Sudan University of Science and Technology
College of Graduate Studies

Journé Commutators and Polynomial Approximations with Characterizations of Hardy Space and BMO

مبدلات جويرني وتقريبات كثيرة الحدود مع تشخيصات فضاء هاردي و BMO

A Thesis Submitted in Fulfillment for the Degree of Ph.D

in Mathematics

By:

Nissreen Alzobir Osman

Supervisor:

Prof. Dr. Shawgy Hussein AbdAlla

December-2019

Dedication

To my family.

Acknowledgements

I would like to thank with all sincerity Allah, and my family for their supports throughout my study. Many thanks are due to my thesis guide, Prof. Dr. Shawgy Hussein AbdAlla.

Abstract

A characterization of product and multi-paraproducts of bounded mean oscillations with upper bound for multi-parameter and higher order Journé commutators are studied. We describe an algorithm of nonlinear piecewise polynomial approximations by refinable functions, of functions of the classes W_p^{α} and multivariate bound variation spaces of Wiener-Young type. We show the classes and maximal function characterization of Hardy spaces associated with operators, nonnegative self-adjoint operators satisfying Gaussian estimates with characterizations of Hardy and bounded mean oscillations by weak factorizations and commutators.

الخلاصة

قمنا بدراسة التشخيص للضرب ومتعدد ضرب الفقرات لترجحات الوسط المحدود مع الحد الأعلى لأجل متعدد-المعامل ومبدلات جويرني ذات الرتبة العليا. تم وصف خوارزمية لتقريبات كثيرة الحدود متعددة التعريف غير الخطية بواسطة دوال قابلة التحسن ولدوال لعائلات W_p^α وفضاءات التغير المحدود متعدد المتغيرات لنوع واينر-ينق. أوضحنا العائلات وتشخيص الدالة الأعظمية لفضاءات هاردي المشاركة مع المؤثرات ومؤثرات المرافق-الذاتي غير السالبة المحققة تقديرات جاوسيان مع التشخيصات لهاردي ومرجحات الوسط المحدود بواسطة التحليل إلى عوامل ضعيفة ومبدلات.

Introduction

We establish a commutator estimate which allows one to concretely identify the product BMO space, $\mathrm{BMO}(\boldsymbol{R}_+^2 \times \boldsymbol{R}_+^2)$, of A. Chang and R. Fefferman, as an operator space on $L^2(\boldsymbol{R}^2)$. The one-parameter analogue of this result is a well-known theorem of Nehari[9].

We present a newly-developed version of the Up-and-Down Algorithm (UDA) designed for nonlinear approximation by piecewise polynomials. Several almost optimal results are obtained about N-term nonlinear approximation by dilated integer translates of a refinable function.

Let L be the infinitesimal generator of an analytic semigroup on $L^2(\mathbb{R}^n)$ with suitable upper bounds on its heat kernels. In Auscher, Duong, and McIntosh (2005) and Duong and Yan (2005), a Hardy space $H_L^1(\mathbb{R}^n)$ and a $\mathrm{BMO_L}(\mathbb{R}^n)$ space associated with the operator L were introduced and studied. We define a class of $H_L^p(\mathbb{R}^n)$ spaces associated with the operator L for a range of $p < 1$ acting on certain spaces of Morrey-Campanato functions defined in New Morrey-Campanato spaces associated with operators and applications by Duong and Yan (2005), and they generalize the classical $H^p(\mathbb{R}^n)$ spaces. Let L be a generator of a semigroup satisfying the Gaussian upper bounds. A new BMO_L space associated with L was recently introduced in [70] and [71]. We discuss applications of the new BMO_L spaces in the theory of singular integration.

We show that the product BMO space can be characterized by iterated commutators of a large class of Calder´on–Zygmund operators. The proof introduces some new paraproducts which have BMO estimates. We characterize L^p boundedness of iterated commutators of multiplication by a symbol function and tensor products of Riesz and Hilbert transforms. We obtain a two-sided norm estimate that shows that such operators are bounded on L^p if and only if the symbol belongs to the appropriate multiparameter BMO class. We extend the results to a much more intricate situation; commutators of multiplication by a symbol function and paraproduct-free Journé operators. We show that the boundedness of these commutators is also determined by the inclusion of their symbol function in the same multiparameter BMO class.

We investigate the order of approximation of functions of the Sobolev-Slobodeckiĭ classes $W_p^a(Q^m)$($Q^m$is the m-dimensional unit cube) by piecewise-polynomial functions. The named space denoted by V_{pq}^k consists of L_q functions on $[0,1)^d$ of bounded p-variation of order $k \in \mathbb{N}$. It generalizes the classical spaces $V_p(0,1)$ $(= V^1_{p\infty})$ and BV $([0,1)^d)$ (V^1_{1q} where $q := \frac{d}{d-1}$) and is closely related to several important smoothness spaces, e.g., to Sobolev spaces over L_p, BV and BMO and to Besov spaces.

For a nonnegative, self-adjoint operator satisfying Gaussian estimates on $L^2(\mathbb{R}^n)$. We give an atomic decomposition for the Hardy spaces $H^p_{L,\max}(\mathbb{R}^n)$ in terms of the nontangential maximal functions associated with the heat semigroup of self-adjoint operator. We provide a deeper study of the Hardy and BMO spaces associated to the Neumann Laplacian Δ_N . For the Hardy space $H^1_{\Delta_N}(\mathbb{R}^n)$ (which is a proper subspace of the classical Hardy space $H^1(\mathbb{R}^n)$) we demonstrate that the space has equivalent norms in terms of Riesz transforms, maximal functions, atomic decompositions, and weak factorizations.

The Contents

Chapter 1
Characterization with Multi-Parameter

We discuss a situation governed by a two- parameter family of dilations, and so the spaces H^1 and BMO have a more complicated structure. We show that the classical Coifman-Meyer theorem holds on any polydisc $\boldsymbol{T}^d$ of arbitrary dimension $d \geq 1$.

Section (1.1): Product BMO by Commutators

Here $\boldsymbol{R}_+^2$denotes the upper half-plane and $BMO(\boldsymbol{R}_+^2 \times \boldsymbol{R}_+^2)$ is defined to be the dual of the real-variable Hardy space H^1 on the product domain $\boldsymbol{R}_+^2 \times \boldsymbol{R}_+^2$. There are several equivalent ways to define this latter space, see [6] for the various characterizations. We will be more interested in the bi-holomorphic analogue of H^1, which can be defined in terms of the boundary values of bi-holomorphic functions on $\boldsymbol{R}_+^2 \times \boldsymbol{R}_+^2$and will be denoted throughout by $H^1(\boldsymbol{R}_+^2 \times \boldsymbol{R}_+^2)$ cf.[11].

In one variable, the space $L^2(\boldsymbol{R})$ decomposes as the direct sum $H^2(\boldsymbol{R}) \oplus \overline{H^2(\boldsymbol{R})}$, where $H^2(\boldsymbol{R})$ is defined as the boundary values of functions in $H^2(\boldsymbol{R}_+^2)$ and $\overline{H^2(\boldsymbol{R})}$ denotes the space of complex conjugate of functions in $H^2(\boldsymbol{R}_+^2)$. The space $\overline{L^2(\boldsymbol{R})}$, therefore, decomposes as the direct sum of the four spaces $H^2(\boldsymbol{R}) \otimes H^2(\boldsymbol{R})$,$\overline{H^2(\boldsymbol{R})} \otimes H^2(\boldsymbol{R})$, $H^2(\boldsymbol{R}) \otimes \overline{H^2(\boldsymbol{R})}$ and $\overline{H^2(\boldsymbol{R})} \otimes \overline{H^2(\boldsymbol{R})}$, where the tensor products are the Hilbert space tensor products. Let $P_{\pm,\pm}$ denote the orthogonal projection of $L^2(\boldsymbol{R}^2)$ onto the holomorphic/anti-holomorphic subspaces, in the first and second variables, respectively, and let H_j denote the one-dimensional Hilbert transform in the jth variable, $j = 1, 2$. In terms of the projections $P_{\pm,\pm}$

$$H_1 = P_{+,+} + P_{+,-} - P_{-,+} - P_{-,-} \text{and } H_2 = P_{+,+} + P_{-,+} - P_{+,-} - P_{-,-}.$$

The nested commutator deternfined by the function b is tile operator $\left[[M_b, H_1], H_2\right]$ acting on $L^2(\boldsymbol{R}^2)$, where, for a function b on the plane, we define $M_b f \coloneqq bf$.

In terms of the projections$P_{\pm\pm}$, it takes the form

$$\frac{1}{4}\left[[M_b, H_1], H_2\right] = P_{+,+} M_b P_{-,-} - P_{+,-} M_b P_{-,+} - P_{-,+} M_b P_{+,-} + P_{-,-} M_b P_{+,+}. \qquad (1)$$

Ferguson and Sadosky [5] established the inequality $\left\|\left[[M_b, H_1], H_2\right]\right\|_{L^2} \leq c\|b\|_{BMO}$. The main result is the converse inequality.

Theorem (1.1.1)[1]: There is a constant $c > 0$ such that $\|b\|_{BMO} \leq c\left\|\left[[M_b, H_1], H_2\right]\right\|_{L^2 \to L^2}$ for all functions b in $BMO(\boldsymbol{R}_+^2 \times \boldsymbol{R}_+^2)$.

As A. Chang and R. Fefferman have established for us, the structure of the space BMO is more complicated in the two-parameter setting, requiring a more subtle approach to this theorem, despite the superficial similarity of the results to the one-parameter setting. The proof relies on three key ideas. The first is the dyadic characterization of the BMO norm given in [2]. The second is a variant of Journ's lemma, [6]. The third idea is that we have the estimates, the second of which was shown in [5],

$$\|b\|_{BOM(\text{rec})} \leq c\left\|\left[[M_b, H_1], H_2\right]\right\|_{L^2 \to L^2} \leq c'\|b\|_{BMO}.$$

An unpublished example of$L$. Carleson shows that the rectangular BMO norm is not comparable to the BMO norm, [4]. We may assume that the rectangular BMO norm of the function b is small. Indeed, this turns out to be an essential aspect of tile argument.

From Theorem (1.1.1) we deduce a weak factorization for the (biholomorphic) space $H^1(\boldsymbol{R}_+^2 \times \boldsymbol{R}_+^2)$. The idea is that if the function b has biholomorphic extension to $\boldsymbol{R}_+^2 \times \boldsymbol{R}_+^2$then for fimctions $f, g \in L^2(\boldsymbol{R}^2)$,

$$\frac{1}{4}\langle[[M_b, H_1], H_2]f, g\rangle = \langle b, \overline{P_{-,-}f}P_{+,+}g\rangle.$$

So in this case, the operator norm of the nested commutator$[[M_b, H_1], H_2]$is comparable to the dual norm

$$\|b\|_* := \sup|\langle f_g, b\rangle|,$$

where the supremum above is over all pairs f, g in the unit ball of $H^2(\boldsymbol{R}_+^2 \times \boldsymbol{R}_+^2)$. On the other hand, since$\|b\|_{BMO}$ and$\|[[M_b, H_1], H_2]\|_{L^2\to L^2}$ are comparable, the dual norm above satisfies

$$\|b\|_* \sim \sup|\langle h, b\rangle|,$$

where the supremum is over all functions h in the unit ball of$H^1(\boldsymbol{R}_+^2 \times \boldsymbol{R}_+^2)$.

Corollary (1.1.2)[1]: Let h be in$H^1(\boldsymbol{R}_+^2 \times \boldsymbol{R}_+^2)$ with$\|h\|_1 = 1$. Then there exist functions $(f_j), (g_j) \subseteq H^2(\boldsymbol{R}_+^2 \times \boldsymbol{R}_+^2)$such that $h = \sum_{j=1}^{\infty} f_j g_jand\sum_{j=1}^{\infty}\|f_j\|_2\|g_j\|_2 \le c$.

We remark that the weak factorization above implies the analogous factorization forH^1 of the bidisk. Indeed, for all$1 \le p < \infty$, the map$u_p: H^p(\boldsymbol{R}_+^2 \times \boldsymbol{R}_+^2) \to H^p(D^2)$ defined by

$$(u_p f)(\mathcal{Z}, w) = \pi^{2/p}\left(\frac{2i}{1-\mathcal{Z}}\right)^{2/p}\left(\frac{2i}{1-w}\right)^{2/p} f(\alpha(\mathcal{Z}), \alpha(w)), \alpha(\lambda) := i\frac{1+\lambda}{1-\lambda},$$

is an isometry with isometric inverse

$$(u_p^{-1} g)(\mathcal{Z}, w) = \pi^{-2/p}\left(\frac{1}{\mathcal{Z}+i}\right)^{2/p}\left(\frac{1}{w+i}\right)^{2/p} g(\beta(\mathcal{Z}), \beta(w)), \beta(\lambda) := \frac{\lambda - i}{\lambda + i}.$$

The dual formulation of weak factorization for$H^1(D^2)$ is a Nehari theorem for the bidisk. Specifically, if$b \in H^2(D^2)$ then the little Hankel operator with symbol b is densely defined on$H^2(D^2)$ by the formula

$$\Gamma_b f = P_{-,-}(\bar{b}f).$$

By

$$\|\Gamma_b\| = \|[[M_{\bar{b}}, H_1], H_1]\|_{L^2\to L^2} \tag{1}$$

and thus, by Theorem(1.1.1), $\|\Gamma_b\|$is comparable to $\|b\|_{BMO}$, which, by definition, is just the norm of b acting on $H^1(D^2)$. So the boundedness of the Hankel operator Γ_b implies that there is a function $\phi \in L^\infty(T^2)$ such that $P_{+,+}\phi = b$. Several variations and complements on these themes in the one-parameter setting have been obtained by Coifman, Rochberg and Weiss [3].

We give the one-dimensional preliminaries for the proof of Theorem(1.1.1), and devoted to the proof of Theorem (1.1.1). One final remark about notation $A \lesssim B$ means that there is an absolute constant C for which $A \le CB$, $A \approx B$ means that $A \lesssim B$ and $B \lesssim A$.

We are indebted to Andreas Seeger.

Several factors conspire to make the one-dimensional case easier than the higher-dimensional case. Before proceeding with the higher-dimensional case, we make several comments about the one-dimensional case, comments that extend and will be useful.

Let H denote the Hilbert transform in one variable, $P_+ = \frac{1}{2}(I+H)$ be the projection of $L^2(\boldsymbol{R})$ onto the positive frequencies, and P_- is $\frac{1}{2}(I-H)$ the projection onto the negative frequencies. We shall in particular rely upon the following basic computation:

$$\frac{1}{2}[M_b, H]\bar{b} = P_-|P_-b|^2 - P_+|P_+b|^2. \tag{2}$$

The frequency distribution of $|P_-b|^2$ is symmetric since it is real-valued. Thus,

$$\|b\|_4^2 \lesssim \|P_-|P_-b|^2 - P_+|P_+b|^2\|_2 \le \|[M_b, H]\|_{2\to 2}\|b\|_2.$$

Moreover, if b is supported on an interval I, we see that

$$\|b\|_2 \le |I|^{1/4}\|b\|_4 \lesssim |I|^{1/4}\|[M_b, H]\|_{2\to 2}^{1/2}\|b\|_2^{1/2},$$

which is the BMO estimate on I. We seek an extension of this estimate in the two-parameter setting.

We use a wavelet proof of Theorem (1.1.1), and specifically use a wavelet with compact frequency support constructed by Y. Meyer[8]. There is a Schwartz function w with these properties:

(a) $\|w\|_2 = 1$.

(b) $\widehat{w}(\xi)$ is supported on $\left[\frac{2}{3}, \frac{8}{3}\right]$ together with the symmetric interval about 0.

(c) $P_\pm w$ is a Schwartz function. We have

$$|w(x)|, \left|P_\pm w(x)\right| \lesssim (1+|x|)^{-n}, n \ge 1.$$

Let $\mathcal{D}$ denote a collection of dyadic intervals on $\boldsymbol{R}$. For any interval I, let $c(I)$ denote its center, and define

$$w_I(x) := \frac{1}{\sqrt{|I|}} w\left(\frac{x - c(I)}{|I|}\right).$$

Set $w_I^\pm := P_\pm w_I$. The central facts that we need about the functions $\{w_I : I \in \mathcal{D}\}$ are these: First, that these functions are an orthonormal basis on $L^2(\boldsymbol{R})$. Second, that we have the Littlewood-Paley inequalities, valid on all L^p, though $p = 4$ will be of special significance for us. These inequalities are

$$\|f\|_p \approx \left\|\left[\sum_{I\in\mathcal{D}} \frac{|\langle f, w_I\rangle|^2}{|I|} 1_I\right]^{1/2}\right\|_p, 1 < p < \infty. \tag{3}$$

Third, that the functions w_I have good localization properties in the spatial variables. That is,

$$|w_I(x)|, \left|w_I^\pm(x)\right| \lesssim |I|^{-1/2}\chi_1(x)^n, n \ge 1, \tag{4}$$

where $\chi_I(x) := (1 + \mathrm{dist}(x, I)/|I|)^{-1}$. We find the compact localization of the wavelets in frequency to be very useful. The price we pay for this utility below is the careful accounting of "Schwartz tails" we shall make in the main argument. Fourth, we have the identity below for the commutator of one w_I with a w_J. Observe that since P_+ is one half of $I + H$, it suffices to replace H by P_+ in the definition of the commutator.

$$w_{I,J} := [w_I, P_+]\overline{w_J} = w_I\overline{w_J^-} - P_+ w_I\overline{w_J} = P_- w_I\overline{w_J^-} - P_+ w_I\overline{w_J^+} = P_- w_I^- w_I\overline{w_J^-} - P_+ w_I^+\overline{w_J^+}$$

$$= \begin{cases} 0 \quad \text{if} |I| \ge 4|J|, \\ P_-|w_I^-|^2 - P_+|w_I^+|^2 \ \text{if} I = J, \\ w_I^-\overline{w_J^-} - w_I^+\overline{w_J^+} \text{if} |J| \ge 4|I|. \end{cases} \tag{5}$$

From this we see a useful point concerning orthogonality. For intervals I, I', J and J', assume$|J| \geq 8|I|$, and likewise forI'andJ'. Then

$$\operatorname{supp}(\widehat{w_I, J}) \cap \operatorname{supp}(\widehat{w_{I'}, J'}) = \emptyset, |I'| \geq 8|I|. \tag{6}$$

Indeed, this follows from a direct calculation. The positive frequency support of $w_I^+ \overline{w_J^+}$is contained in the interval $[(3|I|)^{-1}, 8(3|I|)^{-1}]$. Under the conditions on I and I', the frequency supports are disjoint.

$BMO(\boldsymbol{R}_+^2 \times \boldsymbol{R}_+^2)$ will denote the BMO of two parameters (or product BMO) defined as the dual of (real)$H^1(\boldsymbol{R}_+^2 \times \boldsymbol{R}_+^2)$. The following characterization of the space $BMO(\boldsymbol{R}_+^2 \times \boldsymbol{R}_+^2)$ is due to A. Chang and R. Fefferman [2].

The relevant class of rectangles is $\mathcal{R} = \mathcal{D} \times \mathcal{D}$ all rectangles which are products of dyadic intervals. These are indexed by $R \in \mathcal{R}$. For such a rectangle, write it as a product $R_1 \times R_2$ and then define

$$v_R(x_1, x_2) = w_{R_1}(x_1) w_{R_2}(x_2).$$

A function $f \in BMO(\boldsymbol{R}_+^2 \times \boldsymbol{R}_+^2)$and only if

$$\sup_U \left[|U|^{-1} \sum_{R \subset U} |\langle f, v_R \rangle|^2 \right]^{1/2} < \infty.$$

Here, the sum extends over those rectangles $R \in \mathcal{R}$, and the supremum is over all open sets inthe plane of finite measure. Note that the supremum is taken over a much broader class of sets than merely rectangles in the plane. We denote this supremum as$\|f\|_{BMO}$. In this definition, if the supremum over U is restricted to just rectangles, this defines the "rectangular BMO" space, and we denote this restricted supremum as$\|f\|_{BMO(rec)}$.

We make a further comment on the BMO condition. Suppose that for $R \in \mathcal{R}$, we have non-negative constants a_R for which

$$\sum_{R \subset U} a_R \leq |U|,$$

for all open sets U in the plane of finite measure. Then, we have the John-Nirenberg inequality

$$\left\| \sum_{R \subset U} |R|^{-1} a_R 1_R \right\|_p \lesssim |U|^{1/p}, 1 < p < \infty.$$

See [2]. This, with the Littlewood-Paley inequalities, will be used several times below, and referred to as the John-Nirenberg inequalities.

The function b may be taken to be of Schwartz class. By multiplying b by a constant, we can assume that the BMO norm of b is one. Set $B_{2\to2}$.to be the operator norm of $[[M_b, H_1]H_2]$. We provide a lower bound for $B_{2\to2}$. Let U be an open set of finite measure for which we have the equality

$$\sum_{R \subset U} |\langle b, v_R \rangle|^2 = |U|.$$

As b is of Schwartz class, such a set exists. By invariance under dilations by a factor of two, we can assume that$\frac{1}{2} \leq |U| \leq 1$. In several estimates below, the measure of U enters.

An essential point is that we may assume that the rectangular BMO norm of b is at most ε. The reason for this is that we have the estimate$\|b\|_{BMO(rec)} \lesssim B_{2\to2}$. See[5]. Therefore,

for a small constant ε to be chosen below, we can assume that $\|b\|_{BMO(rec)} \lesssim \varepsilon$, for otherwise we have a lower bound on $B_{2\to 2}$.
Associated to the set U is a set V which contains U and has the properties specified in **Lemma (1.1.3)[1]:** It is critical that the measure of V be only slightly larger than the measure of U, or more exactly,$|V| < (1+\delta)|U|$, for a choice of $0<\delta<1$ to be specified.
Define

$$\mu(R) \coloneqq \sup\{\mu: \mu R \subset V\}, R \subset U.$$

The quantity$\mu(R)$ #(R) measures how deeply a rectangle R is inside V. This quantity enters into the essential Journ's lemma, see [7].
In the argument below, we will be projecting b onto subspaces spanned by collections of wavelets. These wavelets are in turn indexed by collections of rectangles. Thus, for a collection $\mathcal{A} \subseteq \mathcal{R}$, let us denote

$$b^{\mathcal{A}} \coloneqq \sum_{R\in\mathcal{A}} (b, v_R)\, v_R.$$

The relevant collections of rectangles are defined as

$$\mathcal{U} \coloneqq \{R \in \mathcal{R}: R \subset U\},$$
$$\mathcal{V} = \{R \in \mathcal{R} - \mathcal{U}: R \subset V\},$$
$$\mathcal{W} = \mathcal{R} - \mathcal{U} - \mathcal{V}.$$

For functions f and g, we set$\{f,g\} \coloneqq \left[[M_f, H_1], H_2\right]\bar{g}$.
We will demonstrate that for all $\delta, \varepsilon > 0$ there is a constant $K_\delta > 0$ so that

(i) $\left\|\{b^{\mathcal{V}}, b^{\mathcal{U}}\}\right\|_2 \lesssim \delta^{1/4}$,

(ii) $\left\|\{b^{\mathcal{W}}, b^{\mathcal{U}}\}\right\|_2 \le K_\delta \varepsilon^{1/3}$.

Furthermore, we will show that $1 \lesssim \left\|\{b^{\mathcal{U}}, b^{\mathcal{U}}\}\right\|_2$. Since $b = b^{\mathcal{U}} + b^{\mathcal{V}} + b^{\mathcal{W}}$, $\left\|b^{\mathcal{U}}\right\|_2 \lesssim 1$ and $\delta, \varepsilon > 0$ are arbitrary, a lower bound on $B_{2\to 2}$ will follow from an appropriate choice of δ and ε. To be specific, one concludes the argument by estimating

$$1 \lesssim \left\|\{b^{\mathcal{U}}, b^{\mathcal{U}}\}\right\|_2 \lesssim \left\|\{b^{\mathcal{U}} + b^{\mathcal{V}}, b^{\mathcal{U}}\}\right\|_2 + \delta^{1/4}$$
$$\lesssim \left\|\{b^{\mathcal{U}} + b^{\mathcal{V}} + b^{\mathcal{W}}, b^{\mathcal{U}}\}\right\|_2 + \delta^{1/4} + K_\delta \varepsilon^{1/3} \lesssim B_{2\to 2} + \delta^{1/4} + K_\delta \varepsilon^{1/3}.$$

Implied constants are absolute. Choosing δ first and then εappropriately small supplies a lower bound on $B_{2\to 2}$.
The estimate $1 \lesssim \left\|\{b^{\mathcal{U}}, b^{\mathcal{U}}\}\right\|_2$relies on the John-Nirenberg inequality and the two-parameter version of (2), namely

$$\frac{1}{4}[[M_b, H_1], H_2]\tilde{b} = P_{+,+}\left|P_{+,+}b\right|^2 - P_{+,-}\left|P_{+,-}b\right|^2 - P_{-,+}\left|P_{-,+}b\right|^2 + P_{-,-}\left|P_{-,-}b\right|^2.$$

This identity easily follows from the one-variable identities. Here $P_{\pm,\pm}$ denotes the projection onto the positive/negative frequencies in the first and second variables. These projections are orthogonal and moreover, since$\left|P_{\pm,\pm}b\right|^2$is real-valued we have that$\left\|P_{\pm,\pm}\left|P_{\pm,\pm}b\right|^2\right\|_2 \ge \frac{1}{4}\left\|\left|P_{\pm,\pm}b\right|^2\right\|_2$Therefore,$\left\|b^{\mathcal{U}}\right\|_4^2 \lesssim \left\|\{b^{\mathcal{U}}, b^{\mathcal{U}}\}\right\|_2$. It follows that

$$1 \lesssim \left\|b^{\mathcal{U}}\right\|_2 = \left[\sum_{R\in\mathcal{U}} |\langle b, v_R\rangle|^2\right]^{1/2} \lesssim \left\|\left[\sum_{R\in\mathcal{U}} \frac{|\langle b, v_R\rangle|^2}{|R|} 1_R\right]^{1/2}\right\|_4 \lesssim \left\|b^{\mathcal{U}}\right\|_4 1$$
$$\lesssim \left\|\{b^{\mathcal{U}}, b^{\mathcal{U}}\}\right\|_2^{1/2}.$$

The estimate (i) relies on the estimate $|V| < (1+\delta)|U|$. Now, if $R \in \mathcal{V}$, then $R \subset V$ and since b has BMO norm one, it follows that

$$|U| + \left\|b^{\mathcal{V}}\right\|_2^2 = \sum_{R \in \mathcal{U} \cup \mathcal{V}} |\langle b, v_R \rangle|^2 \leq (1+\delta)|U|.$$

Hence $\left\|b^{\mathcal{V}}\right\|_2 \lesssim \delta^{1/2}$. Yet the BMO norm of $b^{\mathcal{V}}$ can be no more than that of b, which is to say one. Interpolating norms we see that $\left\|b^{\mathcal{V}}\right\|_4 \lesssim \delta^{1/4}$, and so

$$\left\|\{b^{\mathcal{V}}, b^{\mathcal{U}}\}\right\|_2 \lesssim \left\|b^{\mathcal{V}}\right\|_4 \lesssim \left\|b^{\mathcal{U}}\right\|_4 \delta^{1/4}.$$

We now turn to the estimate (ii). $b^{\mathcal{U}}$ and $b^{\mathcal{W}}$ live on disjoint sets. But in this argument we are trading off precise Fourier support of the wavelets for imprecise spatial localization, that is the "Schwartz tails" problem. Accounting for this requires a careful analysis, invoking several subcases.

A property of the commutator that we will rely upon is that it controls the geometry of R and R'. Namely, $\{v_{R'}, v_R\} \neq 0$ if and only if writing $R = R_1 \times R_2$ and likewise for R', we have for both $j = 1,2$, $|R'_j| \leq 4|R_j|$. This follows immediately from our one-dimensional calculations, in particular (5). We abbreviate this condition on R and R' as $R' \lesssim R$ and restrict our attention to this case. Orthogonality also enters into the argument. Observe the following. For rectangles $R^k, \tilde{R}^k, k = 1,2$, with $\tilde{R}^k \lesssim R^k$, and for $j = 1$ or $j = 2$,

$$\text{if } 8\left|\tilde{R}_j^1\right| \leq \left|R_j^1\right| \text{ and } 8\left|\tilde{R}_j^2\right| < \left|R_j^2\right|, \qquad \text{then } \langle v_{\tilde{R}^1}\overline{v_{R^1}}, v_{\tilde{R}^2}\overline{v_{R^2}}\rangle = 0. \tag{7}$$

This follows from applying (6) in the jth coordinate.

Therefore, there are different partial orders on rectangles that are relevant to our argument. They are:

(a) $R' < R$ if and only if $8|R'_j| \leq |R_j|$ for $j = 1$ and $j = 2$.

(b) For $j = 1$ or $j = 2$, define $R' <_j R$ if and only if $R' \lesssim R$ and $8|R'_j| \leq |R_j|$ but $R' \nless R$.

(c) $R' \simeq R$ if and only if $\frac{1}{4}|R_j| \leq |R'_j| \leq |R_j|$ for $j = 1$ and $j = 2$.

These four partial orders divide the collection $\{(R', R): R' \in \mathcal{W}, R \in \mathcal{U}, R' \lesssim R\}$ into four subclasses which require different arguments.

In each of these four arguments, we have recourse to this definition. Set $\mathcal{U}_k$, for $k = 0, 1, 2, \ldots$, to be those rectangles in $\mathcal{U}$ with $2^{-k-1} < \mu(R) \leq 2^k, R \in \mathcal{U}_k$. Journe's lemma enters into the considerations. Let $\mathcal{U}' \subset \mathcal{U}_k$ be a collection of rectangles which are pairwise incomparable with respect to inclusion. For this latter collection, we have the inequality

$$\sum_{R \in \mathcal{U}'} |R| \leq K_\delta 2^{\delta/100} \left|\bigcup_{R \in \mathcal{U}'} R\right|. \tag{8}$$

See Journe's [7]. This together with the assumption that b has small rectangular BMO norm gives us

$$\left\|b^{\mathcal{U}_k}\right\|_{BMO} \leq K_\delta 2^{k/100} \varepsilon. \tag{9}$$

This interplay between the small rectangular BMO norm and Journe''s lemma is a decisive feature of the argument.

Essentially, the decomposition into the collections $\mathcal{U}_k$ is a spatial decomposition of the collection $\mathcal{U}$. A corresponding decomposition of $\mathcal{W}$ enters in. Yet the definition of this class differs slightly depending on the partial order we are considering.

For $R' \in \mathcal{W}$ and $R \in \mathcal{U}$ the term $\{v^{R'}, v^R\}$is a linear combination of

$$v_{R'} H_2 H_1 \overline{v_R}, H_2(v_{R'} H_1 \overline{v_R}), (H_1 v_{R'})(H_2 \overline{v_R}), H_1 H_2(v_{R'} \overline{v_R}).$$

Consider the last term. As we are to estimate an L^2-norm, the leading operators $H_1 H_2$ can be ignored. Moreover, the essential properties of wavelets used below still hold for the conjugates and Hilbert transforms of the same. These properties are Fourier localization and spatial localization. Similar comments apply to the other three terms, and so the arguments below applies to each type of term above.

