(P(g))$ we have the obvious estimate $\mathbb{S}(P(g)) \geq |\langle g, h^{(1,0)}_{[0,1[\times[0,1[}\rangle|$. Next recall that $\|P(g)\|_p \sim \|\mathbb{S}(P(g))\|_p$ hence $\|P(g)\|_p \geq c|\langle g, h^{(1,0)}_{[0,1[\times[0,1[}\rangle|$. By (7.26), we have $\langle g, h^{(1,0)}_{[0,1[\times[0,1[}\rangle = 4\epsilon/\pi^2$, hence

$$||P(g)||_p \geq c\epsilon. \tag{7.38}$$

By (7.37) and (7.25)

$$\|g\|_p^{1/p-\eta}\|R_1(g)\|_p^{1/q+\eta} \leq C\epsilon^{1+\eta}. \tag{7.39}$$

Combining (7.38) and (7.39) yields

$$\frac{||P(g)||_p}{\|g\|_p^{1/p-\eta}\|R_1(g)\|_p^{1/q+\eta}} \geq c\epsilon^{-\eta}.$$

Since $\eta > 0$ is fixed and $\epsilon > 0$ is arbitrarily small we verified (7.4).

Next we turn to the case $p \geq 2$. The test function f_ϵ is defined by (7.8). Proposition 7.2 and Proposition 7.3 give the upper bounds

$$\|f_\epsilon\|_p \leq C_p \quad \text{and} \quad \|R_1(f_\epsilon)\|_p \leq C_p\epsilon.$$

Hence for $\eta > 0$

$$\|f_\epsilon\|_p^{1/2-\eta}\|R_1(f_\epsilon)\|_p^{1/2+\eta} \leq C_p\epsilon^{1/2+\eta}.$$

By Proposition 7.4 we have the lower estimate

$$\|P(f_\epsilon)\|_p \geq c_p\epsilon^{1/2}.$$

so that

$$\frac{\|P(f_\epsilon)\|_p}{\|f_\epsilon\|_p^{1/2-\eta}\|R_1(f_\epsilon)\|_p^{1/2+\eta}} \geq c_p\epsilon^{-\eta}.$$

∎

Acknowledgement. During the preparation of this paper J. Lee and P. Müller visited repeatedly the Max Planck Insitute for Mathematics in the Sciences in Leipzig. It is our pleasure to thank this institution for its hospitality and for providing us with excellent working conditions.

References

[1] C. Bennett and R. Sharpley. *Interpolation of Operators*, volume 129 of *Pure and Applied Mathematics*. Academic Press, 1988.

[2] Z. Ciesielski. Haar orthogonal functions in analysis and probability. In *A. Haar memorial conference, Vol. I, II (Budapest, 1985)*, volume 49 of *Colloq. Math. Soc. János Bolyai*, pages 25–56. North-Holland, Amsterdam, 1987.

[3] B. Dacorogna. *Weak continuity and weak lower semicontinuity of nonlinear functionals*, volume 922 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1982.

[4] T. Figiel. Singular integral operators: a martingale approach. In *Geometry of Banach spaces (Strobl, 1989)*, volume 158 of *London Math. Soc. Lecture Note Ser.*, pages 95–110. Cambridge Univ. Press, Cambridge, 1990.

[5] I. Fonseca and S. Müller. $\mathcal{A}$-quasiconvexity, lower semicontinuity, and Young measures. *SIAM J. Math. Anal.*, 30(6):1355–1390 (electronic), 1999.

[6] P. W. Jones. BMO and the Banach space approximation problem. *Amer. J. Math.*, 107(4):853–893, 1985.

[7] Y. Meyer and R. Coifman. *Wavelets*, volume 48 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1997. Calderón-Zygmund and multilinear operators, Translated from the 1990 and 1991 French originals by David Salinger.

[8] C. B. Morrey, Jr. Quasi-convexity and the lower semicontinuity of multiple integrals. *Pacific J. Math.*, 2:25–53, 1952.

[9] P. F. X. Müller. *Isomorphisms between H^1 spaces*, volume 66 of *Instytut Matematyczny Polskiej Akademii Nauk. Monografie Matematyczne (New Series) [Mathematics Institute of the Polish Academy of Sciences. Mathematical Monographs (New Series)]*. Birkhäuser Verlag, Basel, 2005.

[10] S. Müller. Rank-one convexity implies quasiconvexity on diagonal matrices. *Internat. Math. Res. Notices*, (20):1087–1095, 1999.

[11] S. Müller. Variational models for microstructure and phase transitions. In *Calculus of variations and geometric evolution problems (Cetraro, 1996)*, volume 1713 of *Lecture Notes in Math.*, pages 85–210. Springer, Berlin, 1999.

[12] F. Murat. Compacité par compensation. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 5(3):489–507, 1978.

[13] F. Murat. Compacité par compensation. II. In *Proceedings of the International Meeting on Recent Methods in Nonlinear Analysis (Rome, 1978)*, pages 245–256, Bologna, 1979. Pitagora.

[14] F. Murat. Compacité par compensation: condition nécessaire et suffisante de continuité faible sous une hypothèse de rang constant. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)*, 8(1):69–102, 1981.

[15] E. M. Stein. *Singular integrals and differentiability properties of functions.* Princeton Mathematical Series, No. 30. Princeton University Press, Princeton, N.J., 1970.

[16] L. Tartar. Une nouvelle méthode de résolution d'équations aux dérivées partielles non linéaires. In *Journées d'Analyse Non Linéaire (Proc. Conf., Besançon, 1977)*, volume 665 of *Lecture Notes in Math.*, pages 228–241. Springer, Berlin, 1978.

[17] L. Tartar. Compensated compactness and applications to partial differential equations. In *Nonlinear analysis and mechanics: Heriot-Watt Symposium, Vol. IV*, volume 39 of *Res. Notes in Math.*, pages 136–212. Pitman, Boston, Mass., 1979.

[18] L. Tartar. Compacité par compensation: résultats et perspectives. In *Nonlinear partial differential equations and their applications. Collège de France Seminar, Vol. IV (Paris, 1981/1982)*, volume 84 of *Res. Notes in Math.*, pages 350–369. Pitman, Boston, MA, 1983.

[19] L. Tartar. Some remarks on separately convex functions. In *Microstructure and phase transition*, volume 54 of *IMA Vol. Math. Appl.*, pages 191–204. Springer, New York, 1993.

[20] L. Tartar. H-measures and applications. In *Nonlinear partial differential equations and their applications. Collège de France Seminar, Vol. XI (Paris, 1989–1991)*, volume 299 of *Pitman Res. Notes Math. Ser.*, pages 282–290. Longman Sci. Tech., Harlow, 1994.

AMS Subject classification: 49J45 , 42C15, 35B35

Addresses:

Jihoon Lee
Department of Mathematics
Sungkyunkwan University
Suwon, Korea
E-mail:jihoonlee@skku.edu

Paul F.X. Müller
Institut für Analysis
J. Kepler Universität
A-4040 Linz
E-mail: pfxm@bayou.uni-linz.ac.at

Stefan Müller
Max Planck Institute for Mathematics in the Sciences
Inselstr. 22–26
D 04103 Leipzig Germany
E-mail: sm@mis.mpg.de