We consider the case of $R' < R$ for $R' \in \mathcal{W}$ and $\in \mathcal{U}$. The sums we consider are related to the following definition. Set

$$b_{trun}^{\mathcal{U}_k}(x) := \sup_{R'} \left| \sum_{\substack{R \in \mathcal{U}_k \\ R' < R}} \langle b, v_R \rangle v_R(x) \right|.$$

We consider the maximal truncation of the sum over all choices of dimensions of the rectangles in the sum. Thus, this sum is closely related to the strong maximal function M applied to $b^{\mathcal{U}_k}$, so that in particular we have the estimate below, which relies upon(9):

$$\left\| b_{trun}^{\mathcal{U}_k} \right\|_p \lesssim \varepsilon 2^{\frac{k}{100}}, 1 < p < \infty.$$

(By a suitable definition of the strong maximal function$M$, one can deduce this inequality from the L^p-bounds for M.) We apply this inequality far away from the set U. For the set$W_\lambda = R^2 - \bigcup_{R \in \mathcal{U}_k} \lambda R, \lambda > 1$, we have the inequality

$$\left\| b_{trun}^{\mathcal{U}_k} \right\|_{L^p(W)} \lesssim \varepsilon 2^{\frac{k}{100}} \lambda^{-100}, 1 < p < \infty. \qquad (10)$$

We shall need a refined decomposition of the collection $\mathcal{W}$, the motivation for which is the following calculation. Let $\mathcal{W}' \subset \mathcal{W}$. For $n = (n_1, n_2) \in \boldsymbol{Z}^2$, set

$$\mathcal{W}'(n) := \{R' \in \mathcal{W}' : |R_j'| = 2^{n_j}, j = 1,2\}.$$

In addition, let

$$B(\mathcal{W}', n) := \sum_{R' \in \mathcal{W}'(n)} \sum_{\substack{R \in \mathcal{U}_k \\ R' < R}} \langle b, v_{R'} \rangle \overline{\langle b, v_R \rangle} v_{R'} v_R.$$

And set $B(\mathcal{W}') = \sum_{n \in \boldsymbol{Z}^2} B(\mathcal{W}', n)$.

Then, in view of (7), we see that $B(\mathcal{W}', n)$ and $B(\mathcal{W}', n')$are orthogonal if n and n'differ by at least three in either coordinate. Thus,

$$\left\| \sum_{n \in \boldsymbol{Z}^2} B(\mathcal{W}', n) \right\|_2^2 \leq 3 \sum_{n \in \boldsymbol{Z}^2} \|B(\mathcal{W}', n)\|_2^2.$$

The rectangles $R' \in \mathcal{W}(n)$are all translates of one another. Thus, taking advantage of the rapid spatial decay of the wavelets, we can estimate

$$\|B(\mathcal{W}', n)\|_2^2 \lesssim \sum_{R' \in \mathcal{W}(n)} \int \left| \frac{|\langle b, v_{R'} \rangle|}{\sqrt{|R'|}} (\chi_{R'} * 1_{R'}) b_{trun}^{\mathcal{U}_k} \right|^2 dx.$$

In this display, we let $\chi(x_1, x_2) = (1 + x_1^2 + x_2^2)^{-10}$and for rectangles $R, \chi_R(x_1, x_2) = \chi(x_1 |R_1|^{-1}, x_2 |R_2|^{-1})$. Note that χ_R depends only oil the dimensions of R and not its location.

Continuing, note the trivial inequality$\int (\chi_R * f)^2 g dx \lesssim \int |f|^2 \chi_R * g dx$. We. can estimate

$$\|B(\mathcal{W}')\|_2^2 \lesssim \sum_{R'\in\mathcal{W}'} |\langle b, v_{R'}\rangle|^2 \left\{|R'|^{-1}\int_{R'} M\left(\left|b_{trun}^{\mathcal{U}_k}\right|^2\right)dx\right\} \qquad (11)$$

$$\lesssim \left|\bigcup_{R'\in\mathcal{W}'} R'\right| \sup_{R'\in\mathcal{W}'} \operatorname{avg}(R').$$

Here we take $\operatorname{avg}(R') \coloneqq |R'|^{-1}\int_{R'} M\left(\left|b_{trun}^{\mathcal{U}_k}\right|^2\right)$.

The terms $\operatorname{avg}(R')$ are essentially of the order of magnitude ε^2 times the scaled distance between R' and the open set U. To make this precise requires a decomposition of the collection $\mathcal{W}$.

For integers $l > k$ and $m \geq 0$, set $\mathcal{W}(l,m)$to be those $R' \in \mathcal{W}$ which satisfy these three conditions:

(a) First, $\operatorname{avg}(R') \leq \varepsilon^2 2^{-4l}$if $m = 0$ and $\varepsilon^2 2^{-4l+m-1} < \operatorname{avg}(R')\varepsilon^2 2^{-4l+m}$ if $m > 0$.

(b) Second, there is an $R \in \mathcal{U}_k$ with $R' < R$ and $R' \subset 2^{l+1}R$.

(c) Third, for every $R \in \mathcal{U}_k$with$R' < R$, we have $R' \not\subset 2^{l+1}R$. Certainly, this collection of rectangles is empty if $l \leq k$.

We see that

$$\left|\bigcup_{R'\in\mathcal{W}(l,m)} R'\right| \lesssim \min\left(2^{2/p}, 2^{-mp/2}\right), 1 < p < \infty.$$

The first estimate follows since the rectangles $R' \in \mathcal{W}(l,m)$ are contained in the set $\{M1_U \geq 2^{-2l-1}\}$. The second estimate follows from (10).

But then from (11) we see that for $m > 0$,

$$\left\|B\left(\mathcal{W}(l,m)\right)\right\|_2^2 \lesssim \varepsilon^2 2^{-4l+m}\min\left(2^{2/p}, 2^{-mp/2}\right) \lesssim \varepsilon^2 2^{-(m+l)/10}.$$

In the case that $m = 0$, we have the bound 2^{2lp}. This is obtained by taking the minimum to be 2^{2lp} for $p = \frac{5}{4}$ and $0 < m < \frac{11}{8}l$. For $m \geq \frac{11}{8}l$take the minimum to be $2^{-mp/2}$ with $p = 4$.

This last estimate is summable over $0 < k < l$ and $0 < m$ to at most $\lesssim \varepsilon$, and so completes this case.

We treat the case of $R' <_1 R$, while the case of $R' <_2 R$ is the same by symmetry. The structure of this partial order provides some orthogonality in the first variable, leaving none in the second variable. Bounds for the expressions from the second variable are derived from a cognate of a Carleson measure estimate.

There is a basic calculation that we perform for a subset $\mathcal{W}' \subset \mathcal{W}$. For an integer $n' \in \boldsymbol{Z}$define$\mathcal{W}'(n') \coloneqq \left\{R' \in \mathcal{W}' \colon |R_1'| = 2^{n'}\right\}$and

$$B(\mathcal{W}', n') \coloneqq \sum_{R'\in\mathcal{W}'(n')} \sum_{\substack{R\in\mathcal{U}_k \\ R'<_1 R}} \langle b, v_{R'}\rangle, \overline{\langle b, v_R\rangle} v_{R'}\overline{v_R}.$$

Recalling(7), if $n'$and n'' differ by more than 3, then$B(\mathcal{W}', n')$ and $B\left(\mathcal{W}', n''\right)$ are orthogonal.

Observe that for R' and R as in the sum defining $B(\mathcal{W}', n)$, we have the estimate

$$|v_{R'}(x)\overline{v_R}(x)| \lesssim (|R||R'|)^{-\frac{1}{2}} \operatorname{dist}(R', R)^{1000}\, \chi_{R'} * 1_{R'}(x), x\boldsymbol{R}^2. \qquad (12)$$

In this display, we are using the same notation as before,$\chi(x_1, x_2) = (1 + x_1^2 + x_2^2)^{-10}$and for rectangles $R, \chi_R(x_1, x_2) = \chi(x_1|R_1|^{-1}, x_2|R_2|^{-1})$. In addition, $\operatorname{dist}(R', R) \coloneqq$

$M1_R(c(R'))$, with $c(R')$ being the center of R'. (This distance is more properly the inverse of a distance that takes into account the scale of the rectangle R.) Now define

$$\beta(R') \coloneqq \sum_{\substack{R\in\mathcal{U}\\ R'<_1 R}} |R|^{-12} |\langle b, v_R\rangle| \operatorname{dist}(R', R)^{1000}. \tag{13}$$

The main point of these observations and definitions is this. For the function $B(\mathcal{W}') \coloneqq \sum_{n'\in\boldsymbol{Z}} B(\mathcal{W}', n')$, we have

$$\|B(\mathcal{W}')\|_2^2 \lesssim \|B(\mathcal{W}', n')\|_2^2$$

$$\lesssim \sum_{n'\in\boldsymbol{Z}} \int_{\boldsymbol{R}^2} \left[\sum_{R'\in\mathcal{W}'(n')} |\langle b, v_{R'}\rangle| \beta(R') |R'|^{-1/2} \chi_{R'} * 1_{R'} \right]^2 dx$$

$$\lesssim \sum_{n'\in\boldsymbol{Z}} \int_{\boldsymbol{R}^2} \left[\sum_{R'\in\mathcal{W}'(n')} |\langle b, v_{R'}\rangle| \beta(R') |R'|^{-1/2} 1_{R'} \right]^2 dx.$$

At this point, it occurs to one to appeal to the Carleson measure property associated to the coefficients $|\langle b, v_{R'}\rangle| |R'|^{-1/2}$. This necessitates that one proves that the coefficients $\beta(R')$ satisfy a similar condition, which doesn't seem to be true in general. A slightly weaker condition is however true.

To get around this difficulty, we make a further diagonalization of the terms $\beta(R')$ above. For integers $\nu \geq \nu_0, \mu \geq 1$and a rectangle $R' \in \mathcal{W}$, consider rectangles $R \in \mathcal{U}_k$such that

$$R' <_1 R, 2^{-\upsilon} \leq \operatorname{dist}(R'.R) \leq 2^{-\upsilon+1}, 2^{\mu}|R'| = |R|.$$

(The quantity υ_0 depends upon the particular subcollection $\mathcal{W}'$ we are considering.) We denote one of these rectangles as$\pi(R')$.

An important geometrical fact is this. We have$\pi(R') \subset 2^{\upsilon+\mu+10} R_1' \times 2^{\upsilon+10} R_2'$ and in particular, this last rectangle has measure$\lesssim 2^{2\upsilon+\mu}|R'|$.

Therefore, there are at most $O(2^{2\upsilon})$ possible choices for $\pi(R')$. (Small integral powers of 2^{υ} are completely harmless because of the large power of$\operatorname{dist}(R'.R)$ that appears in (13).)

We bound this next expression by a term which includes a power of ε, a small power of 2^{υ} and a power of $2^{-\mu}$. Define

$$S(\mathcal{W}', \nu, \mu) \coloneqq \sum_{n'\in\boldsymbol{Z}} \int_{\boldsymbol{R}^2} \left[\sum_{R'\in\mathcal{W}'(n')} \frac{|\langle b, v_{R'}\rangle\langle b, v_{\pi(R')}\rangle|}{\sqrt{|R'||\pi(R')|}} \chi_{R'} * 1_{R'} \right]^2 dx$$

$$\lesssim \sum_{n'\in\boldsymbol{Z}} \int_{\boldsymbol{R}^2} \left[\sum_{R'\in\mathcal{W}'(n')} \frac{|\langle b, v_{R'}\rangle\langle b, v_{\pi(R')}\rangle|}{\sqrt{|R'||\pi(R')|}} 1_{R'} \right]^2 dx$$

$$= \sum_{n'\in\boldsymbol{Z}} \sum_{R'\in\mathcal{W}'(n')} \frac{|\langle b, v_{R'}\rangle\langle b, v_{\pi(R')}\rangle|}{\sqrt{|R'||\pi(R')|}} \sum_{\substack{R''\in\mathcal{W}'(n')\\ R''\subset R'}} \sqrt{\frac{|R''|}{|\pi(R'')|}} |\langle b, v_{R''}\rangle\langle b, v_{\pi(R'')}\rangle|.$$

The innermost sum can be bounded this way. First$\|b\|_{BMO(rec)} \leq \varepsilon$, so that

$$\sum_{R''\subset R'} |\langle b, v_{R''}\rangle|^2 \leq \varepsilon^2 |R'|.$$

Second, by our geometrical observation about $\pi(R')$,

$$\sum_{R''\subset R'} \frac{|R''|}{|\pi(R'')|} \left|\langle b, \upsilon_{\pi(R'')}\rangle\right|^2 \lesssim \varepsilon^2 2^{2\upsilon}|R'|.$$

In particular, the factor 2^udoes not enter into this estimate.
This means that

$$S(\mathcal{W},\nu,\mu) \lesssim \varepsilon^2 2^{2\upsilon} \sum_{R'\in\mathcal{W}'} \sqrt{\frac{|R'|}{|\pi(R')|}} \left|\langle b, \upsilon_{R'}\rangle\langle b, \upsilon_{\pi(R')}\rangle\right|$$

$$\lesssim \varepsilon^2 2^{2\upsilon-\mu/2} \left[\sum_{R'\in\mathcal{W}'} |\langle b, \upsilon_{R'}\rangle|^2 \sum_{R\in\mathcal{U}_k} |\langle b, \upsilon_R\rangle|^2\right]^{1/2} \lesssim \varepsilon^2 2^{2\upsilon-\mu/2} \left|\bigcup_{R'\in\mathcal{W}'} R'\right|^{1/2}.$$

The point of these computations is that a further trivial application of the Cauchy-Schwarz inequality proves that

$$\|B(\mathcal{W}')\|_2 \lesssim \varepsilon 2^{-100\nu_0} \left|\bigcup_{R'\in\mathcal{W}'} R'\right|^{1/4},$$

where ν_0 is the largest integer such that for all $R' \in \mathcal{W}'$ and $R \in \mathcal{U}_k$, we have $\operatorname{dist}(R',R) \le 2^{-\nu_0}$.
We shall complete by decomposing $\mathcal{W}$ into subcollections for which this last estimate is summable to $\varepsilon 2^{-k}$. Indeed, take $\mathcal{W}_\upsilon$ to be those $R' \in \mathcal{W}$ with $R' \not\subset 2^\upsilon R$ for all $R \in \mathcal{U}_k$ with $R' <_1 R$. And there is an $R \in \mathcal{U}_k$ with $R' \subset 2^{\upsilon+1}R$ and $R' <_1 R$. Certainly, we need only consider $\upsilon \ge k$.

It is clear that this decomposition of $\mathcal{W}$ will conclude the treatment of this partial order.

We now consider the case of $R' \simeq R$, which is less subtle as there is no orthogonality to exploit and the Carleson measure estimates are more directly applicable. We prove the bound

$$\left\|\sum_{R'\in\mathcal{W}} \sum_{\substack{R\in\mathcal{U}\\ R'\simeq R}} \langle b, \upsilon_{R'}\rangle\overline{\langle b, \upsilon_R\rangle}\upsilon_{R'}\overline{\upsilon_R}\right\|_2 \lesssim K_\delta \varepsilon^{1/3}.$$

The diagonalization in space takes two different forms. For $\lambda \ge 2^k$and $R \in \mathcal{U}_k$set $\sigma(\lambda, R)$to be a choice of $R' \in \mathcal{W}$ with $R' \simeq R$ and $R' \subset 2\lambda R$. (The definition is vacuous for $\lambda < 2^k$.) It is clear that we need only consider $\simeq \lambda^2$choices of these functions $\sigma(\lambda, .)\colon \mathcal{U}_k \to \mathcal{W}$. There is an L^1-estimate which allows one to take advantage of the spatial separation between R and $\sigma(\lambda, R)$:

$$\left\|\sum_{R\in\mathcal{U}_k} \langle b, \upsilon_{\sigma(\lambda,R)}\rangle\overline{\langle b, \upsilon_R\rangle}\upsilon_{\sigma(\lambda,R)}\overline{\upsilon_R}\right\|_1 \lesssim \lambda^{-100} \sum_{R\in\mathcal{U}_k} \left|\langle b, \upsilon_{\sigma(\lambda,R)}\rangle\overline{\langle b, \upsilon_R\rangle}\right|$$

$$\lesssim \lambda^{-100} \left[\sum_{R\in\mathcal{U}_k} \left|\langle b, \upsilon_{\sigma(\lambda,R)}\rangle\right|^2 \sum_{R\in\mathcal{U}_k} |\langle b, \upsilon_R\rangle|^2\right]^{1/2}$$

$$\lesssim K_\delta \varepsilon \lambda^{-90}.$$

This estimate uses (9) and is a very small estimate.

To complete this case we need to provide an estimate in L^4. Here, we can be quite inefficient. By Cauchy-Schwarz and the Littlewood-Paley inequalities,

$$\left\|\sum_{R\in\mathcal{U}}\langle b,v_{\sigma(\lambda,R)}\rangle\overline{\langle b,v_R\rangle}v_{\sigma(\lambda,R)}\overline{v_R}\right\|_4$$

$$\lesssim\left\|\left[\sum_{R\in\mathcal{U}}\left|\langle b,v_{\sigma(\lambda,R)}\rangle v_{\sigma(\lambda,R)}\right|^2\right]^{\frac{1}{2}}\right\|_4\left\|\left[\sum_{R\in\mathcal{U}}|\langle b,v_R\rangle\overline{v_R}|^2\right]^{\frac{1}{2}}\right\|_4\lesssim\lambda.$$

This follows directly from the BMO assumption on$b$. Our proof is complete.

Let U be an open set of finite measure in the plane. Let $\mathcal{R}(U)$be all dyadic rectangles in$\mathcal{R}$that are contained in U. For each $R\in\mathcal{R}(U)$and open set $V\supset U$, set

$$\mu(V;R)=\sup\{\mu>0:\mu R\subset V\}.$$

The form of Journs lemma we need is

Lemma (1.1.4)[1]: For each $0<\delta<1$and open set U of finite measure in the plane, there is a set $V\supset U$ for which$|V|<(1+\delta)|U|$, and for all$0<\varepsilon<1$, there is a constant$K_{\delta,\varepsilon}$so that for any subset $\mathcal{R}'\subset\mathcal{R}(U)$ such that $R\not\subset R'$for any two rectangles$R\neq R'\in\mathcal{R}'$, we have the inequality

$$\sum_{R\in\mathcal{R}'}\mu(V;R)^{-\varepsilon}|R|\le K_{\delta,\varepsilon}\left|\bigcup_{R\in\mathcal{R}'}R\right|. \tag{14}$$

Journ~'s lemma is the central tool in verifying the Carleson measure condition, and points to the central problem in two dimensions: that there can be many rectangles close to the boundary of an open set.

Among the references we could find in the literature[7], [10], the form of Journ~'s lemma cited and proved take the set V to be$\left\{M1_U>\frac{1}{2}\right\}$, whizch only satisfies$|V|<K|U|$.

Proof. There are two stages of the proof, with the first stage being the specification of the set V. This must be done with some care, and in a manner that depends upon$\delta>0$. Let us illustrate the difficulty.

At first guess, one would take$V:=\{M1_U>1-\delta\}$, with M being the strong maximal function. But the problem is that the strong maximal function is not bounded on $L^1(\boldsymbol{R}^2)$,so it can't possibly satisfy the desired inequality on its measure.

It is then tempting to define V as some variant of the one-dimensional maximal function. While this maximal function is bounded on $L^1(R)$, as a map into $L^{1,\infty}(\boldsymbol{R})$, the norm is known to exceed one.

The dyadic maximal function, however, maps L^1into $L^{1,\infty}$with norm one. This well known fact we shall utilize in a slightly more general form. Define a grid to be a collection $\mathcal{J}$ of intervals in the real line for which for all $I,I'\in\mathcal{J},I\cap I'\in\{\emptyset,I,I'\}$. For a collection of intervals $\mathcal{J}$, not necessarily a grid, set

$$M^{\mathcal{J}}f(x):=\sup_{I\in\mathcal{J}}1_I(x)|I|^{-1}\int_I f(y)\,dy.$$

Then, for any grid $\mathcal{J}$,$M^{\mathcal{J}}$ maps $L^1(\boldsymbol{R})$into$L^{1,\infty}(\boldsymbol{R})$ with norm one. This is in particular true for the dyadic grid $\mathcal{D}$.

Now, let us take $0<\delta<1$, and in particular take $\delta=(2^d+1)^{-1}$ for integer d. We define shifted dyadic grids, modifying an observation due to M. Christ. For integers $0\le b<d$, and $\alpha\in\{\pm(2^d+1)^{-1}\}$, let

$$\mathcal{D}_{d,b,\alpha} := \left\{2^{kd+b}\left((0,1)+j+(-1)^k\alpha\right) : k \in \boldsymbol{Z}, j \in \boldsymbol{Z}\right\}.$$

One checks that this is a grid. Indeed, it suffices to assume$\alpha = (2^d+1)^{-1}$, and that$b = 0$. Checking the grid structure can be done by induction. And it suffices to check that the intervals in $\mathcal{D}_{d,0,\alpha}$of length one are a union of intervals in $\mathcal{D}_{d,b,\alpha}$ of length 2^{-d}. One need only check this for the interval $(0,1)+\alpha$. But certainly

$$(0,1)+\frac{1}{2^d+1} = \bigcup_{j=1}^{2^d}(0,2^{-d})+j2^{-d}-\frac{1}{2^d(2^d+1)}$$

$$= \bigcup_{j=1}^{2^d}\left(\frac{1}{2^d(2^d+1)},\frac{1}{2^d+1}\right)+\frac{j}{2^d}.$$

What is more important concerns the collections $\mathcal{D}_d := \bigcup_\alpha \bigcup_{b=0}^{d-1}\mathcal{D}_{d,b,\alpha}$.For each dyadic interval $I \in \mathcal{D}, I \pm \delta|I| \in \mathcal{D}_d$. (The problem we are avoiding here is that the dyadic grid distinguishes dyadic rational points. At the point 0, for instance, observe that for all integers$k, (1+\delta)(0,1) \not\subset (0,2^k)$, regardless of how big k is.) Moreover, the maximal function $M^{\mathcal{D}_d}$ maps L^1 into $L^{1,\infty}$ with norm at most $2d \simeq \log\delta$.

We may define V. For a collection of intervals $\mathcal{J}$ and $j = l,2$, set $M_j^{\mathcal{J}}$to be the maximal function associated to $\mathcal{J}$, computed in the coordinate j. Initially, we use only the dyadic grids, setting

$$V_0 = \bigcup_{i\neq j}\left\{M_i^{\mathcal{D}}1\left\{M_j 1_U > 1-\delta\right\} > 1-\delta\right\}.$$

It is clear that$|V_0| < (1+K\delta)|U|$. Invoking the collections $\mathcal{D}_d$, set

$$V = \bigcup_{i\neq j}\left\{M_i^{\mathcal{D}_d}1\left\{M_j^{\mathcal{D}_d}1_{V_0} > 1-\delta\right\} > 1-\delta\right\}.$$

Then$|V| < (1+K\delta\log\delta^{-1})|U|$, and we will work with this choice of V.

The additional important property that V has can be formulated this way. For all dyadic rectangles$R = R_1 \times R_2 \subset V_0$, the four rectangles

$$(R_1 \pm \delta|R_1|)\times(R_2 \pm \delta|R_2|) \subset V. \qquad (15)$$

This follows immediately from the construction of the shifted dyadic grids. The first stage of the proof is complete.

In the second stage, we verify(14). A typical proof of Journ~'s lemma shows that the rectangles in $\mathcal{R}'$ have logarithmic overlap, measured in terms of $\log\mu(V;U)$. We adopt that method of proof. Fix a subset$\mathcal{R}' \subset \mathcal{R}(U)$satisfying the incomparability condition of the lemma, and fix $\mu \geq 1$. Set S to be those rectangles in $\mathcal{R}'$ with $\mu \leq \mu(R) \leq 2\mu$. It suffices to show that

$$\sum_{R\in S}|R| \lesssim (1+\log\mu)^2\left|\bigcup_{R\in S}R\right|.$$

For then this estimate is summed over $\mu \in \{2^k : k \in \boldsymbol{Z}\}$.

In showing this estimate, we can further assume for all$R, R' \in S$, writing$R = R_1 \times R_2$and likewise for R', that if for $j = l,2, |R_j| > |R_j'|$then$|R_j| > 16\mu 6^{-1}|R_j'|$. This is done by restricting $\log_2|R_j|$to be in an arithmetic progression of difference $\simeq \log\mu\delta^{-1}$. This necessitates the division of all rectangles into$\lesssim (1+\log\mu\delta^{-1})^2$ subclasses, and so we prove the bound above without the logarithmic term.

We define a bad class of rectangles $\mathcal{B}=\mathcal{B}(S)$ as follows. For $j=l,2$, let $\mathcal{B}_j(S)$ be those rectangles R for which there are rectangles

$$R^1, R^2, \dots, R^k \in S-\{R\},$$

so that for each $1 \le k \le K, |R_j^k| > |R_j|$ and

$$\left| R \cap \bigcup_{k=1}^{K} R^k \right| > \left(1-\frac{1}{10}\delta\right)|R|.$$

Thus $R \in \mathcal{B}_j$ if it is nearly completely covered by dyadic rectangles in the j th direction of the plane. Set $\mathcal{B}(S)=\mathcal{B}_1(S)\cup\mathcal{B}_2(S)$. It follows that if $R \notin \mathcal{B}(S)$, it is not covered in both the vertical and horizontal directions, hence

$$\left| R \cap \bigcap_{R' \in S-\{R\}} (R')^c \right| \ge \frac{\delta^2}{100|R|}.$$

And so

$$\sum_{R \in S-\mathcal{B}(S)} |R| \le 100\delta^2 \left| \bigcup_{R\in S} R \right|.$$

Thus, it remains to consider the set of rectangles $\mathcal{B}_1(S)$ and $\mathcal{B}_2(S)$. Observe that for any collection $S'\mathcal{B}_j(S') \subset S'$, as follows immediately from the definition. Hence $\mathcal{B}_1\left(\mathcal{B}_2\left(\mathcal{B}_1(S)\right)\right) \subset \mathcal{B}_1\left(\mathcal{B}_1(S)\right)$. And we argue that this last set is empty. As our definition of V and $\mu(V;R)$ is symmetric with respect to the coordinate axes, this is enough to finish the proof.

We argue that $\mathcal{B}_1\left(\mathcal{B}_1(S)\right)$ is empty by contradiction. Assume that R is in this collection. Consider those rectangles R' in $\mathcal{B}_1(S)$ for which (i) $|R_1'| > |R_1|$ and (ii) $R' \cap R \neq \emptyset$ Then

$$\left| R \cap \bigcup_{R' \in \mathcal{B}_1(S)} R' \right| \ge \left(1-\frac{1}{10}\delta\right)|R|.$$

Fix one of these rectangles R' with $|R_1'|$ being minimal. We then claim that $8\mu R' \subset V$, which contradicts the assumption that $\mu(V;R')$ is no more than 2μ.

Indeed, all the rectangles in $\mathcal{B}_1(S)$ are themselves covered by dyadic rectangles in the first coordinate axis. We see that the the set $\left\{M_2^{\mathcal{D}} 1_U > 1-\delta\right\}$ contains the dyadic rectangle $R_1'' \times R_2$, in which R_2 is the second coordinate interval for the rectangle R and R_1'' is the dyadic interval that contains R_1' and has measure $8\mu\delta^{-1}|R_1'| \le |R_1''| < 16\mu\delta^{-1}|R_1'|$. That is, $R_1'' \times R_2$ is contained in V_0. And the dimensions of this rectangle are very much bigger than those of R. Applying (15), the rectangles $\left(R_1'' \pm |R_1''|\right) \times R_2 \pm \delta|R_2|$ are contained in V. And since $8\mu R'$ is contained in one of these last four rectangles, we have contradicted the assumption that $\mu(V;R') < 2\mu$.

Section (1.2): Multi-Parameter Paraproducts

For $n \ge 1$ let $m\left(=m(\tau)\right)$ in $L^\infty(\mathbb{R}^n)$ be a bounded function, smooth away from the origin and satisfying

$$|\partial^\alpha m(\tau)| \lesssim \frac{1}{|\tau|^{|\alpha|}} \tag{16}$$

for sufficiently many multi-indices α. Denote by $T_m^{(1)}$ the n-linear operator defined by

$$T_m^{(1)}(f_1, \ldots, f_n)(x) = \int_{\mathbb{R}^n} m(\tau)\widehat{f_1}(\tau_1) \ldots \widehat{f_n}(\tau_n) e^{2\pi i x(\tau_1+\cdots+\tau_n)} d\tau \quad (17)$$

where $f_1, \ldots, f_n$ are Schwartz functions on the real line$\mathbb{R}$. The following statement of Coifman and Meyer is a classical theorem in Analysis [13], [17], [15].

Theorem (1.2.1)[12]: $T_m^{(1)}$ maps $L^{p_1} \times \ldots \times L^{p_n} \to L^p$ boundedly, as long as $1 < p_1, \ldots, p_n \leq \infty, \frac{1}{p_1} + \cdots + \frac{1}{p_n} = \frac{1}{p}$ and $0 < p < \infty$.

In [18] we considered the bi-parameter analogue of $T_m^{(1)}$ defined as follows. Let $m\big(= (\gamma, \eta)\big)$ in $L^\infty(\mathbb{R}^{2n})$ be a bounded function, smooth away from the subspaces $\{\gamma = 0\} \cup \{\eta = 0\}$and satisfying

$$\left|\partial_\gamma^\alpha \partial_\eta^\beta m(\gamma, \eta)\right| \lesssim \frac{1}{|\gamma|^{|\alpha|}} \frac{1}{|\eta|^{|\beta|}} \quad (18)$$

for sufficiently many multi-indices α and β.

Denote by $T_m^{(2)}$ the n-linear operator defined by

$$T_m^{(2)}(f_1, \ldots, f_n)(x) = \int_{\mathbb{R}^{2n}} m(\gamma, \eta)\widehat{f_2}(\gamma_1, \eta_1) \ldots \widehat{f_n}(\gamma_n, \eta_n) \times e^{2\pi i x[(\gamma_1,\eta_1)+\cdots+(\gamma_n,\eta_n)]} d\gamma d\eta, \quad (19)$$

where $f_1, \ldots, f_n$ are Schwartz functions on the plane$\mathbb{R}^n$. The following theorem has been proven in [18].

Theorem (1.2.2)[12]: $T_m^{(2)}$ maps $L^{p_1} \times \ldots \times L^{p_n} \to L^p$ boundedly, as long as $1 < p_1, \ldots, p_n \leq \infty, \frac{1}{p_1} + \cdots + \frac{1}{p_n} = \frac{1}{p}$ and $0 < p < \infty$.

We generalize Theorem (1.2.2) to the d-parameter setting, for any $d \geq 1$.

In general, if $\xi_1 = \left(\xi_1^i\right)_{i=1}^d, \ldots, \xi_n = \left(\xi_n^i\right)_{i=1}^d$are n generic vectors in $\mathbb{R}^d$, they naturally generate the following d vectors in$\mathbb{R}^n$ which we will denote by $\overline{\xi_1} = \left(\xi_j^1\right)_{j=1}^n, \ldots, \overline{\xi_d} = \left(\xi_j^d\right)_{j=1}^n$. As before, let $m\left(= m(\xi) = m(\bar{\xi})\right)$in $L^\infty(\mathbb{R}^{dn})$be a bounded symbol, smooth away from the subspaces $\left\{\overline{\xi_1} = 0\right\} \cup \ldots \cup \left\{\overline{\xi_d} = 0\right\}$and satisfying

$$\left|\partial_{\overline{\xi_1}}^{\alpha_1} \ldots \partial_{\overline{\xi_d}}^{\alpha_d} m(\bar{\xi})\right| \lesssim \prod_{i=1}^d \frac{1}{\left|\overline{\xi_\iota}\right|^{|\alpha_i|}} \quad (20)$$

for sufficiently many multi-indices $\alpha_1, \ldots, \alpha_d$. Denote by $T_m^{(d)}$ the n-linear operator defined by

$$T_m^{(d)}(f_1, \ldots, f_n)(x) = \int_{\mathbb{R}^{dn}} m(\xi)\widehat{f_1}(\xi_1) \ldots \widehat{f_n}(\xi_n)\, e^{2\pi i x(\xi_1+\cdots+\xi_n)} d\xi \quad (21)$$

where$f_1, \ldots, f_n$ are Schwartz functions on$\mathbb{R}^d$. The main theorem is the following.

Theorem (1.2.3)[12]: $T_m^{(d)}$ maps$L^{p_1} \times \ldots \times L^{p_n} \to L^p$ boundedly, as long as$1 < p_1, \ldots, p_n \leq \infty, \frac{1}{p_1} + \cdots + \frac{1}{p_n} = \frac{1}{p}and0 < p < \infty$.

Classically,$[2, 7, 5]$ an estimate as the one in Theorem (1.2.1) is proved by using the $T(1)$ theorem of David and Journ´e[19] together with the Caldero´n- Zygmund decomposition. In particular, the theory of BMO functions and Carleson measures is involved.

On the other hand, it is well known [2], [16] that in the multi-parameter setting all these results and concepts are much more delicate (BMO, John-Nirenberg inequality, Caldero´n-

Zygmund decomposition). To overcome these difficulties, in [18] we had to develop a completely new approach to prove Theorem(1.2.2). This approach relied on the one dimensional BMO theory and also on Journe's lemma [16],[1], but did not extend to prove the general d-parameter case.
The novelty is that it simplifies the method introduced in [18] and this simplification works equally well in all dimensions. It turned out that one doesn't need to rely on any knowledge of BMO, Carleson measures or Journ´e's lemma in order to prove the estimates in Theorem(1.2.3).
We shall rely on [18] and we chose to present the argument in the same bi-linear bi-parameter setting (so both n and d will be equal to 2). However, it will be clear from the proof that its extension to the n-linear d-parameter case is straightforward.

We recall the discretization procedure from [18] which reduces the study of our operator to the study of some general multi-parameter paraproducts. We present the proof of our main theorem, Theorem (1.2.3) and we give a proof of Lemma (1.2.6) which plays an important role in our simplified construction.

As we promised, assume throughout that $n = d = 2$. In this case, our operator$T_m^{(d)}$ can be written as

$$T_m^{(2)}(f,g)(x) = \int_{\mathbb{R}^4} m(\gamma,\eta)\hat{f}(\gamma_1,\eta_1)\hat{g}(\gamma_2,\eta_2)e^{2\pi i x[(\gamma_1,\eta_1)+(\gamma_2,\eta_2)]}d\gamma d\eta\,. \quad (22)$$

In [18], we decomposed the operator $T_m^{(2)}$ into smaller pieces, well adapted to its bi-parameter structure. This allowed us to reduce its analysis to the analysis of some simpler discretized dyadic paraproducts. We will recall their definitions below.
An interval I on the real line $\mathbb{R}$ is called dyadic if it is of the form $I = 2^k[n, n+1]$ for some $k, n \in \mathbb{Z}$. If $\lambda, t \in [0,1]$ are two parameters and I is as above, we denote by $I_{\lambda,t}$ the interval $I_{\lambda,t} = 2^{k+\lambda}[n+t, n+t+1]$.
Definition (1.2.4)[12]: For $J \subseteq \mathbb{R}$ an arbitrary interval, we say that a smooth function Φ_J is a bump adapted to J, if and only if the following inequalities hold

$$\left|\Phi_J^{(l)}(x)\right| \le C_{l,\alpha}\frac{1}{|J|^l}\frac{1}{\left(1+\frac{\mathrm{dist}(x,J)}{|J|}\right)^{\alpha}}, \quad (23)$$

for every integer $\alpha \in \mathbb{N}$ and for sufficiently many derivatives $l \in \mathbb{N}$. If$\Phi_J$is a bump adapted to J, we say that $|J|^{-1/2}\Phi_J$ is an L^2-normalized bumpadapted to J.
For $\lambda, \mathrm{t}_1, \mathrm{t}_2, \mathrm{t}_3 \in [0,1]$ and $j \in \{1,2,3\}$ we define the discretized dyadic paraproduct $\prod^j_{\lambda,\mathrm{t}_1,\mathrm{t}_2,\mathrm{t}_3}$ of "type$j$" by

$$\prod\nolimits^j_{\lambda,\mathrm{t}_1,\mathrm{t}_2,\mathrm{t}_3}(f,g) = \sum_{I\in\mathcal{D}}\frac{1}{|I|^{1/2}}\langle f,\Phi^1_{I_{\lambda,t_1}}\rangle\langle g,\Phi^2_{I_{\lambda,t_2}}\rangle\Phi^3_{I_{\lambda,t_3}}, \quad (24)$$

where f, g are complex-valued measurable functions on $\mathbb{R}$ and $\Phi^1_{I_{\lambda,t_i}}$,ti are L^2-normalized bumps adapted to $I_{\lambda,t_i}$with the additional property that $\int_{\mathbb{R}}\Phi^i_{I_{\lambda,t_i}}(x)dx = 0$ for $i \neq j, i = 1,2,3$. $\mathcal{D}$ is an arbitrary finite set of dyadic intervals and by$\langle.,.\rangle$ ·,· we denoted the complex scalar product.
Similarly, for$\vec{\lambda}, \vec{t_1}, \vec{t_2}, \vec{t_3} \in [0,1]^2$ and $\vec{j} \in \{1,2,3\}^2$, we define the discretized dyadic bi-parameter paraproduct of "type $\vec{j}$"

$$\Pi^{\vec{j}}_{\vec{\lambda},\vec{t_1},\vec{t_2},\vec{t_3}} = \Pi^{j'}_{\lambda',t'_1,t'_2,t'_3} \otimes \Pi^{j''}_{\lambda'',t''_1,t''_2,t''_3}$$

by

$$\Pi^{\vec{j}}_{\vec{\lambda},\vec{t_1},\vec{t_2},\vec{t_3}}(f,g) = \sum_{R\in\vec{\mathcal{D}}} \frac{1}{|R|^{1/2}} \langle f, \Phi^1_{R_{\vec{\lambda},\vec{t_1}}}\rangle \langle g, \Phi^2_{R_{\vec{\lambda},\vec{t_2}}}\rangle \Phi^3_{R_{\vec{\lambda},\vec{t_3}}}, \quad (25)$$

where this time f, g are complex-valued measurable functions on $\mathbb{R}^2$, $R = I \times J$ are dyadic rectangles and $\Phi^i_{R_{\vec{\lambda},\vec{t_i}}}$ are given by

$$\Phi^i_{R_{\vec{\lambda},\vec{t_i}}} = \Phi^i_{I_{\lambda',t'_i}} \otimes \Phi^i_{J_{\lambda'',t''_i}}$$

for $i = 1,2,3$. In particular, if $i \neq j'$ then $\int_{\mathbb{R}} \Phi^i_{J_{\lambda',t'_i}}(x)dx = 0$ and if $i \neq j''$ then $\int_{\mathbb{R}} \Phi^i_{J_{\lambda'',t''_i}}(x)dx = 0$ $\vec{\mathcal{D}}$ is an arbitrary finite collection of dyadic rectangles. We will also denote by $\Lambda^{\vec{j}}_{\vec{\lambda},\vec{t_1},\vec{t_2},\vec{t_3}}(f,g,h)$ the trilinear form given by

$$\Lambda^{\vec{j}}_{\vec{\lambda},\vec{t_1},\vec{t_2},\vec{t_3}}(f,g,h) = \int_{\mathbb{R}^2} \Pi^{\vec{j}}_{\vec{\lambda},\vec{t_1},\vec{t_2},\vec{t_3}}(f,g)(x,y)h(x,y)dxdy. \quad (26)$$

In [18] we showed that Theorem(1.2.2) can be reduced to the following Proposition.
Proposition (1.2.5)[12]: Fix $\vec{j} \in \{1,2,3\}^2$ and let $1 < p,q < \infty$ be two numbers arbitrarily close to1. Let also, $f \in L^p$, $\|f\|_p = 1$, $g \in L^q$, $\|g\|_q = 1$ and $E \subseteq \mathbb{R}^2$, $|E| = 1$. Then, there exists a subset $E' \subseteq E$ with $|E'| \sim 1$ such that

$$\left|\Lambda^{\vec{j}}_{\vec{\lambda},\vec{t_1},\vec{t_2},\vec{t_3}}(f,g,h)\right| \lesssim 1 \quad (27)$$

uniformly in the parameters $\vec{\lambda}, \vec{t_1}, \vec{t_2}, \vec{t_3} \in [0,1]^2$, where $h := \chi_{E'}$.

It is therefore enough to prove the above Proposition(1.2.5), in order to complete the proof of our main Theorem(1.2.2). Since all the cases are similar, we assume as in [18] that $\vec{j} = (1,2)$.

To construct the desired set E', we need to recall the "maximal-square", "square-maximal" and "square-square" functions considered in [18].

For $(x,y) \in \mathbb{R}^2$ define

$$MS(f)(x,y) = \sup_I \frac{1}{|I|^{1/2}} \left(\sum_{J:R=I\times J\in\vec{\mathcal{D}}} \sup_{\vec{\lambda},\vec{t_1}} \frac{\left|\langle f, \Phi^1_{R_{\vec{\lambda},\vec{t_i}}}\rangle\right|^2}{|J|} \chi_J(y) \right) \chi_I(x), \quad (28)$$

$$SM(g)(x,y) = \left(\sum_I \frac{\sup_{J:R=I\times J\in\vec{\mathcal{D}}} sup_{\vec{\lambda},\vec{t_2}} \frac{\left|\langle g, \Phi^2_{R_{\vec{\lambda},\vec{t_2}}}\rangle\right|^2}{|J|} \chi_J(y)}{|I|} \chi_I(x) \right)^{1/2} \quad (29)$$

and

$$SS(h)(x,y) = \left(\sum_{R\in\vec{\mathcal{D}}} \sup_{\vec{\lambda},\vec{t_3}} \frac{\left|\langle h, \Phi^3_{R_{\vec{\lambda},\vec{t_3}}}\rangle\right|^2}{|R|} \chi_R(x,y) \right)^{1/2}. \quad (30)$$

Then, we also recall (see[19]) the bi-parameter Hardy-Littlewood maximal function

$$MM(F)(x,y) = \sup_{(x,y)\in I\times J} \frac{1}{|I||J|}\int_{I\times J} |F(x',y')|\, dx'dy'. \tag{31}$$

The following simple estimates explain the appearance of these functions. In particular, we will see that our desired bounds in Theorem (1.2.2) can be easily obtained as long as all the indices involved are strictly between 1 and ∞.

We start by recalling the following basic inequality, [18]. If Π^1 is a one-parameter paraproduct of "type1" given by

$$\Pi^1(f_1,f_2) = \sum_I \frac{1}{|I|^{1/2}} \langle f_1,\Phi_I^1\rangle\langle f_2,\Phi_I^2\rangle\Phi_I^3 \tag{32}$$

then we can write

$$|\Lambda^1(f_1,f_2,f_3)| = \left|\int_{\mathbb{R}} \Pi^1(f_1,f_2)(x)f_3(x)dx\right| \lesssim \sum_I \frac{1}{|I|^{1/2}} |\langle f_1,\Phi_I^1\rangle||\langle f_2,\Phi_I^2\rangle||\langle f_3,\Phi_I^3\rangle|$$

$$= \int_{\mathbb{R}} \left(\sum_I \frac{|\langle f_1,\Phi_I^1\rangle|}{|I|^{1/2}} \frac{|\langle f_2,\Phi_I^2\rangle|}{|I|^{1/2}} \frac{|\langle f_3,\Phi_I^3\rangle|}{|I|^{1/2}} \chi_I(x)\right) dx$$

$$\lesssim \int_{\mathbb{R}} M(f_1)(x)S(f_2)(x)S(f_3)(x)\, dx \quad . \tag{33}$$

where M denotes the Hardy-Littlewood maximal function and S is the square function of Littlewood and Paley. In particular, we easily see that $\Pi^1: L^p \times L^q \to L^r$ for any $1 < p,q,r < \infty$ satisfying $1/p + 1/q = 1/r$. Analogous estimates hold for any other type of paraproducts Π^j for $j = 1\ ,2,3$.

Similarly, for the bi-parameter paraproduct $\Pi^{(1,2)}$ of "type(1,2)" formally defined by $\Pi^{(1,2)} = \Pi^1 \otimes \Pi^2$ one obtains the inequalities

$$\left|\Lambda^{(1,2)}(f_1,f_2,f_3)\right| = \left|\int_{\mathbb{R}} \Pi^{(1,2)}(f_1,f_2)(x,y)f_3(x,y)dxdy\right|$$

$$\lesssim \cdots \lesssim \int_{\mathbb{R}^2} MS(f_1)(x,y)SM(f_2)(x,y)S(f_3)(x,y)\, dxdy. \tag{34}$$

and analogous estimates hold for any other type of paraproducts $\Pi^{\vec{j}}$ for $\vec{j} \in \{1,2,3\}^2$. It is important that all these MS, SM and SS functions are bounded on L^p for any $1 < p < \infty$. We recall the proof of this fact here (see[18]). We start with $SM(f_2)(x,y)$. It can be written as

$$SM(f_2)(x,y) = \left(\sum_{\tilde{I}} \frac{\sup_{\tilde{J}} \frac{\left|\langle f_2,\Phi_{\tilde{I}}^2\otimes\Phi_{\tilde{J}}^2\rangle\right|^2}{|\tilde{J}|}\chi_{\tilde{J}}(y)}{|\tilde{I}|}\chi_{\tilde{I}}(x)\right)^{1/2} \tag{35}$$

$$\lesssim \left(\sum_{\tilde{I}} M\left(\frac{\langle f_2,\Phi_{\tilde{I}}^2\rangle}{|\tilde{I}|^{1/2}}\right)^2 (y)\chi_{\tilde{I}}(x)\right)^{1/2}$$

where $\tilde{I}$ and $\tilde{J}$ are the intervals where the corresponding supremums over $\vec{\lambda},\vec{t}_2 \in [0,1]^2$ in (29) are attained.

In particular, by using Fefferman-Stein[14] and Littlewood-Paley [19] inequalities, we have

$$\|SM(f_2)\|_p \lesssim \left\| \left(\sum_{\tilde{I}} M \left(\frac{\langle f_2, \Phi_{\tilde{I}}^2 \rangle}{|\tilde{I}|^{1/2}} \right)^2 (y) \chi_{\tilde{I}}(x) \right)^{1/2} \right\|_p \tag{36}$$
$$\lesssim \left\| \left(\sum_{\tilde{I}} \frac{|\langle f_2, \Phi_{\tilde{I}}^2 \rangle|^2}{|\tilde{I}|} (y) \chi_{\tilde{I}}(x) \right)^{1/2} \right\|_p \lesssim \|f_2\|_p$$

for any $1 < p < \infty$. Then, we observe that the MS function is pointwise smaller than a certain SM type function and hence bounded on L^p, while the SS function is a classical double square function and its boundedness on L^p spaces is well known, [2]. As a consequence, it follows as before that $\Pi^1 : L^p \times L^q \to L^r$ as long as $1 < p, q, r < \infty$ with $1/p + 1/q = 1/r$.

It remains to prove Proposition(1.2.5). First, we state the following Lemma.

It is the main new ingredient which allows us to simplify our previous argument in [18]. Using it, we can decompose our trilinear form in(25) as

$$\Lambda^{\vec{j}}_{\vec{\lambda},\vec{t_1},\vec{t_2},\vec{t_3}}(f,g,h) = \sum_{\vec{k}\in\mathbb{N}^2} 2^{-1000|\vec{k}|} \sum_{R\in\vec{\mathcal{D}}} \frac{1}{|R|^{\frac{1}{2}}} \langle f, \Phi^1_{R_{\vec{\lambda},\vec{t_1}}} \rangle \langle g, \Phi^2_{R_{\vec{\lambda},\vec{t_2}}} \rangle \langle h, \Phi^{3,\vec{k}}_{R_{\vec{\lambda},\vec{t_3}}} \rangle, \tag{37}$$

where the new functions $\Phi^{3,\vec{k}}_{R_{\vec{\lambda},\vec{t_3}}}$ have basically the same structure as the old $\Phi^3_{R_{\vec{\lambda},\vec{t_3}}}$ but they also have the additional property that $\operatorname{supp}\left(\Phi^{3,\vec{k}}_{R_{\vec{\lambda},\vec{t_1}}}\right) \subseteq 2^{\vec{k}} R_{\vec{\lambda},\vec{t_3}}$. We denoted by $2^{\vec{k}} R_{\vec{\lambda},\vec{t_3}} \coloneqq 2^{k_1} I_{\lambda',t_3'} \times 2^{k_2} I_{\lambda'',t_3''}, \vec{k} = (k_1, k_3)$ and $|\vec{k}| = k_1 + k_2$.

Fix now f, g, E, p, q as in Proposition(1.2.5). For each $\vec{k} \in \mathbb{N}^2$ define

$$\Omega_{-5|\vec{k}|} = \left\{(x,y) \in \mathbb{R}^2 : MS(f)(x,y) > C2^{5|\vec{k}|}\right\} \tag{38}$$
$$\bigcup \left\{(x,y) \in \mathbb{R}^2 : SM(g)(x,y) > C2^{5|\vec{k}|}\right\}.$$

Also, define

$$\tilde{\Omega}_{-5|\vec{k}|} = \left\{(x,y) \in \mathbb{R}^2 : MM\left(\chi_{\Omega_{-5|\vec{k}|}}\right)(x,y) > \frac{1}{100}\right\} \tag{39}$$

and then

$$\tilde{\tilde{\Omega}}_{-5|\vec{k}|} = \left\{(x,y) \in \mathbb{R}^2 : MM\left(\chi_{\tilde{\Omega}_{-5|\vec{k}|}}\right)(x,y) > \frac{1}{2|\vec{k}|}\right\}. \tag{40}$$

Finally, we denote by

$$\Omega = \bigcup_{\vec{k}\in\mathbb{N}^2} \tilde{\Omega}_{-5|\vec{k}|}.$$

It is clear that $|\Omega| < 1/2$ if C is a big enough constant, which we fix from now on. Then, define $E' \coloneqq E \setminus \Omega$ and observe that $|E'| \sim 1$. We now want to show that the corresponding expression in(27) is $O(1)$ uniformly in the parameters $\vec{\lambda}, \vec{t_1}, \vec{t_2}, \vec{t_3} \in [0,1]^2$. Since our argument will not depend on these parameters, we can assume for simplicity that they are all zero and in this case we will write Φ_R^i instead of $\Phi_{R_{\vec{\lambda},\vec{t_i}}}^i$ for $i = 1, 2$ and $\Phi_R^{3,\vec{k}}$ instead of c.

Fix then $\vec{k} \in \mathbb{N}^2$ and look at the corresponding inner sum in (37). We split it into two parts as follows. Part I sums over those rectangles R with the property that

$$R \cap \widetilde{\Omega}^c_{-5|\vec{k}|} \neq \emptyset \tag{41}$$

while Part II sums over those rectangles with the property that

$$R \cap \widetilde{\Omega}^c_{-5|\vec{k}|} = \emptyset. \tag{42}$$

We observe that Part II is identically equal to zero, because if $R \cap \widetilde{\Omega}^c_{-5|\vec{k}|} \neq \emptyset$ then $R \subseteq \widetilde{\Omega}_{-5|\vec{k}|}$ and in particular this implies that $2^{\vec{k}}R \subseteq \widetilde{\widetilde{\Omega}}_{-5|\vec{k}|}$ which is a set disjoint from E'. It is therefore enough to estimate Part I only. This can be done by using the technique developed in [18].

Since if $R \cap \widetilde{\Omega}^c_{-5|\vec{k}|} \neq \emptyset$, it follows that

$$\frac{\left|R \cap \Omega_{-5|\vec{k}|}\right|}{|R|} < \frac{1}{100} \text{ or equivalently, } \left|R \cap \Omega^c_{-5|\vec{k}|}\right| > \frac{99}{100}|R|.$$

We describe three decomposition procedures, one for each function f, g, h. Later on, we will combine them, in order to handle our sum. First, define

$$\Omega_{-5|\vec{k}|+1} = \left\{(x,y) \in \mathbb{R}^2 : MS(f)(x,y) > \frac{C2^{5|\vec{k}|}}{2^1}\right\}$$

and set

$$\mathrm{T}_{-5|\vec{k}|+1} = \left\{R \in \vec{\mathcal{D}} : \left|R \cap \Omega_{-5|\vec{k}|+1}\right| > \frac{1}{100}|R|\right\},$$

then define

$$\Omega_{-5|\vec{k}|+2} = \left\{(x,y) \in \mathbb{R}^2 : MS(f)(x,y) > \frac{C2^{5|\vec{k}|}}{2^2}\right\}$$

and set

$$\mathrm{T}_{-5|\vec{k}|+2} = \left\{R \in \vec{\mathcal{D}} \setminus \mathrm{T}_{-5|\vec{k}|+1} : \left|R \cap \Omega_{-5|\vec{k}|+2}\right| > \frac{1}{100}|R|\right\},$$

and so on. The constant $C > 0$ is the one in the definition of the setE' above.

Since there are finitely many rectangles, this algorithm ends after a while, producing the sets $\{\Omega_n\}$ and$\{\mathrm{T}_n\}$ such that $\vec{\mathcal{D}} = \bigcup_n \mathrm{T}_n$.

Independently, define

$$\Omega'_{-5|\vec{k}|+1} = \left\{(x,y) \in \mathbb{R}^2 : SM(g)(x,y) > \frac{C2^{5|\vec{k}|}}{2^1}\right\}$$

and set

$$\mathrm{T}'_{-5|\vec{k}|+1} = \left\{R \in \vec{\mathcal{D}} : \left|R \cap \Omega'_{-5|\vec{k}|+1}\right| > \frac{1}{100}|R|\right\},$$

then define

$$\Omega'_{-5|\vec{k}|+2} = \left\{(x,y) \in \mathbb{R}^2 : SM(g)(x,y) > \frac{C2^{5|\vec{k}|}}{2^2}\right\}$$

and set

$$\mathrm{T}'_{-5|\vec{k}|+2} = \left\{R \in \vec{\mathcal{D}} \setminus \mathrm{T}'_{-5|\vec{k}|+1} : \left|R \cap \Omega'_{-5|\vec{k}|+2}\right| > \frac{1}{100}|R|\right\},$$

and so on, producing the sets $\{\Omega'_n\}$ and $\{\mathrm{T}'_n\}$ such that $\vec{\mathcal{D}} = \bigcup_n \mathrm{T}'_n$. We would like to have such a decomposition available for the function h also. To do this, we first need to construct the analogue of the set $\Omega_{-5|\vec{k}|}$, for it. Pick $N > 0$ a big enough integer such that for every $R \in \vec{\mathcal{D}}$ we have $\left|R \cap \Omega''_{-N}\right| > \frac{99}{100}|R|$ where we defined

$$\Omega''_{-N} = \left\{(x,y) \in \mathbb{R}^2 : SS^{\vec{k}}(h)(x,y) > C2^N\right\}.$$

Here $SS^{\vec{k}}$ denotes the same "square-square" function defined in (30) but with the functions $\Phi^{3,\vec{k}}_{R_{\vec{\lambda},\vec{t_3}}}$ instead of $\Phi^3_{R_{\vec{\lambda},\vec{t_3}}}$. Then, similarly to the previous algorithms, we define

$$\Omega''_{-N+1} = \left\{(x,y) \in \mathbb{R}^2 : SS^{\vec{k}}(h)(x) > \frac{C2^N}{2^1}\right\}$$

and set

$$\mathrm{T}''_{-N+1} = \left\{R \in \vec{\mathcal{D}} : \left|R \cap \Omega''_{-N+1}\right| > \frac{1}{100}|R|\right\},$$

then define

$$\Omega''_{-N+2} = \left\{x \in \mathbb{R}^2 : SS^{\vec{k}}(h)(x) > \frac{C2^N}{2^2}\right\}$$

and set

$$\mathrm{T}''_{-N+2} = \left\{R \in \vec{\mathcal{D}} \setminus \mathrm{T}''_{-N+1} : \left|R \cap \Omega''_{-N+2}\right| > \frac{1}{100}|R|\right\},$$

and so on, constructing the sets $\{\Omega''_n\}$ and $\{\mathrm{T}''_n\}$ such that $\vec{\mathcal{D}} = \bigcup_n \mathrm{T}''_n$.
Then we write Part I as

$$\sum_{\substack{n_1,n_2>-5|\vec{k}| \\ n_3>-N}} \sum_{R \in T_{n_1,n_2,n_3}} \frac{1}{|R|^{3/2}} |\langle f, \Phi^1_R\rangle||\langle g, \Phi^2_R\rangle||\langle h, \Phi^3_R\rangle|, \tag{43}$$

where $T_{n_1,n_2,n_3} := T_{n_1} \cap \mathrm{T}'_{n_2} \cap \mathrm{T}''_{n_3}$. Now, if R belongs to T_{n_1,n_2,n_3} this means in particular that R has not been selected at the previous $n_1 - 1, n_2 - 1$ and $n_3 - 1$ steps respectively, which means that

$$\left|R \cap \Omega_{n_1-1}\right| < \frac{1}{100}|R|, \left|R \cap \Omega'_{n_2-1}\right| < \frac{1}{100}|R| \text{and} \left|R \cap \Omega''_{n_3-1}\right| < \frac{1}{100}|R|$$

or equivalently,

$$\left|R \cap \Omega^c_{n_1-1}\right| > \frac{99}{100}|R|, \left|R \cap \Omega'^c_{n_2-1}\right| > \frac{99}{100}|R| \text{and} \left|R \cap \Omega''^c_{n_3-1}\right| < \frac{99}{100}|R|.$$

But this implies that

$$\left|R \cap \Omega^c_{n_1-1} \cap \Omega'^c_{n_2-1} \cap \Omega''^c_{n_3-1}\right| > \frac{97}{100}|R|. \tag{44}$$

In particular, using (44), the term in (43) is smaller than

$$\sum_{\substack{n_1,n_2>-5|\vec{k}| \\ n_3>-N}} \sum_{R \in T_{n_1,n_2,n_3}} \frac{1}{|R|^{3/2}} |\langle f, \Phi^1_R\rangle||\langle g, \Phi^2_R\rangle||\langle h, \Phi^3_R\rangle| \times \left|R \cap \Omega^c_{n_1-1} \cap \Omega'^c_{n_2-1} \cap \Omega''^c_{n_3-1}\right|$$

$$= \sum_{\substack{n_1,n_2>-5|\vec{k}| \\ n_3>-N}} \int_{\Omega^c_{n_1-1}\cap\Omega'^c_{n_2-1}\cap\Omega''^c_{n_3-1}} \sum_{R\in T_{n_1,n_2,n_3}} \frac{1}{|R|^{3/2}} |\langle f,\Phi^1_R\rangle||\langle g,\Phi^2_R\rangle|$$

$$\times |\langle h,\Phi^3_R\rangle|\chi_R(x,y)dxdy$$

$$\lesssim \sum_{\substack{n_1,n_2>-5|\vec{k}| \\ n_3>-N}} \int_{\Omega^c_{n_1-1}\cap\Omega'^c_{n_2-1}\cap\Omega''^c_{n_3-1}\cap\Omega_{T_{n_1,n_2,n_3}}} MS(f)(x,y)SM(g)(x,y)\times SS^{\vec{k}}(h)(x,y)dxdy$$

$$\lesssim \sum_{\substack{n_1,n_2>-5|\vec{k}| \\ n_3>-N}} 2^{-n_1}2^{-n_2}2^{-n_3}\left|\Omega_{T_{n_1,n_2,n_3}}\right|, \tag{45}$$

Where

$$\Omega_{T_{n_1,n_2,n_3}} := \bigcup_{R\in T_{n_1,n_2,n_3}} R.$$

On the other hand we can write

$$\left|\Omega_{T_{n_1,n_2,n_3}}\right| \le \left|\Omega_{T_{n_1}}\right| \le \left|\left\{(x,y)\in\mathbb{R}^2 : MM\left(\chi_{\Omega_{n_1}}\right)(x,y) > \frac{1}{100}\right\}\right|$$

$$\lesssim \left|\Omega_{n_1}\right| = \left|\left\{\left\{(x,y)\in\mathbb{R}^2 : MS(f)(x,y) > \frac{C}{2^{n_1}}\right\}\right\}\right| \lesssim 2^{n_1 p}.$$

Similarly, we have

$$\left|\Omega_{T_{n_1,n_2,n_3}}\right| \lesssim 2^{n_2 q}$$

and also

$$\left|\Omega_{T_{n_1,n_2,n_3}}\right| \lesssim 2^{n_2 \alpha},$$

for every $\alpha > 1$. Here we used the fact that all the operators SM, MS,$SS^{\vec{k}}$, MM are bounded on L^s (independently of $\vec{k}$) as long as $1 < s < \infty$ and also that$|E'|\sim 1$. In particular, it follows that

$$\left|\Omega_{T_{n_1,n_2,n_3}}\right| \lesssim 2^{n_1 p\theta_1}2^{n_2 q\theta_2}2^{n_3\alpha\theta_3} \tag{47}$$

for any $0 \le \theta_1,\theta_2,\theta_3 < 1$, such that $\theta_1+\theta_2+\theta_3 = 1$.
Now we split the sum in (45) into

$$\sum_{\substack{n_1,n_2>-5|\vec{k}| \\ n_3>-N}} 2^{-n_1}2^{-n_2}2^{-n_3}\left|\Omega_{T_{n_1,n_2,n_3}}\right| + \sum_{\substack{n_1,n_2>-5|\vec{k}| \\ 0>n_3>-N}} 2^{-n_1}2^{-n_2}2^{-n_3}\left|\Omega_{T_{n_1,n_2,n_3}}\right|. \tag{47}$$

To estimate the first term in (47) we use the inequality (46) in the particular case $\theta_1 = \theta_2 = 1/2, \theta_3 = 0$, while to estimate the second term we use (46) for $\theta_j, j = 1\,,2,3$ such that $1 - p\theta_1 > 0$, $1 - q\theta_2 > 0$ and $\alpha\theta_3 - 1 > 0$. With these choices, the sum in (47) is $O\left(2^{10|\vec{k}|}\right)$ and this makes the expression in(37) to be $O(1)$, after summing over$\vec{k}\in\mathbb{N}^2$.
This completes our proof.
It is now clear that our argument works equally well in all dimensions. In the general case, exactly as in [18], one first reduces the study of the operator $T_m^{(d)}$ to the study of generic d-parameter dyadic paraproducts $\Pi^{\vec{j}}$ for $\vec{j} = (j_1,..,j_d) \in \{1,2,3\}^d$formally defined by$\Pi^{\vec{j}} = \Pi^{j_1}\otimes\ldots\otimes\Pi^{j_d}$. Then, one observes as before, by using the linear theory and Fefferman- Stein inequality, that all the corresponding “square and maximal” type

functions which naturally appear in inequalities analogous to (33), (34) are bounded in L^p for $1 < p < \infty$ (in fact, as before, it is enough to observe this in the $SS\dots SMM\dots M$ case, because all the other expressions are pointwise smaller quantities).

Having all these ingredients. Finally, the n-linear case follows in the same way.

Lemma (1.2.6)[12]: Let $J \subseteq \mathbb{R}$ be an arbitrary interval. Then, every bump function $\emptyset_J$ adapted to J can be written as

$$\emptyset_J = \sum_{k\in\mathbb{N}} 2^{-1000k\emptyset_J^k} \tag{48}$$

where for each $k \in \mathbb{N}$,$\emptyset_J^k$ is also a bump adapted to J but with the additional property that $\mathrm{supp}\left(\emptyset_J^k\right) \subseteq 2^k J$. Moreover, if we assume $\int_R \emptyset_J(x)\,dx = 0$ then all the functions $\emptyset_J^k$ can be chosen so that$\int_{\mathbb{R}} \emptyset_J^k(x)\,dx = 0$ for every $k \in \mathbb{N}$.

Proof. Fix $J \subseteq \mathbb{R}$ an interval and let ϕ_J be a bump function adapted to J. Considerψ a smooth function such that $\mathrm{supp}(\psi) \subseteq [-1/2,1/2]$ and $\psi = 1$ on$[-1/4,1/4]$. If $I \subseteq \mathbb{R}$ is a generic interval with center x_I, we denote by$\psi_I$the function defined by

$$\psi_I(x) = \psi\left(\frac{x - x_I}{|I|}\right). \tag{49}$$

Since

$$1 = \psi_J + \left(\psi_{2J} - \psi_J\right) + \left(\psi_{2^2J} - \psi_{2J}\right) + \cdots$$

it follows that

$$\phi_J = \phi_J.\psi_J + \sum_{k=1}^{\infty} \phi_J\,.\left(\psi_{2^kJ} - \psi_{2^{k-1}J}\right)$$

$$= \phi_J.\psi_J + \sum_{k=1}^{\infty} 2^{-1000k}\,.\left[2^{-1000k}\phi_J.\left(\psi_{2^kJ} - \psi_{2^{k-1}J}\right)\right] \coloneqq \sum_{k=0}^{\infty} 2^{-1000k}\phi_J^k$$

and it is easy to see that all theϕ_J^k functions are bumps adapted to$J$, having the property that $\mathrm{supp}\left(\phi_J^k\right) \subseteq 2^k J$.

Suppose now that in addition we have$\int_{\mathbb{R}} \phi_J(x)dx = 0$. This time, we write

$$\phi_J = \phi_J.\psi_J + \phi_J.\left(1 - \psi_J\right)$$

$$= \left[\phi_J.\psi_J - \left(\frac{1}{\int_{\mathbb{R}} \psi_J(x)dx}\,.\int_{\mathbb{R}} \phi_J(x)\psi_J(x)dx\right).\psi_J\right]$$

$$+\left[\left(\frac{1}{\int_{\mathbb{R}} \psi_J(x)dx}.\int_{\mathbb{R}} \phi_J(x)\psi_J(x)dx\right).\psi_J + \phi_J\left(1 - \psi_J\right)\right]$$

$$\coloneqq \phi_J^0 + R_J^0.$$

Clearly, by construction we have that $\int_{\mathbb{R}} \phi_J^0(x)dx = 0$and therefore

$$\int_{\mathbb{R}} R_J^0(x)dx = 0.$$

Moreover, ϕ_J^0 is a bump adapted to the interval J having the property that $\mathrm{supp}(\phi_J^0) \subseteq J$.

On the other hand, since

$$\left|\frac{1}{\int_{\mathbb{R}} \psi_J(x)dx}.\int_{\mathbb{R}} \phi_J(x)\psi_J(x)dx\right| =$$

$$= \left| \frac{1}{\int_{\mathbb{R}} \psi_J(x)dx} . \int_{\mathbb{R}} \phi_J(x)dx \left(1 - \psi_J(x)\right) dx \right| \lesssim 2^{-1000} \tag{50}$$

follows that $\|R_J^0\|_\infty \lesssim 2^{-1000}$.

Then, we perform a similar decomposition for the "rest function" R_J^0, but this time we localize it on the larger interval $2J$. We have

$$R_J^0 = R_J^0 . \psi_{2J} + R_J^0 . \left(1 - \psi_{2J}\right)$$
$$= \left[R_J^0 . \psi_{2J} - \left(\frac{1}{\int_{\mathbb{R}} \psi_J(x)dx} . \int_{\mathbb{R}} R_J^0(x)\psi_{2J}(x)dx \right) . \psi_{2J} \right]$$
$$+ \left[\left(\frac{1}{\int_{\mathbb{R}} \psi_{2J}(x)dx} . \int_{\mathbb{R}} R_J^0(x)\psi_{2J}(x)dx \right) . \psi_{2J} + R_J^0 . \left(1 - \psi_{2J}\right) \right]$$
$$:= 2^{-1000}\phi_J^1 + R_J^1.$$

As before, we observe that $\int_{\mathbb{R}} \phi_J^1(x)dx = 0$ and also $\int_{\mathbb{R}} R_J^1(x)dx = 0$. Moreover, ϕ_J^1 is a bump adapted to J whose support lies in $2J$ and $\|R_J^1\|_\infty \lesssim 2^{-1000.2}$. Iterating this procedure N times, we obtain the decomposition

$$\phi_J = \sum_{k=0}^{N} 2^{-1000k}\phi_J^k + R_J^N \tag{51}$$

where all the functions ϕ_J^k are bumps adapted to J with $\int_{\mathbb{R}} \phi_J^k(x)dx = 0$ and $\text{supp}(\phi_J^k) \subseteq 2^k J$, while $\|R_J^N\|_\infty \lesssim 2^{-1000N}$.

This completes the proof of the Lemma.

Chapter 2
Algorithm and Nonlinear N-Term Approximation

We establish the order of approximation by this algorithm in weighted L_q-spaces. We show an associated with a finite mask and a rather general matrix dilation $A \in GL_n(\mathbb{Z})$.

Section (2.1): Nonlinear Approximation by Piecewise Polynomials

The UDA was firstly developed for nonlinear approximation by compactly supported refinable functions, see [21] and [23]. We present its version intended for approximation by piecewise polynomials. We believe that this modification of the UDA will have important applications to Numerical Analysis and deserves to be presented to experts in this field. On the other hand, an approximation theorem to be proved in the present has important applications to Approximation theory. One can derive from it the corresponding optimal approximation results for functions from Besov spaces (see [22]).

The algorithm considered makes use of a collection $\mathcal{T} := \{\mathcal{T}_j; j \in \mathbb{Z}_+\}$ of subsequent subdivisions of measurable set $\Omega \subset \mathbb{R}^d$. This collection is equipped with the structure of ordered tree. The input of the algorithms consists of an integer $N \geq 1$and a set function

$$F: \mathcal{T} \to X,$$

where X is a subspace of polynomials in$\mathbb{R}^d$. The output is a function

$$F_N: \mathcal{T} \to X$$

such that

$$\operatorname{supp} F_N := \{\omega \in \mathcal{T}; F_N(\omega) \neq 0\} \leq 4N.$$

Using this we then introduce an approximation aggregate

$$T_N(F) := \sum_{\omega} F_N(\omega)\chi_\omega, \qquad (1)$$

where$\chi_\omega$here and below stands for the Characteristic function of$\omega \subset \mathbb{R}^d$.

If, in particular,$X := \mathcal{P}_{s,d}$, the space of polynomials of degree s in$\mathbb{R}^d$, the aggregate$T_N(F)$ becomes a piecewise polynomial of degree s with $4N$ “pieces”.

However,$\operatorname{supp} T_N(F)$does not form a subdivision ofΩ and therefore these pieces relate to subsets that may differ from subsets $\omega \in \mathcal{T}$.

We present the description of the algorithm. We apply the algorithm to establish a general approximation theorem for functions $f \in L_p(\Omega)$, $0 < p < \infty$,presented in a form

$$f = \sum_{\omega \in \mathcal{T}} f\omega\chi_\omega \left(\text{convergence in} L_p\right).$$

In this case the aforementioned function F is defined by $F(\omega) := f_\omega$, $\omega \in \mathcal{T}$,and we estimate the rate of approximation of f by $T_N(F)$in a weighted L_p-norm,$p < q \leq \infty$.In the subsequent [22], this theorem is applied to derive the corresponding approximation results for Besov spaces. The aggregate(1) in this case yields an optimal in order rate of approximation showing the efficiency of the algorithm.

We begin with the introduction of a tree of subdivisions $\mathcal{T}$. Let Ω be a subset of $\mathbb{R}^d$ with finite d-measure$|\Omega|$,and $\mathcal{T}$ is a collection of subsets of Ω.$\mathcal{T}$ is called a tree of subdivisions for Ω, if the following holds.

(i) Every $\omega', \omega'' \in \mathcal{T}$either nonoverlap, i.e.,

$$|\omega' \cap \omega''| = 0,$$

or one of them is contained in the other. This condition introduces an ordered tree structure on $\mathcal{T}$. Actually, we regard subsets of $\mathcal{T}$ as vertices and connect ω', $\omega'' \in \mathcal{T}$by the edge directed from ω' to ω''(written $\omega' \to \omega''$) if $\omega' \subset \omega''$and there is no set of $\mathcal{T}$ situated between them different fromω' andω''.

Assume that $\Omega \in \mathcal{T}$. Then $\mathcal{T}$ is an ordered tree with the root Ω. Hence each $\omega \in \mathcal{T}$can be connected with Ω by a unique array. In other words, there is a collection$\{\omega_j: 1 \le j \le n\} \subset \mathcal{T}$ such that $\omega_1 \to \omega_2 \to \cdots \to \omega_n$(i.e., this collection is an array), and$\omega_1 = \omega$ and $\omega_n = \Omega$. Because of uniqueness of this array one can correctly define a height $h: \mathcal{T} \to \mathbb{Z}_+$ letting

$$h(\omega) := (\mathrm{card} A) - 1, \tag{2}$$

where A is the array connecting ω and Ω. Specially,$h(\Omega) = 0$.

Set now for $j \in \mathbb{Z}_+$

$$\mathcal{T}_j := \{\omega \in \mathcal{T}: h(\omega) = j\}. \tag{3}$$

These form a partition of $\mathcal{T}$:

$$\mathcal{T}_j \cap \mathcal{T}_{j'} = \phi, \text{if} j \ne j' \text{and} \mathcal{T} = \bigcup_{j \in \mathbb{Z}_+} \mathcal{T}_j. \tag{4}$$

(ii) For every $j \in \mathbb{Z}_+$

$$\mathrm{supp}\, \mathcal{T}_j := \bigcup_{\omega \in \mathcal{T}} \omega = \Omega. \tag{5}$$

In other words,$\{\mathcal{T}_j\}$is a sequence of consequent subdivisions of Ω.

To formulate the last condition one set

$$S(\omega) := \{\omega' \in \mathcal{T}: \omega' \to \omega\}. \tag{6}$$

In accordance with the terminology of Graph Theory, each element of this set is a son of ω(and ω is its father).

(iii) There is a constant $C(\mathcal{T})$ such that for every$\omega \in \mathcal{T}$

$$1 < \mathrm{card} S(\omega) \le C(\mathcal{T}). \tag{7}$$

Definition (2.1.1)[20]: A collection $\mathcal{T}$ of subsets of Ω is said to be a tree of subdivision, if it meets the conditions (i)-(iii).

Let $w: \mathcal{T} \to \mathbb{R}_+$be a weight, and $0 < p \le \infty$. Introduce a space $\ell_p^w(\mathcal{T}; X)$ of functions $F: \mathcal{T} \to \mathbb{R}$ defined by the quasinorm

$$\|F\|_{p,w} := \left\{ \sum_{\omega \in : \mathcal{T}} \left(w(\omega) \sup_{\omega} |F(\omega)| \right)^p \right\}^{\frac{1}{p}}. \tag{8}$$

Note that the ω-term of this sum with unbounded ω is finite, only if the polynomial $F(\omega)$ is constant (or $w(\omega) = 0$). To avoid unnecessary complications we assume that Ω is bounded.

The input of the algorithm comprises a fixed $F \in \ell_p^w(\mathcal{T}; X)$and integer $N \ge 1$. Because of homogeneity of(8) we can and do assume that

$$\|F\|_{p,w} = 1. \tag{9}$$

Given F, we introduce a cost function $\mathcal{J}: 2^{\mathcal{T}} \mathbb{R}_+$by

$$\mathcal{J}(S) = \mathcal{J}(F; S) := \left\{ \sum_{\omega \in S} \left(w(\omega) \sup_{\omega} |F(\omega)| \right)^p \right\}. \tag{10}$$

Specially, for the subset

$$\mathcal{T}(\omega) := \{\omega' \in \mathcal{T} : \omega' \subset \omega\} \tag{11}$$

we simplify this notation by setting

$$\mathcal{I}(\omega) := \mathcal{I}(\mathcal{T}(\omega)). \tag{12}$$

Note that $\mathcal{I}(\omega) \neq \mathcal{I}(\{\omega\}) := (w(\omega)|f(\omega)|)^p$, and

$$\mathcal{I}(\Omega) = 1, \tag{13}$$

See (8) and (9).

We first introduce the subtree

$$\mathcal{G}_N := \{\omega \in \mathcal{T} : \mathcal{I}(\omega) \geq N^{-1}\}. \tag{14}$$

$\mathcal{G}_N$ is nonempty and has the root Ω by (13). Since $\mathcal{T}$ is an ordered set, the set $\mathcal{M}_N$ of minimal elements of $\mathcal{G}_N$ is well-defined. Hence for each $\omega \in \mathcal{M}_N$ and every its son ω'

$$\mathcal{I}(\omega) \geq N^{-1}, \text{while } \mathcal{I}(\omega') < N^{-1}. \tag{15}$$

Numerate the elements of $\mathcal{M}_N$ in some order

$$\mathcal{M}_N := \{\omega_j^{\min} : 1 \leq j \leq m_N\}. \tag{16}$$

Since the subsets of $\mathcal{M}_N$ nonoverlap, we have

$$1 = \mathcal{I}(\Omega) \geq \sum_j \mathcal{I}(\omega_j^{\min}) \geq m_N/N,$$

whence

$$m_N \leq N. \tag{17}$$

We partition $\mathcal{G}_N$ in order to obtain a collection of (basic) arrays $\mathcal{B}_N$. An algorithm fulfilling this operation is the main part of our construction.

In its description we will use the notation

$$[\omega, \omega'] := \{\omega_1 \to \omega_2 \to \cdots \to \omega_n\} \tag{18}$$

for the array connecting $\omega (= \omega_1)$ and $\omega' (= \omega_n)$. We also introduce an "open from the top" subarray of this array setting

$$[\omega, \omega') := [\omega, \omega'] \backslash \{\omega'\}. \tag{19}$$

At the first stage we introduce a partition of $\mathcal{G}_N$ into a collection $\mathcal{A} := \{A_j : 0 \leq j \leq m_N\}$ of ("big") arrays satisfying the following conditions:

(a) $\{A_j : 0 \leq j \leq i\}$ is a partition of the set

$$\mathcal{G}_N^i := \bigcup_{s \leq i} [\omega_s^{\min}, \Omega], 0 \leq i \leq m_N;$$

(b) each A_i has a form

$$A_i = [\omega_i^{\min}, \omega),$$

where ω belongs to some $A_{i'}$ with $i' < i$. This ω is called a contact element and is denoted by ω_i^c; hence

$$A_i = [\omega_i^{\min}, \omega_i^c), 1 \leq i \leq m_N. \tag{20}$$

Since $\mathcal{G}_N^i = \mathcal{G}_N$, if $i = m_N$, the collection $\mathcal{A} = \{A_j : 0 \leq j \leq m_N\}$ forms the desired partition of the subtree $\mathcal{G}_N$. Besides, $\mathcal{A}$ determines the set of contact elements

$$C_N := \{\omega_i^c\} \cup \{\Omega\}. \tag{21}$$

As we shall see, some of these may coincide and therefore

$$\operatorname{card} C_N \leq m_N + 1, \tag{22}$$

where the inequality may be strict.

In order to introduce $\mathcal{A}$ we use induction on j starting with

$$A_0 := \{\Omega\} \text{ and } A_1 := [\omega_1^{\min}, \Omega] \backslash A_0 = [\omega_1^{\min}, \Omega).$$

Assume now that we have determined the arrays $A_i, i = 0,1, \dots, j$, satisfying the conditions (a) and (b) with $i \leq j$. Define

$$A_{j+1} := [\omega_{j+1}^{\min}, \Omega] \setminus \left(\bigcup_{i \leq j} A_i \right).$$

Then $\{A_i : 0 \leq i \leq j+1\}$ is Clearly a partition of $\mathcal{G}_N^{j+1}$. Show that A_{j+1} has a form(20). In fact, consider the intersection of $[\omega_{j+1}^{\min}, \Omega]$ with each $[\omega_i^{\min}, \Omega]$, $i \leq j$. Since $\mathcal{G}_N$ is a tree with the root Ω, this intersection is of a form $[\omega_i, \Omega]$, and $\{\omega_i : 1 \leq i \leq j\}$ is a subset of the array $[\omega_{j+1}^{\min}, \Omega]$. Hence this subset inherits the linear order of the last array. If ω_{i_0} is the smallest element of $\{\omega_i\}$ with respect to this order, then

$$A_{j+1} := [\omega_{j+1}^{\min}, \Omega] \setminus \left(\bigcup_{i \leq j} A_i \right) = [\omega_{j+1}^{\min}, \Omega] \setminus \left(\bigcup_{i \leq j} [\omega_i^{\min}, \Omega] \right) = [\omega_{j+1}^{\min}, \omega_{i_0}).$$

Moreover, $\omega_{i_0} \in \bigcup_{i \leq j} A_i$. Hence the induction is complete.
We proceed the refinement of $\mathcal{G}_N$ subdividing each array A_j by the elements of the set $A_j \cap C_N$, $j \geq 1$. We introduce a collection of "open from the top" subarrays $[\omega', \omega'')$ where ω' is either a minimal or contact element, and ω'' is a contact one. The set of these subarrays one denotes by $\mathcal{R}_N$. According to its definition

$$\operatorname{supp} \mathcal{R}_N := \bigcup_{R \in \mathcal{R}_N} R = \mathcal{G}_N \setminus \{\Omega\} \tag{23}$$

and different elements of $\mathcal{R}_N$ do not overlap, i.e., $\mathcal{R}_N$ is a partition of (23).
At the final stage we complete the partition algorithm subdividing each sub-array $R \in \mathcal{R}_N$ into "basic" arrays as follows.

Let $\omega_-(R)$ and $\omega_+(R)$ be, respectively, the bottom and top endpoints of R, i.e.,

$$R = [\omega_-(R), \omega_+(R)]. \tag{24}$$

One defines inductively a collection $\{\omega_\ell^R : 1 \leq \ell \leq \ell^R\}$ beginning with $\omega_1^R := \omega_-(R)$. If ω_ℓ^R has been determined, we Choose $\omega_{\ell+1}^R$ as an element from $(\omega_\ell^R, \omega_+(R)]$ satisfying the conditions

$$\mathcal{I}([\omega_\ell^R, \omega_{\ell+1}^R]) \geq N^{-1} \text{ and } \mathcal{I}([\omega_\ell^R, \omega_{\ell+1}^R)) < N^{-1},$$

and then set

$$B_\ell^R := [\omega_\ell^R, \omega_{\ell+1}^R). \tag{25}$$

This element may not exist in the next cases:
(a) $\quad \omega_\ell^R = \omega_+(R)$ or $\mathcal{I}([\omega_\ell^R, \omega_+(R)]) < N^{-1}$.
We define $\omega_{\ell+1}^R$ as the father of $\omega_+(R)$ and set

$$B_\ell^R := [\omega_\ell^R, \omega_{\ell+1}^R) (= [\omega_\ell^R, \omega_+(R)]).$$

(b) $\quad \omega_\ell^R \neq \omega_+(R)$ and $\mathcal{I}(\{\omega_\ell^R\}) \geq N^{-1}$.
We define $\omega_{\ell+1}^R$ as the father of ω_ℓ^R and introduce B_ℓ^R by(25).
In this case $\omega_{\ell+1}^R \in R$, and the procedure can be continued. Note also that now B_ℓ^R consists of a single point, $B_\ell^R = \{\omega_\ell^R\}$.
Completing the procedure one obtains the partition $\{B_\ell^R : 1 \leq \ell \leq \ell^R\}$ of R into the basic arrays B_ℓ^R. By their definition

$$\mathcal{I}\left(B_\ell^R\{\omega_\ell^R\}\right) < N^{-1}. \quad (26)$$

Note that the argument in(26) is an empty set, if B_ℓ^R is a singleton. Besides, for $\ell < \ell^R$and card$\left(B_\ell^R\right) > 1$

$$\mathcal{I}\left(\hat{B}_\ell^R\right) \geq N^{-1}, \quad (27)$$

provided that$\hat{B}_\ell^R := B_\ell^R \cup \left\{\omega_{\ell+1}^R\right\}$,if B_ℓ^R is not a singleton, and $\hat{B}_\ell^R = B_\ell^R$, otherwise.
Collecting all the basic arrays for all $R \in \mathcal{R}_N$, we lastly obtain the desired set of the basic arrays

$$\mathcal{B}_N := \left\{B_\ell^R : 1 \leq \ell \leq \ell^R, R \in \mathcal{R}_N\right\}.$$

Proposition (2.1.2)[20]: (a) $\mathcal{B}_N$ is a partition of the set$\mathcal{G}_N \backslash \{\Omega\}$.
(b) For each $B := [\omega_-(B), \omega_+(B)]$from$\mathcal{B}_N$

$$\mathcal{I}\left((\omega_-(B), \omega_+(B)]\right) < N^{-1}. \quad (28)$$

(c) It is true that

$$\text{card}\, \mathcal{B}_N \leq 4N + 1. \quad (29)$$

Proof. (a) follows from (23) and the definition of B_R^ℓ.
(b) follows from (26), since the argument in(28) is $B \backslash \{\omega_-(B)\}$.
(c) Using(27) and noting that the mutiplicity of the cover of R by $\left\{\hat{B}_\ell^R\right\}$ is at most 2, one has

$$N^{-1}(\ell_R - 1) \leq \sum_{\ell=1}^{\ell_R - 1} \mathcal{I}\left(\hat{B}_\ell^R\right) < 2\mathcal{I}(R).$$

This implies, see(13),

$$\sum_{R \in \mathcal{R}_N} (\ell_R - 1) < 2N \sum_{R \in \mathcal{R}_N} \mathcal{I}(R) \leq 2N\mathcal{I}(\mathcal{G}_N) \leq 2N,$$

whence

$$\text{card}(\mathcal{B}_N) = \sum_{R\mathcal{R}_N} \ell_R < 2N + \text{card}(\mathcal{R}_N).$$

By the definition of$\mathcal{R}_N$

$$\text{card}(\mathcal{R}_N) \leq \text{card}(C_N) + \text{card}(M_N) \leq 2N + 1.$$

see(17) and(22).
Combining the last estimates we get(29).
The output of the algorithm is a functionF_N on$\mathcal{T}$ defined as follows.
If $\omega := \omega_-(B)$, the bottom endpoint of a basic array$B \in \mathcal{B}_N$,then

$$F_N(\omega) := \left(\sum_{\omega' \in B} F(\omega')\right) \chi_\omega. \quad (30)$$

We also let $F_N(\Omega) := G(\Omega)\chi_\Omega$.
For all other $\omega \in \mathcal{T}$we let

$$F_N(\omega) := 0. \quad (31)$$

Hence$F_N(\omega)(x)$ is a polynomial from X, if $x \in \omega$, and

$$\text{supp}\, F_N \subset \{\omega_-(B) : B \in \mathcal{B}_N\} \cup \{\Omega\}. \quad (32)$$

Let $\mathcal{T}$ and X be defined as above. We introduce, first, a subspace of$L_p(\Omega)$,$0 < p < \infty$, consisting of functions f that can be presented in a form

$$f = \sum_{\omega \in \mathcal{T}} f_\omega \chi_\omega \left(\text{convergence in } L_p\right) \quad (33)$$

with suitable $f_\omega \in X$.
Then we define the space $B_p^w(\mathcal{T})$by finiteness of the Banach norm (quasinorm, if $p < 1$)

$$|f|_{B_p^w(\mathcal{T})} := \inf\left\{\sum_{\omega\in\mathcal{T}} \left(w(\omega) \sup_\omega|f_\omega|\right)^p\right\}^{\frac{1}{p}}, \tag{34}$$

where the infimum is taken over all decompositions (33).
Here$w: \mathcal{T} \to \mathbb{R}_+$is a given weight Assume now that for some $p < q \le \infty$ the following embedding"

$$B_p^w(\mathcal{T}) \subset L_q(d\mu) \tag{35}$$

holds with embedding constant C_{em}. Here μ is a Borel measure supported by Ω.
Under this assumption the following is true.
Theorem (2.1.3)[20]: Given $f \in B_p^w(\mathcal{T})$ and integer $N \ge 1$, there is an N-term linear combination

$$T_N(f) := \sum_\omega f_\omega \chi_\omega$$

with suitable$f_\omega \in X$and$\omega \in \mathcal{T}$such that

$$\|f - T_N(f)\|_{L_q(d\mu)} \le CN^{\frac{1}{q}-\frac{1}{p}}|f|_{B_p^w(\mathcal{T})}. \tag{36}$$

Besides,

$$\|T_N(f)\|_{L_q(d\mu)} \le C|f|_{B_p^w(\mathcal{T})}. \tag{37}$$

Here the constant C depends only on C_{em}, $C(\mathcal{T})$, see(7), and$p^* := \min(1, p)$.
Proof. Assume that(33) is an ε-optimal decomposition for f, i.e.,

$$\left(\sum_{\omega\in\mathcal{T}} \left(w(\omega) \sup_\omega|f_\omega|\right)^p\right)^{\frac{1}{p}} \le (1+\varepsilon)|f|_{B_p^w(\mathcal{T})}. \tag{38}$$

Without loss of generality we assume that

$$\sum_{\omega\in\mathcal{T}} \left(w(\omega) \sup_\omega|f_\omega|\right)^p = 1. \tag{39}$$

Define now a function$F: \mathcal{T} \to X$ by letting

$$F(\omega) := f_\omega, \omega \in \mathcal{T}. \tag{40}$$

By(39), this F satisfies(9) and we take it and an integer $N \ge 1$ as the input of the algorithm. As the output we obtain the function F_N, see(30) and(31) for $F(\omega) := f_\omega$. In turn, F_N gives rise to required approximation aggregate

$$T_{4N+1}(f) := f_{\Omega\chi_\Omega} + \sum_{B\in\mathcal{B}_N} \left(\sum_{\omega\in B} f_\omega\right)\chi_{\omega_-(B)}. \tag{41}$$

Let us show that(41) provides the desired rate of approximation to the function f in $L_q(d\mu)$. Set

$$\phi(S) := \sum_{\omega\in S} f_\omega\chi_\omega, f \subset \mathcal{T}, \tag{42}$$

and simplify this notation for $S: \mathcal{T}(\omega)$, see(11), by setting

$$\phi(\omega) := \phi(\mathcal{T}(\omega))\omega \in \mathcal{T}. \tag{43}$$

Note that $\phi(\omega) \ne \phi(\{\omega\}) := f_\omega\chi_\omega$.
Proposition(2.1.2) and(42) imply

$$f\_T_{4N+1}(f) = \sum_{B\in\mathcal{B}_N} \phi^*(B) + \phi(\mathcal{T}\backslash\mathcal{G}_N), \tag{44}$$

where we let

$$\phi^*(B) := \phi(B) - \left(\sum_{\omega\in B} f_\omega\right)\chi_{\omega_-(B)} = \sum_{\omega\in B} f_\omega \chi_{\omega\backslash\omega_-(B)}. \tag{45}$$

Applying to(44) the $L_q(d\mu)$-norm, we get for $C := \max\left(1, 2^{\frac{1}{q}-1}\right)$

$$\|f\_T_{4N+1}(f)\|_q \le C(J_1 + J_2), \tag{46}$$

Where

$$J_1 := \left\|\sum_{B\in\mathcal{B}_N} \phi^*(B)\right\|_q, J_2 := \|\phi(\mathcal{T}\backslash G_N)\|_q. \tag{47}$$

In order to obtain the required estimate for J_1, show that for different B, B' from $\mathcal{B}_N$

$$|\text{supp}\, \phi^*(B) \cap \text{supp}\, \phi^*(B')| = 0. \tag{48}$$

Let, first, their top endpoints $\omega_+(B)$ and $\omega_+(B')$ nonoverlap. Since by(45)

$$\text{supp}\, \phi^*(B) \subset \omega_+(B)\backslash\omega_-(B) \tag{49}$$

and the similar is true forsupp $\phi^*(B')$, these supports nonoverlap.

In the remaining case the biggest set of one of them, say $\omega_+(B)$, embeds in the smallest set of the other $\omega_-(B')$. Hence $\text{supp}\, \phi^*(B) \subset \omega_+(B) \subset \omega_-(B')$, while by(49) $\text{supp}\, \phi^*(B') \subset \omega_+(B')\backslash\omega_-(B')$. Thus in this case(48) holds, as well.

Applying (48), we get

$$J_1 = \left\{\sum_{B\in\mathcal{B}_N} \|\phi^*(B)\|_q^q\right\}^{\frac{1}{q}}.$$

Using now embedding (35) and remembering the definition of the cost function I, see (10), we have

$$\|\phi^*(B)\|_q \le \left\|\sum_{\omega\in B\backslash\{\omega-(B)\}} |f_\omega|\chi_\omega\right\|_q \le C_{em}\left\{\sum_{\omega\in B\backslash\{\omega_-(B)\}} \left(w(\omega)\sup_\omega|f_\omega|\right)^p\right\}^{\frac{1}{p}}$$
$$= C_{em}\mathcal{I}(B\backslash\{\omega_-(B)\}).$$

Combining this and (28) and (29), we have

$$J_1 \le C_{em}\left\{\sum_{B\in\mathcal{B}_N} \mathcal{I}(B\backslash\{\omega_-(B)\})^{-\frac{q}{p}}\right\}^{\frac{1}{q}} \le C_{em}N^{-\frac{1}{p}}\,\text{card}\,(\mathcal{B}_N)^{\frac{1}{q}} \le 4^{\frac{1}{q}}C_{em}N^{\frac{1}{q}-\frac{1}{p}}$$

According to (38) and (39) this can be rewritten as

$$J_1 \le 4^{\frac{1}{q}}(1+\varepsilon)^{-1}C_{em}N^{\frac{1}{q}-\frac{1}{p}}. \tag{50}$$

To carry out the similar estimate for J_2 we introduce a collection $\{H_j\}$ of subsets of the set

$$T_0 := \mathcal{T}/\mathcal{G}_N \tag{51}$$

which meets the following conditions.

(a) For every j

$$\mathcal{I}(H_j) < \frac{C(\mathcal{T})}{N}. \tag{52}$$

(b) It is true that

$$\text{card}\,(\{H_j\}) \leq N + 1. \tag{53}$$

An algorithm of nonlinear approximation by piecewise polynomials

(c) $\{H_j\}$is a partition of $T_0(:= \mathcal{T}/G_N)$.

We introduce the required collection by induction. In this part of proof we use the following notation: for every $T \subset \mathcal{T}$ and $\omega \in \mathcal{T}$

$$T(\omega) := \{\omega' \in T : \omega' \subset \omega\}.$$

We begin with the set

$$\{\omega \in \mathcal{T} : \mathcal{I}(T_0(\omega)) \geq N^{-1}\}$$

Since $\mathcal{I}(T_0(\omega)) \leq \mathcal{I}(\omega) \to 0$ as $|\omega| \to 0$, see (12) and (13), this set is either empty or finite. In the former case we obtain the desired (trivial) partition putting $H_1 := T_0$. Then $\mathcal{I}(H_1) = \mathcal{I}(T_0(\Omega)) < N^{-1}$, and (52) is true. Otherwise, T_0 contains an element ω_1 of minimal measure. Since for each $\omega \in T_0$

$$\mathcal{I}(T_0(\omega)) \leq \mathcal{I}(\omega) < N^{-1},$$

this $\omega_1 \notin T_0$. Hence we have the disjoint decomposition of $T_0(\omega_1)$:

$$T_0(\omega_1) = \bigcup_{\omega \in S(\omega_1)} T_0(\omega)\,;$$

recall that $S(\omega_1)$is the set of the sons of ω_1, see (6). Besides, minimality of ω_1, implies for each $\omega \in S(\omega_1)$,

$$\mathcal{I}(T_0(\omega)) < N^{-1}$$

Hence it is true that

$$\mathcal{I}(T_0(\omega_1)) = \sum_{\omega \in S(\omega_1)} \mathcal{I}(T_0(\omega)) < \frac{\text{card}\,(S(\omega_1))}{N} \leq \frac{C(\mathcal{T})}{N}$$

see (7). Introduce now H_1 by

$$H_1 := T_0(\omega_1).$$

Then H_1 satisfies (52). To introduce the next set we put $T_1 := T_0 \backslash H_1$ and consider the set

$$\{\omega \in \mathcal{T} : T_1(\omega) \geq N^{-1}\}.$$

If it is empty, put $H_2 := T_1$ to obtain the desired partition $\{H_1, H_2\}$ of T_0.

Otherwise, this set contains an element ω_2 of minimal measure. As before $\omega_2 \notin T_0(:= \mathcal{T}/G_N)$. and therefore

$$\mathcal{I}(T_1(\omega_2)) < \frac{C(\mathcal{T})}{N}$$

Letting $H_2 := T_1(\omega_2)$ we obtain the desired subset satisfying (52) and not intersecting H_1. Besides,

$$\mathcal{I}(H_i) := \mathcal{I}(T_{i-1}(\omega_i)) \geq N^{-1}, \qquad i = 1,2.$$

Proceeding in this way, we lastly obtain the partition $\{H_j : 1 \leq j \leq n+1\}$ of T_0 satisfying the condition (52). Besides,$H_i := T_{i-1}(\omega_i)$, $1 \leq i \leq n$, and therefore

$$\mathcal{I}(H_i) \geq \frac{1}{N}, \qquad 1 \leq i \leq n.$$

This implies

$$\frac{n}{N} \leq \sum_{i=1}^{n} \mathcal{I}(H_i) \leq \mathcal{I}(F_0) \leq \mathcal{I}(\Omega) = 1,$$

and the condition (53) holds as well.

Using now the partition introduced, we estimate J_2 as follows. By the definition of H_j their supports do not overlap:

$$\left|(\operatorname{supp} H_j) \cap (\operatorname{supp} H_{j'})\right| = 0, \qquad j \neq j'$$

Recall that $\operatorname{supp} H := \bigcup_{\omega \in H} \omega, H \subset \mathcal{T}$. Besides, $\operatorname{supp} H_j = \operatorname{supp} \phi(H_j)$see (42). Hence

$$J_2 := \|\phi(\mathcal{T} \backslash \mathcal{G}_N)\|_q = \left\| \sum_{j \leq n+1} \phi(H_j) \right\|_q = \left\{ \sum_{j \leq n+1} \|\phi(H_j)\|_q^q \right\}^{\frac{1}{q}}.$$

By the embedding (35) and the inequality (52), and the definitions (10) and (34) of, respectively, $\mathcal{I}$ and the quasinorm of $B_p^w(\mathcal{T})$ we then have

$$\|\phi(H_j)\|_q \leq C_{em} \mathcal{I}(H_j)^{\frac{1}{p}} < C_{em} C(\mathcal{T})^{\frac{1}{p}} N^{-\frac{1}{p}}.$$

Together with the previous identity and (53) this yields

$$J_2 \leq C_{em} C(\mathcal{T})^{\frac{1}{p}} N^{-\frac{1}{p}} (n+1)^{\frac{1}{q}} \leq 2^{\frac{1}{q}} C_{em} C(\mathcal{T})^{\frac{1}{p}} N^{\frac{1}{q}-\frac{1}{p}}.$$

Combining this with (38), (49) and (46), we get the inequality

$$\|f - T_{4N+1}(f)\|_q \leq C N^{\frac{1}{p}-\frac{1}{q}} |f|_{B_p^w(\mathcal{T})}$$

This clearly implies the required assertion (36)

It remains to establish the second assertion of the theorem, see (37) By(41) and Proposition(2.1.2)

$$\|T_{4N+1}(f)\|_q = \left\| f_{\Omega \chi_\Omega} + \sum_{B \in \mathcal{B}_N} \left(\sum_{\omega \in B} f_\omega \right) \chi_{\omega_-(B)} \right\|_q \leq \left\| \sum_{\omega \in \mathcal{G}_N} |f_\omega| \, \chi_\omega \right\|_q.$$

Estimating the right hand side by the embedding inequality in (35) and then making use of the inequality (38) we have

$$\|T_{4N+1}(f)\|_q \leq C_{em} \left\{ \sum_{\omega \in \mathcal{G}_N} \left(w(\omega) \sup_\omega |f_\omega| \right)^p \right\}^{\frac{1}{p}} \leq C_{em} (1+\varepsilon)^{-1} |f|_{B_p^w(\mathcal{T})}.$$

The proof of Theorem(2.1.3) is completed.

Section (2.2): Refinable Functions

Approximation by nonlinear finite parametric manifolds has turned out to be very important in several areas of analysis. For example, in approximation by rational functions and by splines with free knots in approximation theory; asymptotics of eigenvalues in operator theory; K-divisibility and related topics in interpolation space theory; data compression in signal and image processing; and finite element methods in numerical analysis. This field has become more and more unified, and nowadays the phrase "nonlinear approximation" applies to a quickly developing theory with its own notions and methods. See [30], [34], [35], [27], [38], [47], [50], which cover different aspects of this theory and its applications.

The present considers a problem of N-term nonlinear approximation that dates back to the classical [49] by E. Schmidt published in 1887. The subsequent development of this part of the theory was essentially influenced by [33] of M. Birman and M. Solomyak. The problem under consideration can be presented, in general, in the following way. Given a complicated function f (an image, a solution of an ODE or PDE, etc.) and a library $\mathcal{L}$ of simpler functions, one tries to approximate f by an N-term linear combination of functions in $\mathcal{L}$ with (nearly) optimal degree of approximation. All these functions are elements of some normed space X, and approximation is measured by the norm of X. Usually, the choice of the library is dictated by the context of the original problem. (For instance, when we are working with finite element methods, $\mathcal{L}$ consists of piecewise polynomials. Alternatively, for numerical harmonic analysis, we can use a library of wavelets, and so on.) This means that, in general, the functions in $\mathcal{L}$ are not well fitted to singularities of the target function f, which may prevent us from effectively using linear methods to resolve the approximation problem. It may happen that X is contained in a larger space Y whose topologyor metric is insensitive to the singularities of f. This may enable effective linear approximation of f to be achieved in Y, and, imply that there exists an infinite series composed of scalar multiples of elements of the library $\mathcal{L}$ and such that it converges fairly rapidly to f in$Y$. If this happens, then we use the terms of that series to find an N-term linear combination of elements of $\mathcal{L}$ that is well adapted to f and in fact provides approximation of f in X that is comparable with the approximation of f in Y. We can be achieved by the use of the classical "greedy" algorithm (choose the N terms in the series whose coefficients have the largest absolute values). This simple method is miraculously successful whenever it can invoke the assistance of a powerful tool, the Caldero´n–Zygmund theory. However this assistance is not available when we wish to work in a number of function spaces important for applications (L_1 and L_∞ spaces, Hölder spaces, etc.). We will apply a different algorithm, which allows us to achieve the desired result for a large class of function spaces. This approach was developed in an algorithmic form in collaboration with Inna Kozlov (see, [20]), by using ideas suggested in [28] by Irina Irodova. Here we consider the application of this algorithm to the case of the library $\mathcal{L}_\varphi := \{\varphi_{jk}: j \in \mathbb{Z}, k \in \mathbb{Z}^n\}$ of "matrix dilated" and translated copies of a bounded refinable function $\varphi: \mathbb{R}^n \to \mathbb{R}$. Specifically, we have $\varphi_{jk}(x): \varphi(A^j x - k)$for some matrix A. The function φ is required to satisfy a scaling equation with respect to the matrix A and some finite mask m.

The function f is assumed to belong to a "bad" space $L_p(\mathbb{R}^n)$ with $0 < p < \infty$, but to have a "sparse" expansion

$$f = \sum c_{jk}\varphi_{jk} \tag{54}$$

that converges in L_p. Here "sparse" means that

$$|f|_{pq} := \left\{\sum_{j,k}\left(|\det A|^{\frac{-j}{q}}|c_{jk}|\right)^p\right\}^{\frac{1}{p}}$$

is finite for some $p < q \leq \infty$. Our algorithm processes the coefficients of the expansion(54) to produce an N-term linear combination f_N of functions φ_{jk} that provides the desired approximation in a "good" space$L_q(\mathbb{R}^n)$. In fact, we obtain

$$\|f - f_N\|_q \leq CN^{-\frac{s}{n}}|f|_{pq},$$

where s, p and q are related by

$$\frac{s}{n} = \frac{1}{p} - \frac{1}{q},$$

and $p \leq 1$ if $q = \infty$.

A simple modification of this algorithm enables us to obtain the same rate of approximation also when $\frac{s}{n} > \frac{1}{p} - \frac{1}{q}$. Finally, the algorithm is applied to the case where L_q is replaced by a Sobolev space (simultaneous approximationof f and its derivatives) and to vector-valued refinable functions (in particular, to piecewise polynomial approximation of Birman–Solomyak type). The refinable function φ appearing in these results is assumed to be stable and colorable. The former property is fulfilled, e.g., whenever the set$\{\varphi(x - k): k \in \mathbb{Z}^n\}$ forms a Riesz basis for the L_2-closure of its linear span. The latter property is fulfilled, e.g., whenever the dilation matrix A associated with φ diagonalizes over the field $\mathbb{Q}$, or whenever A is related to the mask m for φ by

$$|\det A| = \#\,\mathrm{supp}\, m.$$

For the first time, approximation of the type considered in the present was studied in the fundamental [39], $q < \infty$, and [40], $q = \infty$, by R. DeVore, V. Popov, and their collaborators. They dealt with a smooth compactly supported regular function φ (i.e., with the dilation $A: x \longmapsto 2x$) whose integer translates are locally linearly independent. Independently, and at about the same time, a similar result was presented in [28] for the special case of multivariate B-splines. The approximated functions belong to the Besov space$B_q^s(\mathbb{R}^n)$. The results of the present include the above two cases, along with many others (anisotropic Besov spaces, simultaneous approximation, wavelets, box splines and piecewise polynomial approximation, approximation by fractal functions, etc.). All these versions may be useful in applications to image processing, where each type of image singularity (edges, fractals, etc.) requires a flexible choice of the corresponding libraries.

We give a detailed description of the approximation algorithm used in these proofs.

The notation introduced here will be used throughout.

(A) Self-affine regions. Let A be an $(n \times n)$-matrix with integral entries (we write $A \in M_n(\mathbb{Z})$). Throughout A is assumed to be expanding, i.e., it has n eigenvalues with moduli larger than 1. Such a matrix will be called a dilation. Given a dilation A and a digit set $\mathcal{D} := \{d_1, \ldots, d_N\} \subset \mathbb{Z}^n$, we define a self-affine set $T = T(A, \mathcal{D})$ as a nonempty compact solution of the set-valued equation

$$A(T) = \bigcup_{d \in \mathcal{D}} (T + d)(= T + \mathcal{D}). \quad (55)$$

In accordance with Hutchinson's theorem [41], there is a unique compact set satisfying (55). It can be found by iterations of a set-valued map $S := S(A, \mathcal{D})$ given by

$$S(\Omega) \coloneqq \bigcup_{d \in \mathcal{D}} A^{-1}(\Omega + d),\ \Omega \subset \mathbb{R}^n. \quad (56)$$

In fact, for an arbitrary nonempty bounded set Ω we have

$$T(A, \mathcal{D}) = \lim_{j \to \infty} S^j(\Omega), \quad (57)$$

with convergence in the Hausdorff metric. This immediately yields the radix representation of the self-affine set:

$$T(A, \mathcal{D}) = \left\{ \sum_{j \in \mathbb{N}} A^{-j} d_j : d_j \in \mathcal{D} \right\}. \quad (58)$$

A straightforward consequence of (55) and (58) is formulated below.

We only deal with self-affine sets of positive Lebesque measure. They will be called self-affine regions for the reason explained by the next important result (see [51] and [42]).

If the set $T := T(A, \mathcal{D})$ is of positive measure, then T is the closure of its interior $T^0, T = \bar{T}^0$, and its boundary $\partial T := T \setminus T^0$ has Lebesque measure zero.

The following examples clarify and motivate the basic definition.

Example (2.2.1)[24]: (Tiles). A self-affine region $T := T(A, \mathcal{D})$is called a tile if translates $T + d$ with distinct $d \in \mathcal{D}$ are essentially disjoint. This means that$|(T + d) \cap (T + d')|$is zero if $d \neq d'$. Tiles arise in many contexts of analysis including subdivision schemes, multivariate wavelet systems, non-Fourier harmonic analysis, and Markov partitions (see [46], [52] and [51]). In the case of a tile, relation(55) implies that

$$\#\mathcal{D} = |\det A|. \quad (59)$$

In its turn, this implies that$|T|$is an integer if T is a region (see [46]). Condition (59) is not sufficient for the positivity of$|T|$. The simplest sufficient condition for this requires that$\mathcal{D}$ be a complete residue system for the factor group$\mathbb{Z}^n / A(\mathbb{Z}^n)$; see [25]. For each tile T there is a translation set $\mathcal{K} \subset \mathbb{Z}^n$ such that the family $\{T + k : k \in \mathcal{K}\}$is essentially disjoint and its union is $\mathbb{R}^n$ (in other words, T tiles $\mathbb{R}^n$). If $|T| = 1$, then the translation set is $\mathbb{Z}^n$ (see [46]).

In most cases, the boundaries of tiles are fractals, i.e., their Hausdorff dimension $\dim_H$ is strictly larger than the topological one. A remarkable example is the so-called "twin dragon" associated with $A := \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ and $\mathcal{D} := \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}$ (see, e.g., [51], where $\dim_H T(T, \mathcal{D}) \approx 1.523$.

Example (2.2.2)[24]: Let $B \in M_n(\mathbb{Z})$, $|\dim B| = 1$, and let $M_i \geq 2$ and $N_i \geq 1$ be integers, $1 \leq i \leq n$. Then the parallelotope

$$\Pi := B\left(\prod_{i=1}^{n} [0, N_i]\right) \tag{60}$$

with vertices in $\mathbb{Z}^n$ is a self-affine region associated with

$$A := B \operatorname{diag}(M_1, \dots, M_n)\, B^{-1} \quad \text{and } \mathcal{D}; = B\left(\prod_{i=1}^{n} J_i\right) \cap \mathbb{Z}^n,$$

where $J_i := [0, (M_i - 1)N_i], 1 \leq i \leq n$.

It is easy to check that Π is a tile only for $N_i = 1$ and $M_i = 2, 1 \leq i \leq n$. In this case Π is the image of the unit cube $[0,1]^n$ under the action of B, the set $\mathcal{D}$ is the set of vertices of this cube, and $A := 2I := \operatorname{diag}(2, \dots, 2)$. Since $|\Pi| = 1$, the translation set is $\mathbb{Z}^n$.

(B) The digraphs $Gr(A, \mathcal{D})$. Any self-affine set $T(A, \mathcal{D})$ gives rise to a digraph $Gr(A, \mathcal{D})$ in the following way. We introduce a sequence of subsets of $\mathbb{Z}^n$ given by

$$\mathcal{D}_0 := \mathcal{D}, \mathcal{D}_j := \left\{ \sum_{i=0}^{j-1} A^i d_i : d_i \in \mathcal{D} \right\}, j \in \mathbb{N}. \tag{61}$$

Then the set of vertices of $Gr := Gr(A, \mathcal{D})$ is given by

$$\mathcal{V} := \mathcal{V}(A, \mathcal{D}) := \{T_{jk} : j \in \mathbb{Z}_+, k \in \mathcal{D}_j\}, \tag{62}$$

Where

$$T_{jk} := A^{-j}(T + k), j \in \mathbb{Z}, k \in \mathbb{Z}^n. \tag{63}$$

Note that $\mathcal{V}$ is a $\mathbb{Z}_+$-graded set, the graduation of which is given by height h, i.e.,

$$h(T_{jk}) := j. \tag{64}$$

In its turn, this yields a partition of $\mathcal{V}$ into the subsets

$$\mathcal{V}_j := \{T_{jk} : k \in \mathcal{D}_j\}, j \in \mathbb{Z}_+. \tag{65}$$

Hence, for $T' \in \mathcal{V}_j$ we obtain

$$T' = A^{-j}\left(T + \sum_{i=0}^{j-1} A^i d_i\right) \quad (66)$$

with suitable digits $d_i \in \mathcal{D}$. Applying(55), we get

$$T' = \bigcup_{d\in\mathcal{D}} A^{-j-1}\left(T + d + \sum_{i=1}^{j} A^i d_{i-1}\right). \quad (67)$$

The subsets occurring in this union will be called the children of T'(and $T'$is their parent), and will be denoted by $ch(T')$. Observe that a child may have more than one parent.

Now, let $T', T'' \in \mathcal{V}$. These vertices determine an edge directed form $T'$to T''if $T'$is a child of T''. This edge will be denoted by $T' \to T''$, and the set of these edges by $\mathcal{E} := \mathcal{E}(A, \mathcal{D})$.

Thus, we have introduced the required digraph (directed graph) $Gr := Gr(A, \mathcal{D})$. In what follows we use the standard terminology of graph theory (see, e.g., [48]). In particular, a directed edge is named an arc, and its endpoints T' and $T''$are the tail and head, respectively. In accordance with its definition, the digraph $Gr(A, \mathcal{D})$ has no loops (an edge joining a vertex to itself) and no pairs of arcs with the same tail and head. Such a digraph is called to be strict (or simplicial).

A sequence $P := \{T_1, \ldots, T_m\} \subset \mathcal{V}$ is a path (or trail) if no vertex occurs in P more than once, and adjacent vertices are joined by an arc. If, moreover, $T_1 \to T_{m+1}$, $1 \le i < m$, this P is called a directed path, and consequently $T_1$and T_m are its tail and head. In this case we use the notation

$$T_p^- := T_1(\text{tail}), T_p^+ := T_m(\text{head}). \quad (68)$$

Vertices $T'$and $T''$are connected by a (directed) path $P := \{T_1, \ldots, T_m\}$ if $T' = T_1$ and $T'' = T_m$ (consequently, if $T' = T_P^-$and $T'' = T_P^+$). If $T' = T_P^-$ and $T' = T''$for a suitable directed path P, then $T'$is called an off spring of $T''$and $T''$is its ancestor. The following result, the proof of which is straightforward, collects the basic properties of the object introduced.

Proposition (2.2.3)[24]: (a) The degree of each vertex^1 of $Gr(A, \mathcal{D}) = (\mathcal{E}, \mathcal{V})$ equals$\#\mathcal{D}$.

(b) A vertex$T' \in \mathcal{V}$, regarded as a set, is the union of all its off springs of the same height. In particular,

$$T' = \bigcup_{T''\in ch(T')} T''. \quad (69)$$

(c) Each$T'$is connected with the set $T := T(A, \mathcal{D})$.

Since the vertices of $Gr(A, \mathcal{D})$ are subsets of $\mathbb{R}^n$, the set inclusion order gives rise to another digraph structure on $\mathcal{V}$. In this case T',T''in $\mathcal{V}$ are connected by an edge directed from T' to T''if $T' \subset T''$and there is no other vertex situated in-between. We denote this digraph by $Gr_0 = (\mathcal{V}_0, \mathcal{E}_0)$. Then $\mathcal{V}_0 = \mathcal{V}$, but, in general, the set of edges $\mathcal{E}$ is a proper subset of $\mathcal{E}_0$. Compatibility of the digraph structure of $Gr(A, T)$ with the set inclusion order is crucial for our approach. Below we introduce a class of self-affine regions for which this property is fulfilled in a sense. For this, we recall the notion of a coloring of a graph. This is a function defined on the set of vertices and with values in a finite set. The elements of this set are regarded as "colors".

Definition (2.2.4)[24]: A graph whose vertices are measurable subsets of $\mathbb{R}^n$ is spatially colorable if there is a coloring of this graph satisfying the following condition. Any two vertices v, w of the same color are either essentially disjoint$(|v \cap w| = 0)$, or$v \subset w$, or $w \subset v$.

The minimum of colors required in this definition is called the (spatial) chromatic number of this graph. For the digraph $Gr(A, \mathcal{D})$ this number is denoted by $\chi_{(A,\mathcal{D})}$. If $Gr_0(A, \mathcal{D})$ is spatially colorable, then, clearly, so is $Gr(A, \mathcal{D})$. Moreover, in this case almost each point of $\mathbb{R}^n$ is contained in at most $\chi(A, \mathcal{D})$subsets in $\mathcal{V}_j$, because these are colored by $\chi(A, \mathcal{D})$colors, and the distinct subsets of the same color and height are essentially disjoint. In other words, the multiplicity

$$\mu(\mathcal{V}_j) := \underset{x \in \mathbb{R}^n}{\text{ess sup}} \left(\sum_{T' \in \mathcal{V}_j} 1_{T'}(x) \right)$$

does not exceed $\chi(A, \mathcal{D})$. We conjecture that the converse is also true, i.e.,$\sup_j \mu(\mathcal{V}_j) < \infty$ implies $\chi(A, \mathcal{D}) < \infty$. It is easily seen that this supremum is finite. Consequently, we conjecture that each digraph $Gr(A, D)$ is spatially colorable.

More examples with an effective upper bound for the chromatic number will be discussed in detail.

Example (2.2.5)[24]: If $T(A, \mathcal{D})$ is a tile, then

$$\chi(A, \mathcal{D}) = 1. \tag{70}$$

In fact, in this case $Gr(A, \mathcal{D}) = Gr_0(A, \mathcal{D})$ is a rooted tree with the root $T(A, \mathcal{D})$.

Example (2.2.6)[24]: Assume that $Gr(A, \mathcal{D})$ has the following property: if the heights of two vertices $T', T' \in \mathcal{V}$ differ by one, and$|T' \cap T''| \neq 0$, then the smaller vertex is a subset of the bigger.

In this case $Gr(A, \mathcal{D}) = Gr_0(A, \mathcal{D})$ and $\chi(A, \mathcal{D}) < \infty$. Since the intersection of subsetsin $\mathcal{V}_j$ is a union of subsets in $\mathcal{V}_{j+1}, j \in \mathbb{Z}_+$, it is natural to name such a self-affine region $T(A, \mathcal{D})$ asemitile.

Example (2.2.7)[24]: The self-affine region of Example (2.2.2) is spatially colorable if the greatest common divisors (M_i, N_i) of M_i, N_i are 1 (see Proposition(2.2.32)).

(C) Refinable functions. A refinable function $\varphi: \mathbb{R}^n \to \mathbb{R}$ associated with a dilation A and mask $m: \mathbb{Z}^n \to \mathbb{R}$is a solution of the scaling equation

$$\varphi(x) = \sum_{k \in \mathbb{Z}^n} m(k) \varphi(Ax - k), x \in \mathbb{R}^d. \tag{71}$$

A rather complete account of properties of regular refinable functions, i.e., such that $A := 2I := \mathrm{diag}(2, \ldots, 2)$, was presented in [36]. Some of these properties can be established in the general case by the same arguments. In particular, this concerns the properties listed.

Throughout, the mask m is assumed to be finite, i.e.,

$$\# \operatorname{supp} m \coloneqq \#\{k \in \mathbb{Z}^n : m(k) \neq 0\} < \infty. \tag{72}$$

This implies immediately that φ is compactly supported.

We also assume that φ is a bounded and nontrivial solution of(71), i.e.,

$$\|\varphi\|_\infty \coloneqq \operatorname*{ess\,sup}_{\mathbb{R}^n} |\varphi| < \infty \text{ and} |\operatorname{supp} \varphi| \neq 0, \tag{73}$$

where $\operatorname{supp} \varphi := \{x \in \mathbb{R}^n : \varphi(x) \neq 0\}$.

We introduce a library $\mathcal{L}_\varphi$ by

$$\mathcal{L}_\varphi \coloneqq \{\varphi_{jk}(x) \coloneqq \varphi(A^j x - k) : j \in \mathbb{Z}, k \in \mathbb{Z}^n\}. \tag{74}$$

Since $\operatorname{supp} \varphi_{jk} = A^j(\operatorname{supp} \varphi + k)$, we have

$$|\operatorname{supp} \varphi_{jk}| = |\det A|^j |\operatorname{supp} \varphi|. \tag{75}$$

Now, from(71) we derive a similar "jth level" scaling equation,$j \in \mathbb{N}$.For this, we extend the mask m from $\mathbb{Z}^n$ (identified with$\{1\} \times \mathbb{Z}^n$) to$\mathbb{Z} \times \mathbb{Z}^n$by setting

$$m^*(j, k) = 0 \ \text{ if} j \leq 0, k \in \mathbb{Z}^n,$$

and defining $m^*(j, k)$ for $j > 1$ by the chain rule:

$$m^*(j, k) := \sum_{k = Ak' + k''} m(k'') m^*(j - 1, k). \tag{76}$$

This extension is well defined because the mask is finite.

Using this and applying the scaling equation (71) repeatedly, we obtain

$$\varphi = \sum_{k \in \mathbb{Z}^n} m^*(j, k) \varphi_{jk}, j \geq 1. \tag{77}$$

Equation (71) also implies the embedding

$$A(\operatorname{supp} \phi) \subset \operatorname{supp} \varphi + \operatorname{supp} m. \tag{78}$$

In the following cases, equality occurs here (and $\operatorname{supp} \varphi$ is a self-affine region associated with A and $\mathcal{D} \coloneqq \operatorname{supp} m$).

(a) The mask is nonnegative (hence, $\varphi \geq 0$ a.e.).

(b) The family of integer translates $\{\varphi(x - k) : k \in \mathbb{Z}^n\}$ is locally linearly independent, i.e., the nonzero restrictions of these translates to an arbitrary open cube are linearly independent.

In general, $\operatorname{supp} \varphi$ is not a self-affine region, but it is related to such a region in the following way.

Proposition (2.2.8)[24]: If $\mathcal{D} \supset \operatorname{supp} m$, then

$$\operatorname{supp} \varphi \subset T(A, \mathcal{D}). \tag{79}$$

Proof: Using the set-valued operation S with $\mathcal{D} \supset \operatorname{supp} m$, from(78) we deduce that

$$\operatorname{supp} \varphi \subset S(\operatorname{supp} \varphi).$$

Iterating and applying(57), we obtain

$$\operatorname{supp} \varphi \subset S^j(\operatorname{supp} \varphi) \to T(A, \operatorname{supp} m), j \to \infty .$$

Definition (2.2.9)[24]: A refinable functionφ with dilation A and mask m is said to be colorable if there is a digit set $\mathcal{D}$ such thatsupp $m \subset \mathcal{D}$ and

$$\chi(A, \mathcal{D}) < \infty.$$

We put

$$\chi(\varphi) := inf \chi(A, \mathcal{D}) : \operatorname{supp} m \subset \mathcal{D}. \tag{80}$$

Example (2.2.10)[24]: Let $T := T(A, \mathcal{D})$ be a tile, and let $\varphi := 1_T$ be the characteristic function of T. Then

$$\varphi(x) = \sum_{k \in \mathcal{D}} \varphi(Ax - k)$$

almost everywhere (see Example(2.2.1)). Since in this casesupp $m = \mathcal{D}$and $\chi(A, D) = 1$, such φ is a colorable refinable function with $\chi(\varphi) = 1$.

Example (2.2.11)[24]: Suppose the dilation A of φ is $\mathbb{Z}$-similar to a diagonal matrix whose eigenvalues are rational numbers.

Now we define yet another notion used.

Definition (2.2.12)[24]: A refinable function φ is p-stable,$0 < p \leq \infty$, if for each sequence $\lambda := \{\lambda(k) : k \in \mathbb{Z}^n\} \subset \mathbb{R}$ we have

$$C_1 \|\lambda\|_p \leq \left\| \sum \lambda(k) \varphi(x - k) \right\|_p \leq C_2 \|\lambda\|_p \tag{81}$$

with $C_1, C_2 > 0$ independent of λ.

Since in our case φ is bounded and compactly supported, the right inequality is trivially true for all p. It is well known (see Lemma(2.2.20)) that in this case the left inequality is true for all p provided it is valid for one. For this reason, we shall call such φ a stable refinable function (i.e., this term means that φ is bounded, compactly, supported, and p-stable).

Proposition (2.2.13)[24]: Assume thatφ is stable. Then its extended mask(76) satisfies

$$\sup_{j,k} |m^*(j, k)| < \infty. \tag{82}$$

Proof: Using the ∞-stability ofφ and(77), we get

$$|m^*(j,k)| \le C \left\| \sum_k m^*(j,k)\varphi(x-k) \right\|_\infty = C \left\| \sum_k m^*(j,k)\varphi(A^j x - k) \right\|_\infty = C\|\varphi\|_\infty$$

with C independent of j and k.

Remark (2.2.14)[24]: (a) A compactly supported $\varphi \in L_p(\mathbb{R}^n)$satisfies the p-stability condition in(81) if and only if for each$\xi \in (\mathbb{R}^n)^*$ there exists $k \in \mathbb{Z}^n$ such that

$$\hat{\varphi}(\xi + 2\pi k) \neq 0. \tag{83}$$

Here $\hat{\varphi}$ stands for the Fourier transform of φ; see [44].

(b) If the set of integer translates of $\varphi \in L_\infty(\mathbb{R}^n)$is locally linearly independent, then φ is stable; see [39].

We recall the notion of the Strang–Fix condition for a regular refinable function. A function φ satisfies this condition with respect to a finite-dimensional translation invariant subspace P of polynomials if for each $p \in P$ and suitable constants $\lambda(k)$ we have

$$p(x) = \sum_{k\in\mathbb{Z}^n} \lambda(x)\varphi(x-k).$$

For φ compactly supported, this is equivalent to the condition that

$$\left(D^\ell \hat{\varphi}\right)(2\pi k) = 0, k \in \mathbb{Z}^n \backslash \{0\},$$

for all $\ell \in \mathbb{Z}_+^n$ such that $x^\ell \in P$ (see [36]). For the general case of φ associated with an arbitrary dilation and a finite mask, the corresponding condition was presented in $[J]$ (see also [37]).

(D) $B(\varphi)$-spaces. Using the library $\mathcal{L}_\varphi$ (see(74)), for $0 < p < \infty$ we introduce the linear space $\Sigma_p(\varphi)$, of measurable (classes of) functions on $\mathbb{R}^n$ represented as

$$f = \sum c_{jk}\varphi_{jk} \,(\text{convergence in } L_p). \tag{84}$$

Assuming that φ is compactly supported, for$f \in \Sigma_p(\varphi)$ we set

$$\|f\|_{B_p^s(\varphi)} := \inf\left\{ \sum_{j,k} \left(\left|\operatorname{supp}\varphi_{jk}\right|^{\frac{1}{q}} \left|c_{jk}\right| \right)^p \right\}^{\frac{1}{p}}, \tag{85}$$

where the infimum is taken over all expansions as in(84), and $s > 0, 0 < p < q \le \infty$ are related by

$$\frac{s}{n} = \frac{1}{p} - \frac{1}{q}. \tag{86}$$

It is readily seen that (85) yields a Banach (quasi)norm on the linear space $B_p^s(\varphi)$ of all $f \in \Sigma_p(\varphi)$ with finite(85).

More generally, we define the space $B_p^{s\theta}(\varphi)$by the quasinorm

$$\|f\|_{B_p^{s\theta}(\varphi)} := \inf\left\{\sum_{k\in\mathbb{Z}^n}\left(\sum_{j\in\mathbb{Z}}|c_{jk}|.\left(|\operatorname{supp}\varphi_{jk}|^{\mu}\right)^p\right)^{\frac{\theta}{p}}\right\}^{\frac{1}{\theta}}, \tag{87}$$

where the infimum is taken over all expansions(84). Here $0 < \theta\, , p \le \infty\, , s > 0$, and

$$\mu := \frac{s}{n} - \frac{1}{p}.$$

Clearly, this coincides with $B_p^s(\varphi)$ if $\theta = p$. We only deal with the latter space and with the space $B_p^{s\infty}(\varphi)$ defined by

$$\|f\|_{B_p^{s\infty}(\varphi)} \inf\left(\sup_{j,k}\left(|c_{jk}||\operatorname{supp}\varphi_{jk}|^{\mu}\right)\right).$$

These definitions and notation are motivated by the following result of [39]; a partial case of multivariate B-splines was proved independently in [28].

Theorem(2.2.15)[24]: Assume that φ is a bounded regular refinable function of finite mask and obeying the following conditions:

(a) The set $\{\varphi(x-k): k \in \mathbb{Z}^n\}$ is locally linearly independent.

(b) For some $r > s$, the function$\varphi$is subject to the Strang–Fix condition with respect to the space of polynomials of degree less than$r$.

Then, up to equivalence of (quasi)norms,

$$B_p^{s\theta}(\varphi) = B_p^{s\theta}(\mathbb{R}^n).$$

Using a "pseudonorm" associated with A (see [45]), one can define a generalized Besov space $B_p^{s,A}(\mathbb{R}^n)$and conjecture a similar result. Such a pseudonorm is a nonnegative function $\nu := \nu_A$on$\mathbb{R}^n$ satisfying the conditions

(a) $\nu(-x) = \nu(x)$, and $\nu(x) = 0$ if and only if $x = 0$;

(b) $\nu(Ax) = |\det A|^{\frac{1}{n}}\nu(x)$.

For instance, if $A := \operatorname{diag}(M_1, \dots, M_n), |M_i| > 1$, then

$$\nu(x) := \sum_{i=1}^{n} |x_i|^{a_i}, \tag{88}$$

Where

$$a_i := \frac{\log|\det A|}{n \log M_i}$$

In particular, ν is equivalent to the standard norm if A is isotropic, i.e., $\mathbb{Z}$-similar to a diagonal matrix with all eigenvalues of the same modulus.

The required Besov space is defined via its (quasi)norm

$$\|f\|_{B_p^{s,A}} := \left\{\|f\|_p + \int_{\mathbb{R}_+} \left(\frac{\omega_r^A(t; f; L_p)}{t^s}\right)^p \frac{dt}{t}\right\}^{\frac{1}{p}},$$

where $r > s$ and

$$\omega_r^A(t; f; L_p) := \sup_{\nu(x) \le t} \|\Delta_x^r f\|_p.$$

Since all pseudonorms associated with A are equivalent (see [45]), this space does not depend on the choice of ν (up to equivalence of (quasi)norms).

It can be shown that, if ν is as in(88) and $r > s \max a_i$, then the space $B_p^{s,A}(\mathbb{R}^n)$ coincides with the anisotropic Besov space $B_p^{s_1,\dots,s_n}(\mathbb{R}^n)$, where $s_i := sa_i$.This leads to the following conjecture. If φ is stable and r is sufficiently large, then

$$B_p^{s,A}(\mathbb{R}^n) = B_p^s(\varphi).$$

This conjecture can be extended to the case where $s \le 0$ and $p \ge 1$. Now $B_p^s(\varphi)$is a space of tempered distributions defined by formula(85) where the infimum is taken over all expansions(84) with convergence in the sense of distributions. The remaining space is defined via the norm

$$\|f\|_{B_p^{s,A}} := \left\{\sum_{k\in\mathbb{Z}} \left(a^{sj}\left\|\theta_j * f\right\|_p\right)^p\right\}^{\frac{1}{p}},$$

where $a := |\det A|^{\frac{1}{n}}$, $\hat{\theta}_j(\xi) := \theta(B^j, \xi)$,$\xi \in (\mathbb{R}^n)^*$, $B := A^T$, and θ is a nonnegative C_0^∞-function supported on $\{x \in \mathbb{R}^n : a^{-1} < \nu_B(x) < a\}$. Recall that s, p, q are related by(86), so that $0 < q \le p$ in this case.

For $A := 2I$, this definition gives the standard Besov space $B_p^s(\mathbb{R}^n), 1 \le p \le \infty$(see, e.g., [32]), while for $A := \operatorname{diag}(M_1, \dots, M_n)$ it determines the corresponding anisotropic Besov space $B_p^{s_1,\dots,s_n}(\mathbb{R}^n)$ with $s_i \in \mathbb{R}, 1 \le p \le \infty$. The problem presented by the conjecture above, along with other properties of the scale$\{B_p^s(\varphi)\}$, will be studied elsewhere.

φ is a nontrivial bounded refinable function with a given dilation A and a finite mask m. We recall that the extended mask$m^*$is defined by (76). The library$\mathcal{L}_\varphi := \{\varphi_{jk}; j \in \mathbb{Z}, k \in \mathbb{Z}^d\}$generated by this$\varphi$ can be graded as follows:

$$\mathcal{L}_\varphi := \bigcup_{N=1}^{\infty} \mathcal{L}_\varphi(N), \tag{89}$$

where $\mathcal{L}_\varphi(N)$ is the family of all N-term linear combinations of φ_{jk}.

Now, assume that

(a) φ is a stable and colorable refinable function (see Definitions(2.2.9) and(2.2.12)); (b) the numbers $0 < p < q \leq \infty$ and $s > 0$ are related by

$$\frac{s}{n} = \frac{1}{p} - \frac{1}{q}, \tag{90}$$

and $p \leq 1$ if $q = \infty$. Under these assumptions, the following is true.

To formulate a consequence of Theorem (2.2.27), we introduce the best approximation

$$\mathcal{E}_N(f; L_q) := \inf\{\|f - f_N\|_q f_N \in \mathcal{L}_\varphi(N)\}. \tag{91}$$

Suppose that assumption(a) of Theorem(2.2.27) is fulfilled, but assumption (90) is replaced by

$$\frac{s}{n} > \frac{1}{p} - \frac{1}{q} \tag{92}$$

with $0 < p < q \leq \infty$and $s > 0$.

Under these assumptions, the following is true (see [39], [40]).

Here, the crucial point is the so-called Bernstein's inequality, which was first introduced and named after S. Bernstein in [26] devoted to approximation by rational functions with free poles. This inequality must look like this:

$$\|f\|_{B_p^s(\varphi)} \leq CN^{-\frac{s}{n}}\|f\|_q \,, f \in \mathcal{L}_\varphi(N), \tag{93}$$

with C independent of f and N and s, p, q related by(90).

This inequality can also be established. We shall study this issue in a forthcoming.

Remark (2.2.16)[24]: Let A be diagonalizable with the eigenvalues $M_i > 1, 1 \leq i \leq n$. Assume that for some $\sigma > 0$ the numbers

$$\ell_i := \frac{\sigma \log|\det A|}{\log M_i}, 1 \leq i \leq n, \tag{94}$$

are integers. In this case

$$\sigma = \langle \bar{\ell} \rangle := \left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{M_i}\right)^{-1}$$

is the harmonic mean of $\bar{\ell} := (\ell_1, .., \ell_n)$. Assume that the numbers p, q, s satisfy the condition

$$\frac{s - \sigma}{n} = \frac{1}{p} - \frac{1}{q} > 0 \tag{95}$$

and that $p \leq 1$ if $q = \infty$. Then the assertions of Theorem(2.2.27) and Corollary(2.2.30) remain true if we replace the L_q-norm by the anisotropic Sobolev norm

$$\|f\|_{W_q^{\bar{\ell},A}} := \sum_{i=1}^{n} \left\|D_i^{\ell_i} f\right\|_q, \tag{96}$$

where D_i is the derivative in the direction determined by the ith eigenvector of A. Of course, we assume that, moreover, φ belongs to $W_q^{\bar{\ell},A}(\mathbb{R}^n)$.

In this case the method yields simultaneous approximation of f and its derivatives by$f_N$and the corresponding derivatives of f_N.

A similar version of Theorem(2.2.29) is also true under the condition

$$\frac{s-\sigma}{n} > \frac{1}{p} - \frac{1}{q} > 0. \tag{97}$$

In the isotropic case, i.e., for a diagonalized dilation with equal eigenvalues, all the results stated above are valid for the isotropic Sobolev space W_q^{ℓ}; in this case $\sigma = \ell$.

Simple changes in the proofs of the main results that lead to simultaneous approximation are discussed.

We present two results facilitating the proof of Theorem (2.2.27).

Let $Gr = (\mathcal{V}, \mathcal{E})$ be a digraph whose vertices are subsets of $\mathbb{R}^n$, and let $c: \mathcal{V} \to R$ be its coloring. For a color $\gamma \in \Gamma$, set

$$\mathcal{V}(\gamma) ;= \{v \in \mathcal{V} : c(v) = \gamma\}. \tag{98}$$

The elements of this set are named γ-vertices. The family $\{\mathcal{V}(\gamma) : \gamma \in \Gamma\}$ forms a partition of$\mathcal{V}$,

$$\mathcal{V} = \coprod_{\gamma \in \Gamma} \mathcal{V}(\gamma). \tag{99}$$

Here and below $\coprod$ stands for disjoint union.

A vertex $v \in \mathcal{V}(\gamma)$ is called a γ-root if$v$is not a subset of another γ-vertex. The collection of γ-roots is denoted by $\mathcal{R}(\gamma)$.

Given a γ-root$R$, we introduce the set

$$\mathcal{V}_R(\gamma) \coloneqq \{v \in \mathcal{V}(\gamma) : v \subset R\}. \tag{100}$$

Proposition (2.2.17)[24]: Let $Gr := Gr(A, \mathcal{D})$ be the spatially colorable digraph of a self-affine set $T: T(A, \mathcal{D})$with the set of vertices $\mathcal{V}$ and the set of edges $\mathcal{E}$. There exists a coloring $c: \mathcal{V} \to \Gamma$ such that the following is true:

(a) two distinct γ-roots are essentially disjoint, i.e., their intersection is of measure zero;

(b) each $\mathcal{V}_R(\gamma)$is a tree with respect to the set inclusion order;

(c) the family$\{\mathcal{V}_R(\gamma) : R \in \mathcal{R}(\gamma\}$forms a partition of $\mathcal{V}(\gamma)$:

$$\mathcal{V}(\gamma) = \coprod_{R \in \mathcal{R}(\gamma)} \mathcal{V}_R(\gamma). \tag{101}$$

Proof: Fix a coloring $c: \mathcal{V} \to \Gamma$ satisfying the condition of Definition(2.2.4). Then any two vertices of the same color are either essentially disjoint, or the smaller of them is a subset of the larger. This immediately implies assertions (a) and(c). Now, equip$\mathcal{V}_R(\gamma)$with the set inclusion order structure. This gives rise to a digraph $Gr_R(\gamma) \coloneqq (\mathcal{V}_R(\gamma)\mathcal{E}_R(\gamma))$, with the set of edges defined as follows.

A pair$T', T'' \in \mathcal{V}_R(\gamma)$is an edge directed from T' to T'' if $T' \subset T''$and there are no other vertices of$\mathcal{V}_R(\gamma)$in-between.

To establish assertion(b), it suffices to show that every two vertices $T', T'' \in \mathcal{V}_R(\gamma)$can be joined by a unique (undirected) path. For this, we choose a vertex $\tilde{T} \in \mathcal{V}_R(\gamma)$of largest height containing $T'$and T''. Since all vertices of$\mathcal{V}_R(\gamma)$ are subsets of$R$, it does exist. Now we set $T_1 := T'$and let T_2 be a parent of T_1. The latter is unique, because each distinct γ-root containing T_1 should either be a subset of T_2 or contain T_2. Let T_3 be the parent of T_2 and so on up toT_n; all these are of the same height as T. Since$|\tilde{T} \cap T_n| > |T'| > 0$, one of them is a subset of the other. But their heights are equal, whence$\tilde{T} = T_n$. In the same way we define a sequence $\hat{T}_1 \coloneqq T'', \hat{T}_2, \dots, \hat{T}_m = \tilde{T}$. Then the sequence$\{T_1, \dots, T_2, \hat{T}_{m-1}, \dots, \hat{T}_1\}$is a unique path connecting $T'$with T''.Consequently,$Gr_R(\gamma)$ is a tree, and it is rooted in$R$, since all vertices of$\mathcal{V}_R(\gamma)$ are subsets of $R \in \mathcal{V}_R(\gamma)$.

Let $\mathcal{F} \coloneqq \{F_j; j \in \mathbb{Z}\}$be a family of subspaces of$L_p(\mathbb{R}^n)$, $0 < p \le \infty$, satisfying the conditions

$$F_j \subset F_{j+1}, j \in \mathbb{Z}, \text{and} \sup_j E_j(f) \neq 0 \text{ if } f \neq 0. \tag{102}$$

Here the best approximation $E_j(f)$is given by

$$E_j(f) \coloneqq \inf_{g \in F_j} \|f - g\|_p. \tag{103}$$

We introduce an approximation space $\mathcal{A}_p^s(\mathcal{F})$, $s > 0$, by the quasinorm

$$\|f\|_{\mathcal{A}_p^s(\mathcal{F})} \coloneqq \left\{ \sum_{j \in \mathbb{Z}} \left(a^{js} E_j(f) \right)^p \right\}^{\frac{1}{p}}, \tag{104}$$

where $a > 1$ is fixed. Let $p < q \le \infty$be defined by the relation

$$\frac{s}{n} =: \frac{1}{p} - \frac{1}{q}, \tag{105}$$

and assume that $p \le 1$ if $q = \infty$.Also, assume that

$$\|f\|_\infty \le C a^{\frac{jn}{p}} \|f\|_p, f \in F_j, j \in \mathbb{Z}, \tag{106}$$

with a constant independent of f and j.

Under these assumptions, the following is true.

Theorem (2.2.18)[24]: $\mathcal{A}_p^s(\mathcal{F}) \subset L_q(\mathbb{R}^n)$.

For the proof, see [22].

From this result we deduce the corresponding embedding for the space $B_p^s(\varphi)$. For this, we need to present $B_p^s(\varphi)$as an approximation space(104) with a suitable approximation family $\mathcal{F}$. Let $\mathcal{F}_j$ be the linear subspace of $L_q(\mathbb{R}^n)$formed by the functions represented as

$$f = \sum_{k\in\mathbb{Z}^n} c_j(k)\varphi_{jk} \text{ (convergence in } L_p). \tag{107}$$

The scaling equation(71) implies that$\mathcal{F}_j \subset \mathcal{F}_{j+1}$. Below we also prove that for $a \coloneqq |\det A|^{\frac{1}{n}}$we have

$$\|f\|_\infty \le c a^{\frac{jn}{p}} \|f\|_p, f \in F_j, j \in \mathbb{Z}, \tag{108}$$

with a constant independent of f and$j$. This implies that the supremum in(102) is equal to the p-norm of$f$. Hence, the family $\mathcal{F}$ of the subspaces F_j chosen above satisfies condition(102). Let an approximation space $\mathcal{A}_p^{s\theta}(\mathcal{F})$,$0 < \theta$,$p \le \infty$, $s > 0$, be introduced by

$$\|f\|_{\mathcal{A}_p^{s\theta}(\mathcal{F})} \coloneqq \left\{\sum_{j\in\mathbb{Z}} \left(a^{js}E_j(f)\right)^\theta\right\}^{\frac{1}{\theta}}, \tag{109}$$

where $a \coloneqq |\det A|^{\frac{1}{n}}$.

Note that $\mathcal{A}_p^{sp}(\mathcal{F})$coincides with the space $\mathcal{A}_p^s(\mathcal{F})$of(104). The next result compares the space(109) with that in(87).

Proposition (2.2.19)[24]: If$\varphi$is stable, then

$$\mathcal{A}_p^{s\theta}(\mathcal{F}) = B_p^{s\theta}(\varphi). \tag{110}$$

Proof. By [29] (for the particular case under consideration see also [30]), for $\mathcal{F}$ satisfying(102) the equivalence

$$\|f\|_{\mathcal{A}_p^{s\theta}(\mathcal{F})} \approx \inf\left\{\sum_{j\in\mathbb{Z}} \left(a^{js}\|f_j - f_{j-1}\|\right)^\theta\right\}^{\frac{1}{\theta}} \tag{111}$$

is valid uniformly in$f \in \mathcal{A}_p^{s\theta}(\mathcal{F})$. Here the infimum is taken over all expansions

$$f = \sum_{j\in\mathbb{Z}} \left(f_j - f_{j-1}\right) \text{ (convergence in } L_p)$$

with $f_j \in \mathcal{F}_j$. Since for such$f_j$we have

$$f_j - f_{j-1} = \sum c_j(k)\varphi_{jk}$$

(see(107) and(71)), the set of these expansions for f coincides with that involved in the definition of $B_p^{s\theta}(\varphi)$(see(87)).

We show that if $\mu := s - \frac{n}{p}$, then

$$a^{js}\left\|f_j - f_{j-1}\right\|_p \le C\left\{\sum_{k\in\mathbb{Z}}\left(\left|\operatorname{supp}\varphi_{jk}\right|^{\mu}\left|c_j(k)\right|\right)^p\right\}^{\frac{1}{p}}. \tag{112}$$

Raising to the power θ and summing over j, and then applying(111) and(87), we obtain

$$\|f\|_{\mathcal{A}_p^{s\theta}(\mathcal{F})} \le C\|f\|_{B_p^{s\theta}(\varphi)}. \tag{113}$$

Next, a change of variables reduces(112) to the case where $j = 0$, that is, to the inequality

$$\left\|\sum_{k\in\mathbb{Z}^n} c(k)\varphi(x-k)\right\|_p \le C\left\{\sum_{k\in\mathbb{Z}}|x(k)|^p\right\}^{\frac{1}{p}},$$

which is true by the stability of φ(see(81)). Since the stability of φ provides the inequality reverse to(112), we also have

$$\|f\|_{B_p^{s\theta}(\varphi)} \le C\|f\|_{\mathcal{A}_p^{s\theta}(\mathcal{F})}.$$

Together with(113), this completes the proof of the proposition to within inequality (108). By a change of variables, the latter reduces to the estimate

$$\left\|\sum_{k\in\mathbb{Z}^n} c(k)\varphi(x-k)\right\|_\infty \le C\left\|\sum_{k\in\mathbb{Z}^n} c(k)\varphi(x-k)\right\|_p. \tag{114}$$

For the proof of(114) we need the following fact.

Lemma (2.2.20)[24]: The family$\{\operatorname{supp}\varphi(x-k) : k\in\mathbb{Z}^n\} = \{k + \operatorname{supp}\varphi : k\in\mathbb{Z}^n\}$is C-disjoint with $C = C(\varphi)$. [2]

Proof: Let

$$m(\varphi) \coloneqq \operatorname*{ess\,sup}_x\left(\sum_{k\in\mathbb{Z}^n} 1_{\operatorname{supp}\varphi}(x-k)\right)$$

be the multiplicity of this family. Since φ is compactly supported,$m(\varphi) < \infty$. Then the result in [31] implies that the family under consideration is C-disjoint with $C \le C(n)m(\varphi)$.

Using this lemma and the stability of φ(see(81)), we now bound the left-hand side of(114) by

$$C(\varphi)\sup_k|c(k)| \le C(\varphi)\left\{\sum_{k\in\mathbb{Z}^n}|c(k)|^p\right\}^{\frac{1}{p}} \le C_1(\varphi)\left\|\sum_{k\in\mathbb{Z}^n} c(k)\varphi(x-k)\right\|_p$$

Proposition(2.2.19) is established.

Now(110) and(108) allow us to apply Theorem(2.2.18) in order to obtain the required result.

Corollary (2.2.21)[24]: Ifs, p, q satisfy the condition of Theorem(2.2.18) (see(105)), then

$$B_p^s(\varphi) \subset L_q(\mathbb{R}^n).$$

Remark (2.2.22)[24]: We shall use this embedding in the form of the inequality

$$\left\| \sum_{k \in \mathbb{Z}^n} c(j,k)\varphi_{jk} \right\|_q \leq C \left\{ \sum_{j,k} \left(|c(j,k)|. \left|\operatorname{supp} \varphi_{jk}\right|^{\frac{1}{q}} \right)^p \right\}^{\frac{1}{p}}, \tag{115}$$

which follows from the definition of the (quasi)norm in$B_p^s(\varphi)$(see(85)).

That is, it can be divided into at most C subfamilies of pairwise essentially disjoint subsets.

Remark (2.2.23)[24]: Let $A, \bar{\ell}$, and φ be as in Remark(2.2.16). Then an analog of inequality (108)looks like this:

$$\|f\|_{W_q^{\bar{\ell}}} \leq c a^{jn\left(\frac{1}{p}+\sigma\right)} \|f\|_p, j \in \mathbb{Z}$$

(see(96) and(94)). In its turn, this leads to the inequality

$$\left\| \sum_{j,k} c(j,k)\varphi_{jk} \right\|_{W_q^{\bar{\ell}}} \leq C \left\{ \sum_{j,k} \left(|c(j,k)| \left|\operatorname{supp} \varphi_{jk}\right|^{\frac{1}{q}} \right)^p \right\}^{\frac{1}{p}} \tag{116}$$

with s, σ, p, q related by(98).

Reducing Theorem (2.2.27) to a special case. Let $f \in B_p^s(\varphi)$. In what follows we assume, as we may, that

$$\frac{1}{2} \leq \|f\|_{B_p^s(\varphi)} < 1, \tag{117}$$

and therefore (see(85)), there is a representation

$$f = \sum_{j,k} c_{jk}\varphi_{jk} \text{ (convergence in } L_p) \tag{118}$$

such that

$$\nu(f) := \left\{ \sum_{j,k} \left(|c_{jk}|. \left|\operatorname{supp} \varphi_{jk}\right|^{\frac{1}{q}} \right)^p \right\}^{\frac{1}{p}} = 1. \tag{119}$$

The required reduction of Theorem(2.2.27) will be attained in two steps.

We show that Theorem(2.2.27) can be derived from the following result.

Proposition (2.2.24)[24]: Suppose that conditions (a) and(b) of Theorem(2.2.27) are fulfilled. Also, assume that a function g has a representation of the type(118) with coefficients d_{jk} satisfying the conditions

$$\nu(g):=\left\{\sum_{j,k}\left(\left|d_{jk}\right|.\left|\operatorname{supp}\varphi_{jk}\right|^{\frac{1}{q}}\right)^{p}\right\}^{\frac{1}{p}}<\infty, \tag{120}$$

$$\operatorname{supp} d:=\left\{(j,k):d_{jk}\neq 0\right\}\subset\left\{(j,k):\operatorname{supp}\varphi_{jk}\subset\operatorname{supp}\varphi\right\}. \tag{121}$$

Then for each integer $N\geq 1$ there is a linear combination $g_N\in\mathcal{L}_M(\varphi)$ such that

$$\operatorname{supp} g_N\subset\operatorname{supp}\varphi, \tag{122}$$

$$M\leq CN_{\nu}(g)^{p},\text{in particular}, g_N=0\text{if } M<1, \tag{123}$$

$$\|g-g_N\|_q\leq CN^{-\frac{s}{n}}\nu(g). \tag{124}$$

Here and in the sequel C stands for a constant depending only on φ and $p^*:=\min(1,p)$. This C may change from line to line.

We show that this proposition implies Theorem(2.2.27). Suppose φ and f satisfy the conditions of that theorem and(117)–(119) are fulfilled. Given $N\geq 1$, we choose the largest $j_0\in\mathbb{Z}$ such that

$$\sum_{k\in\mathbb{Z}^n}\sum_{j\leq j_0}\left(\left|c_{jk}\right|\left|\operatorname{supp}\varphi_{jk}\right|\right)^{p}\leq N^{-\left(1-\frac{p}{q}\right)}.$$

Set$f_-:=\sum_k\sum_{j\leq j_0}$ and$f_+:=f_f_-$. Together with the embedding(115) and conditions(119) and(117), this choice of$j_0$leads to the inequalities

$$\|f_-\|_q\leq C\|f_-\|_{B_p^s(\varphi)}\leq C\nu(f_-)^{\frac{1}{p}}\leq CN^{\frac{1}{q}-\frac{1}{p}}\leq 2CN^{-\frac{s}{n}}\|f\|_{B_p^s(\varphi)}.$$

Hence, it suffices to prove the result for the functionf_+. Since the assumptions and claims of Theorem(2.2.27) are invariant under the transformation$F(x)\to|\det A|^{-\frac{j_0}{q}}F\left(A^{-j_0}x\right)$, we may assume that $j_0=0$. Then

$$f_+=\sum_{k\in\mathbb{Z}^n}g_k, \tag{125}$$

Where

$$g_k=\sum c_{jk},\varphi_{jk}, \tag{126}$$

with(j,k')running over the set$\{(j,k'): \operatorname{supp}\varphi_{jk'} \subset \operatorname{supp}\varphi_{0k}\}$.After shifting by k, the function g_k will satisfy the assumptions of Proposition(2.2.24). This implies the existence of a linear combination $g_{N,k} \in \mathcal{L}_{M_k}(\varphi)$such that

$$\operatorname{supp} g_{N,k} \subset \operatorname{supp} g_k \ (\subset \operatorname{supp}\varphi_{0k}), \tag{127}$$

$$M_k \leq CN\nu(g_k)^p, \tag{128}$$

$$\left\|g_k - g_{N,k}\right\|_q \leq CN^{-\frac{s}{n}}\nu(g_k). \tag{129}$$

Now we set

$$f_N := \sum_k g_{N,k}.$$

Since the family $\{\operatorname{supp} g_k : k \in \mathbb{Z}^n\}$is C-disjoint (see(127) and Lemma(2.2.20)), relations(125)– (129) imply that

$$\|f_+ - f_N\|_q \leq C\left\{\sum_k \left\|g_k - g_{N,k}\right\|_q^q\right\}^{\frac{1}{q}} \leq CN^{-\frac{s}{n}}\left\{\sum_k \nu(g_k)^q\right\}^{\frac{1}{q}}.$$

The definition of g_k and the Jenssen inequality yield

$$\left\{\sum_k \nu(g_k)^q\right\}^{\frac{1}{q}} \leq \left\{\sum_{k,j}\left(\left|c_{jk}\right|\left|\operatorname{supp}\varphi_{jk}\right|^{\frac{1}{q}}\right)^p\right\}^{\frac{1}{p}} = \nu(f_+).$$

Since $\nu(f_+) \leq \nu(f) = 1 \leq 2\|f\|_{B_p^s(\varphi)}$(see(117)), the linear combinationg_N approximates $f = f_+ + f_-$ in $L_p(\mathbb{R}^n)$with the required rate. Moreover, by(128), the number of its terms is at most $\sum_k M_k \leq CN\sum \nu(g_k)^p = CN\nu(f_+)^p \leq CN$. Thus, Proposition (2.2.24) implies Theorem(2.2.27)

In its turn, Proposition(2.2.24) is a consequence of the result presented below.

Suppose φ satisfies the conditions of Theorem(2.2.27).Then there is a digit set $\mathcal{D} \supset \operatorname{supp} m$such that the chromatic number of $T := T(A,\mathcal{D})$ is bounded, and

$$\chi(\varphi) = \chi(A,\mathcal{D}) < \infty. \tag{130}$$

Recall that for this T we have

$$\operatorname{supp}\varphi \subset T, \tag{131}$$

Whence

$$\operatorname{supp}\varphi_{jk} \subset T_{jk} := A^{-j}(T+k). \tag{132}$$

We apply Proposition(2.2.17) to the partition of $\mathcal{V}$ into the collection of trees $\mathcal{V}_R(\gamma)$ with $R \in \mathcal{R}(\gamma)$ and γ belonging to the set Γ of colors. Here $\mathcal{V}$ is the set of vertices of the digraph $Gr := Gr(A,\mathcal{D})$. We shall index the functions $\varphi_{jk}$and the coefficients c_{jk} of the corresponding expansions by the subscripts $T' \in \mathcal{V}$, setting

$$\varphi_{T'} := \varphi_{jk} \text{ and} c_{T'} := c_{jk} \text{ if } T' = T_{jk}. \tag{133}$$

We formulate a result implying Proposition(2.2.24).

Proposition (2.2.25)[24]: Suppose that conditions(a) and(b) of Theorem(2.2.27) are fulfilled. Also, assume that a function g admits a representation

$$g = \sum_{T' \in \mathcal{V}} d_{T'} \varphi_{T'} \tag{134}$$

such that

$$\nu(g) := \left\{ \sum_{T' \in \nu} \left(|d_{T'}| |\operatorname{supp} \varphi_{T'}|^{\frac{1}{q}} \right)^p \right\}^{\frac{1}{p}} < \infty, \tag{135}$$

and, moreover,

$$\operatorname{supp} d := \{T' \in \mathcal{V} : d_{T'} \neq 0\} \subset \mathcal{V}_R(\gamma) \tag{136}$$

for a given γ-root R.

Then for every $N \geq 1$ there is a linear combination$g_N \in \mathcal{L}_M(\varphi)$such that

$$\operatorname{supp} g_N \subset R, \tag{137}$$

$$M \leq CN\nu(g)^p, \tag{138}$$

$$g_N = 0 \text{ if } N\nu(g)^p < 1, \tag{139}$$

$$\|g - g_N\|_q \leq CN^{-\frac{s}{n}} \nu(g), \tag{140}$$

provided $g_N \neq 0$. We derive Proposition (2.2.24) from Proposition (2.2.25). Assume that a function g satisfies the hypothesis of Proposition (2.2.24). Using the notation introduced above and condition(121), we can rewrite the representation for g as

$$g = \sum_{T' \in \mathcal{V}} d_{T'} \varphi_{T'}. \tag{141}$$

This is possible because

$$\mathcal{V} = \{T_{jk} : \operatorname{supp} \varphi_{jk} \subset \operatorname{supp} \varphi\}$$

(see(131) and(132)). Inequality(120) can be rewritten in a similar way. Using the partition(99) of the set $\mathcal{V}$, we write

$$g = \sum_{\gamma \in \Gamma} g_\gamma, \tag{142}$$

where$g_\gamma$is given by

$$g_\gamma = \sum_{T' \in \mathcal{V}(\gamma)} d_{T'} \varphi_{T'}. \tag{143}$$

If we prove Proposition(2.2.24) for each function g_γ, then(142) allows us to establish the same for the function g up to the multiplicative constant $\#\Gamma$. Since all the assertions of that proposition are homogeneous, without loss of generality we may assume that

$$\nu(g_\gamma) = 1, \tag{144}$$

and derive the desired result for this g_γ. Using the partition(101) of $\mathcal{V}(\gamma)$, we obtain

$$g_\gamma = \sum_{R\in\mathcal{R}(\gamma)} g_{\gamma,R}, \tag{145}$$

where $g_{\gamma,R}$ is given by

$$g_{\gamma,R} := \sum_{T'\in\mathcal{V}_R(\gamma)} d_{T'}\varphi_{T'}. \tag{146}$$

Each function $g_{\gamma,R}$ satisfies the assumptions of Proposition(2.2.25), and we conclude that for each $N \geq 1$ there is a linear combination $g_{N,\gamma,R} \in \mathcal{L}_{M(\gamma,R)}(\varphi)$ such that the following is true:

$$\operatorname{supp} g_{N,\gamma,R} \subset \operatorname{supp} g_{\gamma,R}, \tag{147}$$

$$M(\gamma,R) \leq CN\nu(g_{\gamma,R})^p < 1, \tag{148}$$

$$\left\|g_{\gamma,R} - g_{N,\gamma,R}\right\|_q \leq CN^{-\frac{s}{n}}\nu(g_{\gamma,R}), \tag{149}$$

provided $g_{N,\gamma,R} \neq 0$.

Let $\mathcal{R}_+$ and $\mathcal{R}_0$ be, respectively, the sets of all $R \in \mathcal{R}(\gamma)$ such that $g_{N,\gamma,R} \neq 0$ and $g_{N,\gamma,R} = 0$. We define the required approximant by the formula

$$g_{N,\gamma,R} := \sum_{R\in\mathcal{R}_+} g_{N,\gamma,R}. \tag{150}$$

Then, by(147) and(132),

$$\operatorname{supp}(g_{\gamma,R} - g_{N,\gamma,R}) \subset \operatorname{supp} g_{\gamma,R} \subset \bigcup\{T' : T' \in \mathcal{V}_R(\gamma)\}.$$

The latter union is a subset of the γ-root R (see(100)), and the set of all these γ-roots is essentially pairwise disjoint. Consequently, the family $\{\operatorname{supp}(g_{\gamma,R} - g_{N,\gamma,R}) : R \in \mathcal{R}(\gamma)\}$ has the same property. This implies the identity

$$\left\|g_\gamma - g_{N,\gamma}\right\|_q = \left\{\sum_{R\in\mathcal{R}_+} \left\|g_\gamma - g_{N,\gamma,R}\right\|_q^q + \left\|\sum_{R\in\mathcal{R}_0} g_{\gamma,R}\right\|_q^q\right\}^{\frac{1}{q}}. \tag{151}$$

The first sum is estimated by applying(149) and then Jenssen's inequality. This and(144) yield the desired inequality

$$\sum_{R\in\mathcal{R}_+} \leq CN^{-\frac{s}{n}}\left\{\sum_{R\in\mathcal{R}_+}\nu\big(g_{\gamma,R}\big)^p\right\}^{\frac{1}{p}} \leq CN^{-\frac{s}{n}}\big(g_\gamma\big) = CN^{-\frac{s}{n}}. \tag{152}$$

To estimate the second sum in(151) by the same bound, we enumerate all $R \in \mathcal{R}_0$ in a sequence$\{R_i: i \in \mathbb{N}\}$ such that the numbers $\nu_i := \nu\big(g_{\gamma,R_i}\big)$become monotone nonincreasing. Then we choose an interval $I_1 := (0, i_1]$ with $i_1 \in \mathbb{Z}_+$such that

$$\sum_{i\in I_1} \nu_i^P < N^{-1},$$

While

$$\sum_{i\in I_1^+} \nu_i^p \geq N^{-1}$$

for $I_2^+ := [0, i_1 + 2]$. The interval I_1 may be empty; in this case $I_1^+ \cap \mathbb{N} = \{i_1 + 1\}$.Also, it may happen that $I_1 = (0, +\infty)$and in this case $I_1^+ = \emptyset$. If $I_1^+$is nonempty, we continue this construction by choosing an interval $I_2 := [i_1 + 2, i_2]$ such that

$$\sum_{i\in I_2} \nu_i^p < N^{-1},$$

while

$$\sum_{i\in I_2^+} \nu_i^p \geq N^{-1}$$

for $I_2^+ := [i_1 + 2, i_2 + 1]$.Since$\sum \nu_i^p \leq \nu\big(g_\gamma\big)^p = 1$, this procedure yields a finite set of subsequent intervals $I_1, I_1^+, \dots, I_\ell, I_\ell^+$, where I_ℓ is unbounded, $I_\ell^+ = \emptyset$, and $I_m^+ \backslash I_m$contains the only integer $i_m^+ := i_m + 1, 1 \leq m < \ell$.

For these intervals we have

$$\sum_{i\in I_m} \nu_i^p < N^{-1}, \sum_{i\in I_m^+} \nu_i^p \geq N^{-1} \tag{153}$$

with $1 \leq m \leq \ell$ in the first inequality and $1 \leq m < \ell$ in the second. Observe also that the definition of $\mathcal{R}_0$ and(148) imply that

$$\nu_i < N^{-\frac{1}{p}}, i \in \mathbb{N}. \tag{154}$$

Now we set

$$\psi_m := \sum_{i\in I_m} g_{\gamma,R_i}, \psi_m^+ := g_{\gamma,R_{i_m^+}}, 1 \leq m \leq \ell.$$

Since the supports of these functions are pairwise essentially disjoint, we have

$$\left\|\sum_{R\in\mathcal{R}_0} g_{\gamma,R}\right\|_q = \left\{\sum_{m=1}^{\ell}\|\psi_m\|_q^q + \sum_{m=1}^{\ell}\|\psi_m^+\|_q^q\right\}^{\frac{1}{q}}.$$

Applying the embedding(115) to each term on the right, we bound this quantity by

$$C\left\{\sum_{m=1}^{\ell}\nu(\psi_m^+)^q + \sum_{m=1}^{\ell}\|\psi_m\|_q^q\right\}^{\frac{1}{q}}.$$

Inequality(153) implies the estimate

$$\sum_{m=1}^{\ell}\nu(\psi_m)^q = \sum_{m=1}^{\ell}\left(\sum_{i\in I_m}\nu_i^p\right)^{\frac{q}{p}} \le \ell N^{-\frac{q}{p}},$$

while(154) yields

$$\sum_{m=1}^{\ell}\|\psi_m^+\|_q^q = \sum_{m=1}^{\ell}\nu_{i_m^+}^q < \ell N^{-\frac{q}{p}}.$$

To estimate ℓ , we use the second inequality in(153) to obtain

$$(\ell-1)^{N^{-1}} \le \sum_{m<\ell}\sum_{i\in I_m^+}\nu_i^p \le \sum_{i=1}^{\infty}\nu_i^p \le 1,$$

whence $\ell \le N+1$.

Collecting all these inequalities, we get

$$\left\|\sum_{R\in\mathcal{R}_0} g_{\gamma,R}\right\|_q \le C\left(\ell N^{-\frac{q}{p}}\right)^{\frac{1}{q}} \le CN^{\frac{1}{q}-\frac{1}{p}} = CN^{-\frac{s}{d}}.$$

Together with(152), this proves the required estimate(124) for g_γ.

Now we estimate the number of terms in the linear combination $g_{N,\gamma}$.

$$M(\gamma) := \sum_{R\in\mathcal{R}_+} M(\gamma,R) \le CN\sum_{R\in\mathcal{R}(\gamma)}\nu\left(g_{\gamma,R}\right)^p = CN\nu\left(g_{\gamma,R}\right)^p = CN.$$

Thus, assertion(123) is also true for g_γ.

We define the required approximant g_N for the function g as in(142) by setting

$$g_N = \sum_{\gamma\in\Gamma} g_{N,\gamma}$$

(see(150)). Then inequalities(138) and(140) for $g_\gamma$and the linear combination $g_{N,\gamma}$ imply the same results for g and g_N with an additional factor of#Γ. Consequently,(123) and(124) are true in this case, and it remains to check(122). By(147) and(142),

$$\operatorname{supp} g_N \subset \bigcup_{\gamma\in\Gamma} \operatorname{supp} g_{\gamma,R} \subset \bigcup_{\gamma} \bigcup_{R\in\mathcal{R}(\gamma)} \operatorname{supp} g_{\gamma,R}.$$

In its turn, supp $g_{\gamma,R}$is a subset of the set $\cup\{\operatorname{supp}\varphi_{T'} : T' \in \mathcal{V}_R(\gamma)\}$(see(146)), and the union of all γ-roots $R,\gamma \in \Gamma$, coincides with the set $\mathcal{V}$ of all vertices; see Proposition(2.2.17).

Finally, $\mathcal{V} = \{T' : \operatorname{supp}\varphi_{T'} \subset \operatorname{supp}\varphi\}$, whencesupp $g_\gamma \subset \operatorname{supp}\varphi$, and we see that

$$\operatorname{supp} g_N \subset \operatorname{supp}\varphi.$$

This proves the final assertion(122) of Proposition(2.2.24).

We introduce a nonlinear method of approximation that will be used in the proof of Proposition(2.2.25). The input of the corresponding approximation algorithm consists of an integer $N \geq 1$ and a function

$$d: T' \to d_{T'} \in \mathbb{R}, T' \in \mathcal{V}_R(\gamma), \tag{155}$$

satisfying the condition

$$\nu(d) \coloneqq \left\{ \sum_{T'\in\mathcal{V}_R(\gamma)} \left(|d_{T'}| |\operatorname{supp}\varphi_{T'}|^{\frac{1}{q}} \right)^p \right\}^{\frac{1}{p}} < \infty. \tag{156}$$

Recall that $\mathcal{V}_R(\gamma)$ is the set of vertices of the tree $Gr_R(\gamma) = \big(\mathcal{V}_R(\gamma), \mathcal{E}_R(\gamma)\big)$rooted at R (see Proposition(2.2.17)(b)).

Given N and the function d, we introduce the cost function$\mathcal{J}$defined on the subsets $\Omega \subset \mathcal{V}_R(\gamma)$by

$$\mathcal{J}(\Omega) \coloneqq \sum_{T'\in\Omega} \left(|d_{T'}| |\operatorname{supp}\varphi_T|^{\frac{1}{q}} \right)^p. \tag{157}$$

For the subset

$$\mathcal{V}_R(\gamma; T') \coloneqq \left\{ T'' \in \mathcal{V}_R(\gamma) : T'' \subset T' \right\} \tag{158}$$

with$T' \in \mathcal{V}_R(\gamma)$, we simplify this notation as follows:

$$\mathcal{J}(T') \coloneqq \mathcal{J}\big(\mathcal{V}_R(\gamma; T')\big). \tag{159}$$

Note that $\mathcal{J}(T') \neq \mathcal{J}(\{T'\})$and

$$\mathcal{J}(T) = \nu(d)^p < \infty. \tag{160}$$

It is readily seen that(158) is the set of vertices of a subtree of the tree $\mathcal{T}_R(\gamma)$with the root T'.

Now, assuming that N is such that

$$\mathcal{I}(R) \geq N^{-1}, \tag{161}$$

we define a set $G_N \subset \mathcal{V}_R(\gamma)$by

$$G_N := \{T' \in \mathcal{V}_R(\gamma): \mathcal{I}(T') \geq N^{-1}\}. \tag{162}$$

Since $\mathcal{I}(T') \geq \mathcal{I}(T'')$ if$T'' \subset T'$, this G_N is the set of vertices of a subtree of $Gr_R(\gamma)$, which is finite because $\mathcal{I}(T') \to 0, |T'| \to 0$(see(160)). Finally,$R \in G_N$ by(161), and itis the root of that subtree.

Using the set inclusion order on G_N, we introduce the set $\mathcal{M}_N$ of minimal elements of G_N. Minimality implies that, if $T' \in \mathcal{M}_N$ and T'' is an offspring of T', then (in$Gr_R(\gamma)$)

$$\mathcal{I}(T') \geq N^{-1}, \text{while} \mathcal{I}(T'') < N^{-1} \tag{163}$$

We enumerate the elements of $\mathcal{M}_N$ in some order:

$$\mathcal{M}_N := \{T_j^{\min}: 1 \leq j \leq m_N\}. \tag{164}$$

Being vertices of the tree $Gr_R(\gamma)$, these elements are either essentially disjoint, or one embeds into the other. The latter is impossible because of minimality (see (163)). Consequently, the subsets of $\mathcal{M}_N$ are pairwise (essentially) disjoint. This implies that

$$\nu(d)^p = \mathcal{I}(R) \geq \sum_j \mathcal{I}(T_j^{\min}) \geq \frac{m_N}{N},$$

Whence

$$m_N \leq N\nu(d)^p. \tag{165}$$

We partition G_N to obtain a collection$\mathcal{B}_N$ of directed paths (called basic paths). In the description of the corresponding partition algorithm, we shall use the following notation.

Let T_1 and T_2 be the tail and the head (respectively) of a directed path P in the tree $Gr_R(\gamma)$. Since P is uniquely determined by its endpoints, we write $P := [T_1, T_2]$. We recall (see(68)) that the endpoints of P are denoted by T_P^- and T_P^+. So,$P := [T_P^-, T_P^+]$.We also introduce subpaths of P "open from the head or tail" by setting

$$[T_P^-, T_P^+) := P \backslash \{T_P^+\} \text{and } (T_P^-, T_P^+] := P \backslash \{T_P^-\},$$

and so forth. We start with splitting$G_N$into a collection $\mathcal{A} := \{A_j: 0 \leq j \leq m_N\}$of "long" paths satisfying the following conditions.

(a) The subcollection$\{A_j: 0 \leq j \leq i\}$is a partition of the set

$$G_N^i := \bigcup_{j \leq 1} \left|T_j^{\min}, R\right|, i \leq m_N.$$

(b) Each long path A_j with $j \geq 1$ is of the form$A_j = [T_j^{\min}, T')$, where T'belongs to a suitable A_j with $j' < j$. This T' is called a contact vertex and is denoted by T_j^{c}. Thus,

$$A_j := [T_j^{\min}, T_j^{\mathrm{c}}), 1 \leq j \leq m_N. \tag{166}$$

Since $G_N^i = G_N$ for $i \coloneqq m_N$, the collection $\mathcal{A}$ forms the desired partition of the subtree G_N. Moreover, $\mathcal{A}$ determines the set of contact vertices

$$C_N \coloneqq \{T_i^c\} \cup \{R\}. \tag{167}$$

Some of these may coincide; therefore, the inequality

$$\#C_N \le m_N + 1, \tag{168}$$

can be strict.

In order to introduce $\mathcal{A}$, we use induction on j, starting with

$$A_0 := \{R\} \text{ and} A_1 [T_1^{\min}, R] \backslash A_0 [T_1^{\min}, R).$$

Next, assuming that some A_i satisfying$(a)$and(b)has been determined for $i = 0, 1, \dots, j$, we introduce A_{j+1} by $A_{j+1} \coloneqq [T_{j+1}^{\min}, R] \backslash (\cup_{i \le j} A_i)$. Then, clearly, the collection$\{A_i : 0 \le i \le j+1\}$ forms a partition of G_N^{j+1}. We show that $A_{j+1}$is of the form(166). Indeed, consider the intersection of$[T_{j+1}^{\min}, R]$with each path$[T_i^{\min}, R], i \le j$. Since G_N is a tree rooted at R, this intersection is of the form $[T_i, R]$, and the set of the tails$\{T_i : 1 \le i \le j\}$ is a subset of the path$[T_{j+1}^{\min}, R]$. Therefore, the set of tails inherits thelinear order of this path. If $T_{i_0}$is the smallest element of $\{T_i\}$with respect to this order, then

$$A_{j+1} = [T_{j+1}^{\min}, R] \backslash \left(\bigcup_{i \le j} [T_i^{\min}, R] \right) = [T_{j+1}^{\min}, T_{i_0}).$$

Moreover, $T_{i_0} \in \cup_{i \le j} A_i$, which completes the induction.

We refine G_N by subdividing each long path A_j with the help of the contact vertices belonging to $A_j \cap C_N$. In this way we introduce a collection of subpaths$[T'.T'')$, where $T'$is either a minimal element, or a contact vertex, and $T''$is a contact vertex. The set of such “intermediate” subpaths is denoted by $\mathcal{P}_N$. In accordance with this definition, we have

$$\operatorname{supp} \mathcal{P}_N \coloneqq \bigcup \{P : P \in \mathcal{P}_N\} = G_N \backslash \{R\}, \tag{169}$$

and different subpaths in $\mathcal{P}_N$ do not intersect. In other words, $\mathcal{P}_N$ is a partition of $G_N \backslash \{R\}$.

We complete the partition of G_N by subdividing each subpath $P \in \mathcal{P}_N$ into basic paths as follows.

Inductively, we define a collection of vertices$\{T_\ell(P) \in P : 1 \le \ell \le \ell_P\}$ beginning with$T_1(P) \coloneqq T_P^-$. If$T_\ell(P)$has been determined, then we choose$T_{\ell+1}(P)$as a vertex in$(T_\ell(P), T_P^+]$satisfying

$$\mathcal{I}([T_\ell(P), T_{\ell+1}(P)]) \ge N^{-1}, \text{while} \mathcal{I}([T_\ell(P), T_{\ell+1}(P))) < N^{-1}.$$

Using this, we define a basic path $B_\ell(P)$by

$$B_\ell(P) \coloneqq [T_\ell(P), T_{\ell+1}(P)). \tag{170}$$

The vertex $T_{\ell+1}(P)$ can be undetermined in the following two cases.

(a) The vertex $T_\ell(P)$ coincides with the head T_P^+, or $\mathcal{I}([T_\ell(P), T_P^+]) < N^{-1}$.

Then we define $T_{\ell+1}(P)$ as a parent of T_P^+; in the subtree G_N this parent is unique.

(b) The vertex $T_\ell(P)$ is distinct from T_P^+, but $\mathcal{I}(\{T_\ell(P)\}) \geq N^{-1}$.

Then we define $T_{\ell+1}(P)$ as a parent of $T_\ell(P)$.

In the two cases above, we introduce the basic path $B_\ell(P)$ by the same formula (170). Observe that in case (a) we have

$$B_\ell(P) = [T_\ell(P), T_P^+],$$

and the procedure is completed with $\ell_P := \ell$.

In case (b), the basic path $B_\ell(P)$ is the singleton $\{T_\ell(P)\}$, while $T_{\ell+1}(P)$ still belongs to P and the procedure can be continued.

Completing the procedure, we arrive at the partition $\{T_\ell(P): 1 \leq \ell \leq \ell_P\}$ of P into the basic paths $B_\ell(P) := \big[T_\ell(P), T_{\ell+1}(P)\big)$. Their definition implies that

$$\mathcal{I}\big(T_\ell(P), T_{\ell+1}(P)\big) < N^{-1} \tag{171}$$

for $\ell \leq \ell_P$, and

$$\mathcal{I}([T_\ell(P), T_{\ell+1}(P)]) \geq N^{-1} \tag{172}$$

for $\ell \leq \ell_P$, provided $B_\ell(P) := \big[T_\ell(P), T_{\ell+1}(P)\big)$ contains more than one vertex. For a singleton $B_\ell(P) := \{T_\ell(P)\}$ the first inequality makesno sense, while the second becomes

$$\mathcal{I}\big(B_\ell(P)\big) \geq N^{-1}. \tag{173}$$

Collecting all the basic paths of all $P \in \mathcal{P}$, we introduce the desired set

$$\mathcal{B}_N := \{B_\ell(P): 1 \leq \ell \leq \ell_P, P \in \mathcal{P}_N\}. \tag{174}$$

The following result describes its main features.

Proposition (2.2.26)[24]: (a) B_N is a partition of the set $G_N \backslash \{R\}$ into directed paths.

(b) For each $B := [T_B^-, T_B^+]$ in $\mathcal{B}_N$ we have

$$\mathcal{I}\big((T_B^-, T_B^+]\big) < N^{-1}. \tag{175}$$

(c) The cardinality of $\mathcal{B}_N$ satisfies

$$\#\mathcal{B}_N \leq (4N + 1)\nu(d)^p. \tag{176}$$

Proof: (a) follows from (169) and the definition of the basic paths; (b) follows from (171), because $(T_B^-, T_B^+] = \big(T_\ell(P), T_{\ell+1}(P)\big)$ provided $B := B_\ell(P)$. To prove (c), note that the family $\{[T_\ell(P), T_{\ell+1}(P)]: 1 \leq \ell \leq \ell_P\}$ covers P with multiplicity of at most 2.Therefore, by (172),

$$N^{-1}(\ell_P - 1) \leq \sum_{\ell < \ell_P} \mathcal{I}([T_\ell(P), T_{\ell+1}(P)]) \leq 2\mathcal{I}(P),$$

which leads to the inequality

$$\#\mathcal{B}_N = \sum_{P\in\mathcal{P}_N} \ell_P \le 2N \sum_{P\in\mathcal{P}_N} \mathcal{I}(P) + (\#\mathcal{P}_N).$$

By the definitions of the cost function I (see(157)) and the partition$\mathcal{P}_N$, the first term on the right is at most

$$2N\mathcal{I}(G_N) \le 2N\mathcal{I}(R) = 2N\nu(d)^p$$

(see(160)). Moreover, by(165) and(168),

$$\#\mathcal{P}_N \le (\#C_N) + (\#M_N) \le (2N+1)\nu(d)^p.$$

Combining these inequalities, we get the desired estimate(176).

We define the required approximant g_N. First, we consider the (trivial) case of $N \ge 1$ satisfying

$$\mathcal{I}(R) < N^{-1}; \tag{177}$$

then the required approximant is given simply by

$$g_N := 0. \tag{178}$$

Otherwise, the family $\mathcal{B}_N$ is determined, and each basic path $B \in \mathcal{B}_N$ gives rise to a part of the linear combination g_N as follows.

For$T' \in B$, let$j$ denote the height$h(T')$ (see(11)). Then

$$j \le j_B := h(T_B^-).$$

Identity(24) yields

$$\varphi_{T'} = \sum_{h(T'')=j_B} m(T', T'')\varphi_{T''} \tag{179}$$

With

$$m(T', T'') := m^*(j_B - j, k) \text{for} T'' := T_{j_B,k}, \tag{180}$$

Each vertex T''occurring here is an offspring of T'in the digraph $Gr(A, \mathcal{D})$. This and (179) imply the embeddings

$$\operatorname{supp}\varphi_{T'} \subset \bigcup\{\operatorname{supp}\varphi_{T''} : m(T', T'') \ne 0\}$$

$$\subset \bigcup\{T'' : T''\text{is an offspring of } T'\} = T'.$$

In particular,

$$\operatorname{supp}\varphi_{T''} \subset \operatorname{supp}\varphi_{T'} \subset T' \subset T_B^+, \tag{181}$$

provided $m(T', T'') \ne 0$ and $T' \in B$.

Now, for each $T' \in B$ we define a function $\phi_{T'}$by

$$\phi_{T'} := \sum_{T''} m(T', T'')\varphi_{T''}, \tag{182}$$

where T'' runs over the set of indices in (179) satisfying

$$\left|\operatorname{supp} \varphi_{T''} \cap T_B^-\right| \neq 0. \tag{183}$$

This definition and (181) yield

$$\operatorname{supp}(\varphi_{T'} - \phi_{T'}) \subset T_B^+ \backslash T_B^-. \tag{184}$$

Now we are in a position to introduce the output of the algorithm, namely, the linear combination g_N given by

$$g_N := d_R\varphi_R + \sum_{B\in\mathcal{B}_N}\sum_{T'\in B} d_{T'}\phi_{T'}. \tag{185}$$

This completes the construction of the algorithm.

Theorem (2.2.27)[24]: For each $f \in B_p^s(\varphi)$ and each integer $N \geq 1$, there is a function $f_N \in \mathcal{L}_\varphi(N)$ such that

$$\|f - f_N\|_q \leq CN^{-\frac{s}{n}}\|f\|_{B_p^s(\varphi)}, \tag{186}$$

where C is a constant depending only on φ and $p^* := \min(1, p)$.

Proof. Suppose g satisfies the assumptions of Proposition (2.2.25). In particular,

$$g = \sum_{T'\in\mathcal{V}_R(\gamma)} d_{T'}\varphi_{T'}, \tag{187}$$

where the coefficients are such that

$$\nu(g) := \left\{\sum_{T'}\left(|d_{T'}||\operatorname{supp} \varphi_{T'}|^{\frac{1}{q}}\right)^p\right\}^{\frac{1}{p}} < \infty. \tag{188}$$

Furthermore, $0 < p < q \leq \infty$ are related by

$$\frac{s}{n} = \frac{1}{p} - \frac{1}{q}, \tag{189}$$

where $p \leq 1$ if $q = \infty$.

Since the assumptions of Theorem (2.2.27) are also fulfilled, we have

$$\left\|\sum_{T'\in\Omega} d_{T'}\varphi_{T'}\right\|_q \leq C\left\{\sum_{T'\in\Omega}\left(|d_{T'}||\operatorname{supp} \varphi_{T'}|^{\frac{1}{q}}\right)^p\right\}^{\frac{1}{p}} < \infty. \tag{190}$$

(see (115) and (189)).

We begin at once with the nontrivial case of $N \geq 1$ satisfying

$$N\nu(g)^p \geq 1. \tag{191}$$

Taking such N and the function $d: T' \to d_{T'}$ defined by the expansion (187) as an input of our algorithm, we have $\mathcal{I}(R) = \nu(d)^p = \nu(g)^p \geq N^{-1}$. Then condition (160) is satisfied, and the output of the algorithm gives the linear combination g_N defined by (185). By (181), we have supp $g_N \subset R$, and this proves the first assertion of Proposition (2.2.25) (see (137)).

To prove the second, we first estimate the number $M(B)$ of terms in the linear combination $\sum_{T' \in B} d_{T'} \phi_{T'}$. In accordance with (182) and (183), each term of $\phi_{T'}$ with $T' \in B$ is a linear combination of functions $\varphi_{T''}$, where T'' runs over the set

$$\tilde{B} := \{T'' \in \mathcal{V}: |\mathrm{supp}\, \varphi_{T''} \cap T_B^-| \neq 0, h(T'') = j_B\}; \tag{192}$$

recall that $j_B := h(T_B^-)$. Hence, $M(B) \leq \#\tilde{B}$; the subsets of $\tilde{B}$, in turn, are colored by at most $\chi(\varphi)$ colors, and those of the same color (and height) do not essentially intersect. Therefore, $\#B \leq \chi(\varphi)$. Together with (176), this implies that the number of terms in (185) is at most $1 + (\#\mathcal{B}_N)\chi(\varphi) \leq CN\nu(d)^p$. It remains to estimate the error function $g - g_N$. For this, we put

$$F(\Omega) := \sum_{T' \in \Omega} d_{T'} \varphi_{T'}, \Omega \subset \mathcal{V}_R(\gamma). \tag{193}$$

For the set $\mathcal{V}_R(\gamma; T') := \{T'' \in \mathcal{V}_R(\gamma): T'' \subset T'\}$ with $T' \in \mathcal{V}_R(\gamma)$, we simplify this notation by putting

$$F(T') := F(\mathcal{V}_R(\gamma; T')). \tag{194}$$

Since $\mathcal{B}_N$ is a partition of $G_N \backslash \{R\}$, the error function can be written as

$$g - g_N = \sum_{B \in \mathcal{B}_N} F^*(B) + F(G_N^c), \tag{195}$$

where $G_N^c := \mathcal{V}_R(\gamma) \backslash G_N$ and

$$F^*(B) := F(B) - \sum_{T' \in \mathrm{B}} d_{T'} \phi_{T'}. \tag{196}$$

Taking the L_q-norm in (195), we obtain

$$\|g - g_N\|_q \leq C(J_1 + J_2), \tag{197}$$

where C depends on $q^* := \min(1, q)$, and the J_k are given by

$$J_1 := \left\| \sum_{B \in \mathcal{B}_N} F^*(B) \right\|_q, J_2 := \|F(G_N^c)\|_q.$$

If we prove that $J_k \leq CN^{-\frac{s}{d}} \nu(g), k = 1,2$, then the desired inequality (140) and Proposition (2.2.25) will be established.

In order to estimate J_1, we show that for distinct paths B, B' in $\mathcal{B}_N$ we have

$$|\operatorname{supp} F^*(B) \cap \operatorname{supp} F^*(B')| = 0. \tag{198}$$

First, suppose that the heads T_B^+ and $T_{B'}^+$ are essentially disjoint sets. Since

$$F^*(B) = \sum_{T' \in \mathrm{B}} d_{T'}(\varphi_{T'} - \phi_{T'}), \tag{199}$$

the embedding(184) implies that

$$\operatorname{supp} F^*(B) \subset T_B^+ \backslash T_B^-, \tag{200}$$

and a similar inclusion is true for the second support. Hence,(198) is fulfilled in this case. Now, if T_B^+ and $T_{B'}^+$ essentially intersect, then the head of one, say $T_{B'}^+$, embeds into the tail $T_B^-$of the other. Indeed, the basic paths $B$and$B'$are disjoint parts of a long path $A_j \in \mathcal{A}$ in this case. Consequently, $\operatorname{supp} F^*(B') \subset T_{B'}^+ \subset T_B^-$, while, by(200), the second support is a subset of $T_B^+ \backslash T_B^-$. Thus,(198) is fulfilled in this case as well.

Applying(198), we get$J_1 = \{\sum_{B \in B_N} \|F^*(B)\|_q^q\}^{\frac{1}{q}}$. We show that

$$\|F^*(B)\|_q \leq CN^{\frac{1}{p}}; \tag{201}$$

combined with(176) and(189), this yields the required estimate ofJ_1:

$$J_1 \leq CN^{\frac{1}{p}} \nu(g)(\#\mathcal{B}_N)^{\frac{1}{q}} \leq CN^{-\frac{s}{n}} \nu(g). \tag{202}$$

To prove(201), note that $\phi_{T'} = \varphi_{T_B^-}ifT' = T_B^-$ (see(182)). Hence, the vertex T' in(199) runs over the set $B^- := B\{T_B^-\}$, and we can write

$$\|F^*(B)\|_q \leq C(q) \left\{ \|F(B^-)\|_q + \left\| \sum_{T' \in B^-} d_{T'} \phi_{T'} \right\|_q \right\}. \tag{203}$$

We bound the right-hand side of(203) by $C\mathcal{J}(B^-)^{\frac{1}{p}}$. Since $B^- := (T_B^-, T_B^+]$, assertion(175) implies that $\mathcal{J}(B^-) \leq N^{-1}$. Therefore, the above bound for(203) gives the required inequality(201) and proves estimate(202) for J_1. In the subsequent considerations, the embedding(190) will be used in the equivalent form involving the cost function$\mathcal{J}$ (see(157)):

$$\left\| \sum_{T' \in \Omega} d_{T'} \phi_{T'} \right\|_q \leq c\mathcal{J}(\Omega)^{\frac{1}{p}}. \tag{204}$$

This immediately yields

$$\|F(B^-)\|_q \leq c\mathcal{J}(B^-)^{\frac{1}{p}}. \tag{205}$$

To estimate the second term in(203), we use(182) and(192) to write

$$\sum_{T' \in B^-} d_{T'} \phi_{T'} = \sum_{T'' \in \tilde{B}} \left(\sum_{T' \in B^-} m(T', T'') d_{T'} \right) \varphi_{T''}. \tag{206}$$

By Proposition(2.2.13), we have $\left|m\left(T',T''\right)\right| \leq \|m^*\|_\infty < \infty$, and it has already been proved that $\#\tilde{B} \leq \chi(\varphi)$. Therefore, application of the embedding(205) to the right-hand side of(206) yields

$$\left\|\sum_{T'\in B^-} d_{T'}\phi_{T'}\right\|_q \leq C\|m^*\|_\infty\chi(\varphi)\left\{\sum_{T'\in B^-}|d_{T'}|\right\}\left|\operatorname{supp}\varphi_{T_B^-}\right|^{\frac{1}{q}}. \tag{207}$$

Here we have taken into account the fact that $\left|\operatorname{supp}\varphi_{T''}\right| = \left|\operatorname{supp}\varphi_{T_B^-}\right|$ if $T'' \in \tilde{B}$. Now, suppose that $p \leq 1$ (hence,$q = \infty$). Applying Jenssen's inequality, we bound the right-hand side of(207) by $\{\sum_{T'\in B^-}|d_{T'}|^p\}^{\frac{1}{p}} = C\mathcal{J}(B^-)^{\frac{1}{p}}$. For $p \leq 1$, this and (203)imply the required inequality

$$\|F^*(B)\|_q \leq C\mathcal{J}(B^-)^{\frac{1}{p}}. \tag{208}$$

If $p > 1$ (and $q < \infty$), we present the path B as a sequence $T_1 \subset T_2 \subset\cdots\subset T_\ell$ with $T_1 = T_B^-$ and $T_\ell = T_B^+$. The heights of consequent vertices of B differ at least by 1. Therefore,

$$\left|\operatorname{supp}\varphi_{T_{i+1}}\right|/\left|\operatorname{supp}\varphi_{T_i}\right| \geq |\det A| =: a.$$

This and Hölder's inequality lead to the following bound for the right-hand side of(207):

$$C\left\{\sum_{i>1}\left|d_{T_i}\right|\left|\operatorname{supp}\varphi_{T_i}\right|^{\frac{1}{q}}.a^{-\frac{1}{q}}\right\} \leq C\left\{\sum_{i>1}\left(\left|d_{T_i}\right|\left|\operatorname{supp}\varphi_{T_i}\right|^{\frac{1}{q}}\right)^p\right\}^{\frac{1}{p}}\left\{\sum_{i>1}a^{-\frac{ip'}{q}}\right\}^{\frac{1}{p'}}$$
$$\leq C\mathcal{J}(B^-)^{\frac{1}{p}},$$

Thus,(208) is also true in this case.

It remains to estimate the term J_2 in(197) in a similar way. For this, we use a lemma, the proof of which will be presented later on.

Now, we derive the desired bound for $J_2 := \|F(G_N^c)\|_q$. Assertion(b) of the lemma implies that

$$\left|\operatorname{supp}F\left(S_j\right) \cap \operatorname{supp}F\left(S_{j'}\right)\right| = 0 \text{ if } i \neq j' \text{ and } j,j' < \ell.$$

Together with assertion(a), this yields

$$J_2 \leq C\left(\sum_j\left\|F\left(S_j\right)\right\|_q^q\right)^{\frac{1}{q}}.$$

Next,(204) and(209) imply the inequality

$$\left\|F\left(S_j\right)\right\|_q \leq C\mathcal{J}\left(S_j\right)^{\frac{1}{p}} \leq CN^{-\frac{1}{p}}.$$

Combining this with(210), we obtain

$$J_2 \le CN^{\frac{1}{q}-\frac{1}{p}}\nu(g) = CN^{-\frac{s}{n}}\nu(g).$$

Thus, Proposition(2.2.25) is proved.

Lemma (2.2.28)[24]: There is a collection $S \coloneqq \{S_j: 1 \le j \le \ell\}$of subsets of $G_N^c$with the following properties:

(a) S is a partition of G_N^c;

(b) if T and T'' belong to distinct subsets of $S\backslash\{S_\ell\}$, they are essentially disjoint;

(c) for each $S_j \in S$ we have

$$\mathcal{I}(S_j) \le \frac{C}{N}; \tag{209}$$

(d) the following inequality is valid:

$$\#S \le (N+1)\nu(g)^p. \tag{210}$$

Proof. We use the following notation.

LetT'be a vertex of $Gr = (\mathcal{V}, \mathcal{E})$(it may be out of the set of vertices$\mathcal{V}_R(\gamma)$). If $\Omega$is a subset of$G_N^c \subset \mathcal{V}_R(\gamma)$, we put

$$\Omega(T') \coloneqq \{T'' \in \Omega: T'' \subset T'\}. \tag{211}$$

Note that T'may fail to belong to this set, and that

$$\Omega(R) = \Omega. \tag{212}$$

Now we define the first element$S_1$of$S$. For this, we introduce the set

$$\Omega_1 \coloneqq \{T' \in \mathcal{V}: \mathcal{I}(G_N^c(T')) \ge N^{-1}\}. \tag{213}$$

Since, for$T' \in \mathcal{V}_R(\gamma)$,

$$\mathcal{I}(G_N^c(T')) \le \mathcal{I}(T') \to 0 \text{as } |T'| \to 0$$

(see(157) and (159)), the set introduced above is empty or finite. In the former case we obtain the required partition of G_N^c by putting

$$S_1 \coloneqq G_N^c \text{and} S \coloneqq \{S_1\}.$$

Since R is not an element of the (empty) set Ω_1, we have

$$\mathcal{I}(S_1) = \mathcal{I}(T') < N^{-1}$$

(see(212) and(213)).

Now, suppose $\Omega_1 \ne \phi$. Then there is an element $T_1 \in \Omega_1$of the largest (finite) height, say $h(T_1) \coloneqq j_{\max}$. Since

$$\mathcal{I}(G_N^c(T')) \le \mathcal{I}(T') < N^{-1} \tag{214}$$

(see(159)), T_1 does not belong to G_N^c. It follows that

$$\Omega_1(T_1) \subset \bigcup_{T' \in E(T_1)} \Omega_1(T'), \tag{215}$$

where the set of indices is given by

$$E(T_1) := \{T' \in \mathcal{V} : |T' \cap T_1| \neq 0 \text{ and } h(T') = j_{\max} + 1\}. \tag{216}$$

Indeed, choose T'' in $\Omega_1(T_1)$ and show that $T'' \in \Omega_1(T')$ for a suitable $T' \in E(T_1)$. Let T''' be the ancestor of T'' in the digraph $Gr = (\mathcal{V}, \mathcal{E})$ sharing its height with T_1. Since $h(T'') > h(T_1) = h(T'''')$, the set T'' is a subset of a child T' of T'''. Since T'' is also a subset of T_1, we have

$$|T_1 \cap T'| > |T''| > 0 \text{and} h(T') = h(T''') + 1 = j_{\max} + 1.$$

Thus, $T' \in E(T_1)$, and T'' embeds in T' and is an element of $\Omega_1(T_1) \subset \Omega_1$. Then, by the definition(211), T'' belongs to $\Omega_1(T')$, which proves(215).

By the maximality of $h(T_1)$, we have $\mathcal{I}(\Omega_1(T_1)) < N^{-1}$ for each $T' \in E(T_1)$ (see(213)); hence,

$$\mathcal{I}(\Omega_1(T_1)) \leq \sum_{T' \in E(T_1)} \mathcal{I}(\Omega_1(T')) < \frac{\#E(T_1)}{N}.$$

To estimate the cardinality of $E(T_1)$, observe that its subsets are colored in at most $\chi(\varphi)$ colors, and that distinct subsets of the same color and height are essentially disjoint.

Applying the result of [31], we obtain $\#E(T_1) \leq C(n)\chi(\varphi)$. If we define $S_1 \in S$ by $S_1 := \Omega_1(T_1)$, then the last two inequalities yield $\mathcal{I}(S_1) \leq \frac{C}{N}$, i.e., S_1 satisfies (209).

To introduce the next element of the collection S, we put $\Omega_2 := G_N^c \backslash S_1$ and considerthe set $\Omega_3 := \{T' \in \mathcal{V}; \Omega_2(T') \geq N^{-1}\}$. If this set is empty, then $S_2 := \Omega_2$ and $S := \{S_1, S_2\}$. Since $R \notin \Omega_3 \neq \emptyset$, we have $\mathcal{I}(S_2) = \mathcal{I}(\Omega_2(R)) < N^{-1}$, and (209) is fulfilled for S.

Now, suppose that the finite set Ω_3 is not empty, and let T_2 be its vertex of maximal height. Then, as before,

$$\mathcal{I}\Omega_2(T_2) < \frac{\#E(T_2)}{N} \leq \frac{C}{N},$$

and we put $S_2 := \Omega_2(T_2)$. Again,(209) is true for S_2. Moreover, by definition, the S_j satisfy

$$\mathcal{I}(S_j) := \mathcal{I}\left(\Omega_j(T_j)\right) \geq N^{-1}, j = 1,2.$$

Proceeding in this way, we arrive at the partition $S\{S_j : 1 \leq j \leq \ell\}$ of G_N^c satisfying condition(209). We have $S_j := \Omega_j(T_j), 1 \leq j < \ell$, whence

$$\mathcal{I}(S_j) \geq N^{-1}, 1 \leq j < \ell.$$

This implies the inequality

$$(\ell-1)N^{-1} \le \sum_{j=1}^{\ell-1} \mathcal{I}(S_j) \le \mathcal{I}(G_N^c) \le \nu(g)^p,$$

and combining this with(191), we obtain(210).

It remains to check assertion(b). Note that S is a partition of $G_N^c \subset \mathcal{V}_R(\gamma)$, and the subsets of $\mathcal{V}_R(\gamma)$are either essentially disjoint, or one is a subset of the other. Therefore, we must show that the latter is impossible for T', T''belonging to distinct S_j with $j < \ell$.But if T' and T'' with $T' \subset T''$ belong to distinct collections $S_j \coloneqq \Omega_j(T_j)$ with $j < \ell$, then T' belongs to their intersection (see(209)), which is empty. This contradiction proves(b).

The proof of Proposition(2.2.25) (and the main theorem) is complete.

Theorem (2.2.29)[24]: For each $N \ge 1$ and each $f \in B_p^{s\infty}(\varphi)$, there is$f_N \in \mathcal{L}_\varphi(N)$ such that

$$\|f - f_N\|_q \le CN^{-\frac{s}{n}}\|f\|_{B_p^{s\infty}(\varphi)} \tag{217}$$

with $C = C(\varphi, p^*)$.

Proof. Theorem (2.2.29) given $f \in B_p^{s\infty}(\varphi)$with

$$\frac{s}{n} > \frac{1}{p} - \frac{1}{q} > 0 \tag{218}$$

and$N \ge 1$, we must find $f_N \in \mathcal{L}_\varphi(M)$such that

$$\|f - f_N\|_q \le CN^{-\frac{s}{n}}\|f\|_{B_p^{s\infty}(\varphi)} \tag{219}$$

and, moreover,

$$M \le CN. \tag{220}$$

As in the proof of Theorem(2.2.27) (see Proposition(2.2.24)), we reduce the required result to the case of f having a representation

$$f = \sum_{j,k} c_{jk}\varphi_{jk}, \tag{221}$$

wherej, k run over the set

$$\{j \in \mathbb{Z}_+, k \in \mathbb{Z}^n : \operatorname{supp}\varphi_{jk} \subset \operatorname{supp}\varphi\}. \tag{222}$$

Furthermore, we may assume that, uniformly in f,

$$\|f\|_{B_p^{s\infty}(\varphi)} \approx \sup_{j,k}|c_{jk}||\operatorname{supp}\varphi_{jk}|^\mu \tag{223}$$

with$\mu \coloneqq \frac{1}{p} - \frac{s}{n}$(see(77)).

By Proposition(2.2.19), putting $a \coloneqq |\det A|^{\frac{1}{n}}$, we have

$$\sup_{j\geq 0} a^{js}E_j(f) \approx \sup_{j,k}\left|c_{jk}\left|\operatorname{supp}\varphi_{jk}\right|^{\mu}\right|, \tag{224}$$

provided f is of the form(221) with j, k in the set(222). Because of this choice of$f$, the best approximation$E_j(f)$is now the distance in $L_p(\mathbb{R}^n)$from$f$ to the linear span$F_j^0 := \operatorname{span}\{\varphi_{jk}: \operatorname{supp}\varphi_{jk} \subset \operatorname{supp}\varphi\}$. Clearly, the dimension of this space is bounded by $C(n)m_j(\varphi)\frac{|\operatorname{supp}\varphi|}{|\operatorname{supp}\varphi|} = C(n)m_j(\varphi)a^{jn}$, where $m_j(\varphi)$is the multiplicity of the family of $\operatorname{supp}\varphi_{jk} \subset \operatorname{supp}\varphi$(see Lemma(2.2.20)). Since $m_j(\varphi) = m_0(\varphi)$, we have

$$\dim F_j^0 \leq Ca^{jn}, j \in \mathbb{Z}_+. \tag{225}$$

Letf_j be an optimal element of F_j^0, i.e.,

$$E_j(f) = \|f - f_j\|_p. \tag{226}$$

We choose $J \in \mathbb{Z}_+$such that

$$a^{Jn} \leq N < a^{(J+1)n} \tag{227}$$

and then set$g_J := f - f_J$. By this definition,

$$E_j(g_J) = \begin{cases} E_J(f) \text{ if} j \leq J, \\ E_j(f) \text{ if} j > J. \end{cases} \tag{228}$$

Let σ be defined by

$$\frac{\sigma}{n} := \frac{1}{p} - \frac{1}{q}. \tag{229}$$

By(218),$\sigma < s$, so that $B_p^{s\infty}(\varphi) \subset B_p^{\sigma}(\varphi)$. Applying Theorem(2.2.27) in combination with Proposition (2.2.19) to $g_J \in B_p^{\sigma}(\varphi)$, we obtain

$$\|g_J - g_{N,J}\|_q \leq CN^{-\frac{\sigma}{n}}\left\{\sum_{j\in\mathbb{Z}_+}\left(a^{i\sigma}E_j(p)\right)^p\right\}^{\frac{1}{p}}$$

with a suitable approximant $g_{N,J}$belonging to$\mathcal{L}_{\varphi}(N)$.

Putting $f_N := f_J + g_{J,N}$ and using(228), we rewrite this as

$$\|f - f_N\|_q \leq CN^{-\frac{\sigma}{n}}\left\{\sum_{j\geq J}\left(a^{j\sigma}E_j(f)\right)^p\right\}^{\frac{1}{p}}.$$

Since the sum on the right-hand side is bounded by $Ca^{-J(s-\sigma)}\sup\_j \geq 0a^{js}E_j(f)$, and this quantity, in turn, is less than $CN^{-\frac{s-\sigma}{n}}\|f\|_{B_p^{s\infty}(\varphi)}$ by (223), (224), and(227), we obtain the required estimate(219). To complete the proof, it remains to use (225) and(227) to conclude that the number M of terms in the linear combination$f_N$is at most$N + Ca^{Jn} \leq CN$. So,(220) is also true.

Corollary (2.2.30)[24]: Under the assumptions of Theorem(2.2.27), but with$p < 1$ if$q = \infty$, the inequality

$$\left\{\sum_{N\geq 1}\left(N^{\frac{s}{n}} \in \mathcal{E}_N\big(f; L_q\big)\right)^p N^{-1}\right\}^{\frac{1}{p}} \leq C\|f\|_{B_p^s(\varphi)} \tag{230}$$

is true with $C = C(\varphi, p^*)$.

Proof. We deduce the claim of this corollary from Theorem(2.2.27) by real interpolation. For this, we use an interpolation theorem for the approximation scale $\mathcal{A}_p^{s\theta}(\mathcal{F})$introduced by (109). This result was proved for regular refinable functions in [39]. The general case is derived by the same argument. Thus, the following is true:

$$\left(\mathcal{A}_{p_0}^{s_0\theta_0}(\mathcal{F}), \mathcal{A}_{p_1}^{s_1\theta_1}(\mathcal{F})\right)_{\lambda_p} = \mathcal{A}_p^{sp}(\mathcal{F}), \tag{231}$$

where$s{:} = (1-\lambda)s_0 + \lambda s_1$and

$$\frac{1}{p} := \frac{1-\lambda}{p_0} + \frac{\lambda}{p_1} = \frac{1-\lambda}{\theta_0} + \frac{\lambda}{\theta_1}.$$

Proposition(2.2.19) allows us to rewrite this as

$$\left(B_{p_0}^{s_0\theta_0}(\varphi), B_{p_1}^{s_1\theta_1}(\varphi)\right)_{\lambda_p} = B_p^s(\varphi). \tag{232}$$

Now we introduce a new scale of approximation spaces $\mathcal{N}_p^{s\theta}(\varphi)$ determined by the approximation family$\{\mathcal{L}_\varphi(N){:}\, N \in \mathbb{N}\}$(see(89)and(91)). Thus,

$$\|f\|_{\mathcal{N}_p^{s\theta}(\varphi)} := \left\{\sum_{N\geq 1} N^{-1}\left(N^{\frac{s}{n}}\mathcal{E}_N\big(f; L_q\big)\right)^\theta\right\}^{\frac{1}{p}}.$$

In these terms, Theorem(2.2.27) can be rewritten as the embedding

$$B_p^s(\varphi) \subset \mathcal{N}_q^{s\infty}(\varphi), \tag{233}$$

Where

$$\frac{s}{n} = \frac{1}{p} - \frac{1}{q}, \text{and } p \leq 1 \text{ if } q = \infty. \tag{234}$$

Let $p > 1$, and let p, q, s satisfy(234). We choose $p_0, p_1 > 1$ so close to p that $\frac{1}{p} = \frac{1-\lambda}{p_0} + \frac{\lambda}{p_1}$ for suitable $0 < \lambda < 1$, and that $\frac{s_i}{n} := \frac{1}{p_i} - \frac{1}{q}$, $i = 0{,}1$, be strictly positive. By(234), $s = (1-\lambda)s_0 + \lambda s_1$. Applying(233) with these s_i, p_i and $q, i = 0{,}1$, and then using(232), we obtain

$$B_p^s(\varphi) \subset \left(\mathcal{N}_q^{s_0\infty}(\varphi), \mathcal{N}_q^{s_1\infty}(\varphi)\right)_{\lambda p}.$$

By the Peetre–Sparr theorem (see, e.g., [32]), the right-hand side is equal to$\mathcal{N}_q^{sp}(\varphi)$, i.e.,

$$B_p^s(\varphi) \subset \mathcal{N}_q^{sp}(\varphi).$$

Recalling the definition of the (quasi)norm of $\mathcal{N}_q^{sp}(\varphi)$, we obtain the required inequality

$$\left\{\sum_{N\geq 1}\left(N^{\frac{s}{n}}\mathcal{E}_N(f;L_q)\right)^p N^{-1}\right\}^{\frac{1}{p}} \leq C\|f\|_{B_p^s(\varphi)}$$

in the case where $p > 1$.

In the case of $p < 1$ the proof is similar.

Here we summarize the conclusions presented in Remarks(2.2.16) and(2.2.23). The dilation A is now diagonalizable with integral eigenvalues$M_i > 1$ and with eigenvectors ei forming a basis of$\mathbb{R}^n$. We assume that$\varphi \in W_\infty^{\bar{\ell},A}(\mathbb{R}^n)$; here

$$\|f\|_{W_\infty^{\bar{\ell},A}(\mathbb{R}^n)} \coloneqq \sum_{i=1}^{n}\sum_{k_i=0}^{\ell_i}\left\|D_i^{k_i}f\right\|_q, \tag{235}$$

where$1 \leq q \leq \infty$,$\bar{\ell} \in \mathbb{Z}_+^n$, and$D_i$ stands for the derivative in the direction e_i. Note that φ with such A is colorable (see Example(2.2.7)). We set

$$\sigma \coloneqq \langle\bar{\ell}\rangle \coloneqq \left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{\ell_i}\right)^{-1}$$

and assume that$\bar{\ell}$and A are related by

$$\frac{\ell_i}{\langle\bar{\ell}\rangle} = \frac{\log|\det A|}{n\log M_i}, 1 \leq i \leq n. \tag{236}$$

Changing the proof of Theorem(2.2.27) in only one point, namely, replacing the L_q-norm in inequality(108) by the norm(235) (see Remark(2.2.23)), we arrive in this setup at the next result.

Theorem (2.2.31)[24]: Under the above assumptions on the stable refinable φ and on A and$\bar{\ell}$, the following is true.

Suppose that $s > 0, \bar{\ell}$, and$0 < p < q \leq \infty$satisfy

$$\frac{s-\langle\bar{\ell}\rangle}{n} = \frac{1}{p} - \frac{1}{q} \tag{237}$$

and that $p \leq 1$ if $q = \infty$.

Then for each integer $N \geq 1$ and $f \in B_p^s(\varphi)$there is an approximant $f_N \in \mathcal{L}_\varphi(N)$ such that

$$\|f - f_N\|_{W_q^{\bar{\ell},A}(\mathbb{R}^n)} \leq CN^{-\frac{s-\langle\bar{\ell}\rangle}{n}}\|f\|_{B_p^s(\varphi)} \tag{238}$$

with C independent of f and$N$.

If A is isotropic, i.e., for $M_i = M, 1 \leq i \leq n$, the assumption (236) is clearly true. In this case, the space $W_q^{\bar{\ell},A}$ can be replaced by the Sobolev space W_q^{ℓ}.

In the case of the dilation A; the result will be presented in a forthcoming. Hence, in this case we have

$$B_p^s(\varphi) = B_p^{\bar{s}}(\mathbb{R}^n),$$

where $s = (s_1, \dots, s_n)$ is defined by $s_i := \frac{\log|\det A|}{n \log M_i}$, $1 \leq i \leq n$, and $B_p^{\bar{s}}(\mathbb{R}^n)$ is the standard anisotropic Besov space determined by the partial moduli of continuity of orders $k_i > s_i$. Note that $s = \langle \bar{s} \rangle := \left(\frac{1}{n}\sum_{i=1}^{n}\frac{1}{s_i}\right)^{-1}$; therefore, inequality (238) can be rewritten as

$$\|f - f_N\|_{W_q^{\bar{\ell},A}(\mathbb{R}^n)} \leq CN^{-\frac{\langle \bar{s} \rangle - \langle \bar{\ell} \rangle}{n}} \|f\|_{B_p^{\bar{s}}(\mathbb{R}^n)}.$$

Finally, in the case where $1 < q < \infty$, application of the Mikhlin–Hörmander multiplier theorem allows us to replace $W_q^{\bar{\ell},A}(\mathbb{R}^n)$ by the standard anisotropic Sobolev space $W_q^{\bar{\ell}}(\mathbb{R}^n)$.

The method of the proof remains valid for vector-valued φ ϕ. In this case $\varphi: \mathbb{R}^n \to \mathbb{R}^{\ell}$ is a bounded nontrivial solution of the equation

$$\varphi(x) = \sum_k m(k)\varphi(x - k)$$

with finite mask $m: \mathbb{R}^n \to M_{\ell}(\mathbb{R})$, where the target space is the linear space of real matrices of size $\ell \times \ell$. The definitions of the stability and colorability of φ requires trivial modifications, while the decomposition of $f \in L_p(\mathbb{R}^n)$, $0 < p < \infty$, that was used to introduce $B_p^s(\varphi)$, is now written as

$$f = \sum_{j,k} a_{jk} \cdot \varphi_{jk},$$

where the a_{jk} are vectors in $\mathbb{R}^{\ell}$ and $x \cdot y$ is the scalar product in this space. The following example shows that the case under consideration includes the piecewise polynomial extension of the Birman–Solomyak result ([33]).

Let $\vec{p} := (p_1, \dots, p_{\ell})$ with $\ell = \ell(k, n)$ be a vector-valued function on $\mathbb{R}^n$ whose components form a basis in the space $\mathcal{P}_k(\mathbb{R}^n)$ of polynomials in $x_1, \dots, x_n$ of degree $k - 1$. It is easily seen that the function

$$\vec{\varphi} = 1_{[0,1]^n} \cdot \vec{p}$$

satisfies the scaling equation

$$\vec{\varphi}(x) = \frac{1}{2^n} \sum_{k \in \{0,1\}^n} m(k)\vec{\varphi}(2x - k),$$

where $m(k)$ is the $(\ell \times \ell)$-matrix representing the operator $x \longmapsto \frac{1}{2}(x + k)$ in the basis $\{p_1, \dots, p_{\ell}\}$ of the space $\mathcal{P}_k(\mathbb{R}^n)$.

It is clear that $\vec{\varphi}$ is stable and colorable. Consequently, in this case the analog of Theorem(2.2.29) states that for $f \in B_p^s(\varphi)$there is a piecewise polynomial $f_N = \sum_{Q\in\pi} p_Q 1_Q$ of degree $k-1$, where π is an N-term collection of dyadic subcubes of $[0,1]^n$, that approximates f with the approximation rate$O\left(N^{-\frac{s}{n}}\right)$.

Let φ be an (A,m)-refinable function. In accordance with Definition(2.2.9), its colorability depends on the existence of a spatially colorable $T(A,\mathcal{D})$ with $\mathcal{D} \supset \operatorname{supp} m$. Therefore, a crucial point is to find a fairly large class of such sets. Here we introduce two such classes; weuse methods of coloring related to geometric and algebraic properties of the data(A,m).

Let $T:=T(A,\mathcal{D})$ be a tile (see Example(2.2.1)). In this case the digraph$Gr(A,\mathcal{D})$is a tree rooted at$T$, and its structure comes from the set inclusion order. Therefore,$Gr(A,\mathcal{D})$ is spatially colorable by a single color, i.e., $\chi(A,\mathcal{D})=1$.

Now, let $T:=T(A,\mathcal{D})$ be a semitile (see Example(2.2.6)). Thus, each pair of vertices in $\mathcal{V}:=\mathcal{V}(A,\mathcal{D})$ with heights differing by one is either (essentially) disjoint, or the smaller is a subset of the larger. Let $\mu_j$be the multiplicity of the family$\{T_{jk}: k \in \mathbb{Z}^n\}$. Then, by a change of variable, we obtain

$$\mu_j \coloneqq \operatorname*{ess\,sup}_x \left\{\sum_{k\in\mathbb{Z}^n} 1_T\left(A^j x - k\right)\right\} = \mu_0 < \infty.$$

By [31], this family is a union of at most $C(n)\mu_0$disjoint subfamilies. In other words, this family can be colored in at most $C(n)\mu_0$colors in such a way that the subsets of the same color be disjoint. Using this coloring for each level $\{T_{jk} \in \mathcal{V}\}$of height j, we obtain the required result.

The same approach shows that $Gr(A,\mathcal{D})$ is spatially colorable under the following weaker assumption: for every $T', T'' \in \mathcal{V}(A,\mathcal{D})$ with heights differing by a fixed $j_0 \geq 0$ $T'$and$T''$are either disjoint, or the smaller is a subset of the larger.

In this case we use the previous set of colors, say Γ, to turn it into a new one, defined as$\Gamma \times \{0,1,\dots,j_0-1\}$.

It can be shown that the class of digraphs introduced below can also be spatially colored in this way. However, we present another method, which yields an efficient estimate for $\chi(A,\mathcal{D})$.

Let Π be the parallelotope of Example(2.2.2) (see(60)). Thus, $\Pi \coloneqq B(\prod_{i=1}^n [0,N_i])$, where $B \in M_n(\mathbb{Z})$ is a unimodular matrix. Then$\Pi = T(A,\mathcal{D})$, where

$$A \coloneqq B \operatorname{diag}(M_1,\dots,M_n)\, B^{-1} \text{ and} \mathcal{D} \coloneqq B\left(\prod_{i=1}^n J_i\right) \cap \mathbb{Z}^n.$$

We recall that $N_i \geq 1$, $M_i \geq 2$ are integers and $J_i \coloneqq [0,(M_i-1)N_i]$.

Proposition (2.2.32)[24]: Assume that the greatest common divisor of$M_i$and$N_i$satisfies

$$(M_i, N_i) = 1\,, 1 \le\ i \le n.$$

Then

$$\chi(A, \mathcal{D}) \le N_1 \ldots N_n.$$

Proof: We begin with the following result, which is a straightforward consequence of the definitions.

Lemma (2.2.33)[24]: (a) Let $T(A_i, \mathcal{D}_i)$,$1 \le i \le m$, be a family of self-affine sets, and let

$$A := \operatorname{diag}(A_1, \ldots, A_m)\,, \mathcal{D} := \prod_{i=1}^{m} \mathcal{D}_i.$$

Then

$$\chi(\mathrm{A}, \mathcal{D}) \le \prod_{i=1}^{m} \chi(A_i, \mathcal{D}_i).$$

(b) Let $B \in M_n(\mathbb{Z})$ be unimodular. Then

$$\chi(BAB^{-1}, B\mathcal{D}) = \chi(A, \mathcal{D}).$$

Using this, we reduce the proof of the general result to the following lemma, which was proved for $M = 2$ in [40].

Lemma (2.2.34)[24]: Denote$\Pi := [0, N]$, $A := [M]$, and$\mathcal{D} := [0, (M-1)N] \cap \mathbb{Z}$, and let $N \ge 1$ and $M \ge 2$ be integers. Assume that

$$(N, M) = 1\,. \tag{239}$$

Then $\Pi = T(A, \mathcal{D})$ is spatially colorable in at most N colors.

Proof: By the Gauss lemma, each $k \in \mathbb{Z}$ has a unique representation $k = Mk' + N\ell$ with $k' \in \mathbb{Z}$ and$\ell \in \{0,1, \ldots, M-1\}$. Applying this to k' and so on, for $x \coloneqq M^{-j}k$with$k \in \mathbb{Z}$,$j \in \mathbb{Z}_+$, we obtain aunique representation

$$x = k_0(x) + N \sum_{i=1}^{j} \ell_i(x) M^{-i}$$

with$k_0(x) \in \mathbb{Z}$ and$\ell_i(x) \in \{\,0,1, \ldots, M-1\}$.

Now we define a function c on the set$\{x \coloneqq kM^{-j} : k \in \mathbb{Z}, j \in \mathbb{Z}_+\}$by

$$c(x) \equiv k_0(x)(N), c(x) \in \{0,1, \ldots, N-1\}.$$

By this definition,

$$c\big(kM^{-j}\big) = c\big(k' M^{-j}\big) \text{if and only if} k \equiv k'(N). \tag{240}$$

In this case a vertex I of the digraph$Gr(A, \mathcal{D})$ is an interval of the form $M^{-j}[\upsilon, \upsilon + N]$ with suitable $\upsilon \in \mathbb{Z}$ and$j \in \mathbb{Z}_+$; therefore, the endpoints x_I, y_I of that interval satisfy the condition $c(x_I) = c(y_I)$.We define the desired coloring of $Gr(A, \mathcal{D})$ by letting

$$c(I) := c(x_I)$$

and show that the condition of Definition(2.2.9) is satisfied. Let $I \coloneqq M^{-j}[v, v+N]$ and $I' \coloneqq M^{-j'}[v', v'+N]$be vertices of this digraph sharing the same color, and let $j \geq j'$.Consider the lattice

$$L := \{(M^{j-j'}v + Nk)M^{-j} : k \in \mathbb{Z}\}.$$

Since$c(vM^{-j}) = c(v'M^{-j'})$, the congruence$M^{j-j'}v \equiv v'(N)$is true (see(239) and(240)). Therefore, L contains all points $y \coloneqq kM^{-j}$with $k \in \mathbb{Z}$, and $c(y) = c(x_{I'})(= c(I'))$. The endpoints x_I, y_I of I are points of this type and hence belong to L. Since the length of I is equal to the step of L, and the endpoints of $I'$are also in$L$, there are only two possibilities: either $x_I, y_I \in I'$and then $I \subset I'$, or these endpoints do not belong to the interior of$I'$and$|I' \cap I| = 0$. Consequently, $Gr(A, \mathcal{D})$ is spatially colorable, and

$$\chi(A, \mathcal{D}) \leq \#\text{image}(c) = N.$$

Suppose the dilation$A$ of$\varphi$is diagonalized with eigenvaluesλ_i and eigenvectorsv_i, $1 \leq i \leq n$.

Proposition (2.2.35)[24]: If$\lambda_i \in \mathbb{Q}$,$1 \leq i \leq n$, then φ is colorable.

Proof: Let$\lambda_i \coloneqq \frac{m_i}{d}$, $m_i, d \in \mathbb{Z}$. Replacing A by a $\mathbb{Z}$-similar matrix, we may assume that $\lambda_i > 1$, so that $M_i > d \geq 1$. By our assumptions, all vectors v_i can be taken in $\mathbb{Z}^n$, and they form a basis of$\mathbb{R}^n$. Consider the parallelepiped

$$\Pi \coloneqq \left\{\sum_{i=1}^{n} c_i v_i : 0 \leq c_i \leq dN_i\right\}$$

with integers N_i to be chosen later. Then $Av_i \coloneqq \frac{m_i}{d}, v_i$, whence

$$A(\Pi) = \left\{\sum_{i=1}^{n} c_i v_i : 0 \leq c_i \leq M_i N_i\right\} = \bigcup_{d \in \mathcal{D}} (\Pi + d),$$

where the digit set $\mathcal{D}$ is given by

$$\mathcal{D} \coloneqq \left\{\sum_{i=1}^{n} c_i v_i \in \mathbb{Z}^n : c_i \in [0, (M_i - 1)N_i] \cap \mathbb{Z}\right\}$$

with$M_i \coloneqq m_i - d + 1 (\geq 1)$.

In other words, $\Pi = T(A, \mathcal{D})$. Now we choose N_i such that$(N_i, M_i) = 1$,$1 \leq i \leq n$. Then, by Proposition (2.2.32),

$$\chi(A, \mathcal{D}) \leq N_1 \cdots N_n.$$

Taking N_i sufficiently large and shifting by a suitable vector $v \in \mathbb{Z}^n$, we reduce the general statement to the case of m satisfying

$$\operatorname{supp} m \subset \mathcal{D} + \upsilon.$$

Since $T(A, \mathcal{D} + \upsilon) = T(A, \mathcal{D}) + A(I - A)^{-1}\upsilon$ (see (58)), the refinable function φ associated with (A, m) is colorable (see Definition $(2.2.9)$).

Chapter 3
Classes of Hardy Spaces and Comparison of the Classical BMO

We establish a duality theorem between the $H_L^p(\mathbb{R}^n)$ spaces and the Morrey-Campanato spaces. We obtain the boundedness of fractional integrals on $H_L^p(\mathbb{R}^n)$ and give the inclusion between the classical $H^p(\mathbb{R}^n)$ spaces and the $H_L^p(\mathbb{R}^n)$ spaces associated with operators. We obtain BMO_L estimates and interpolation results for fractional powers, purely imaginary powers and spectral multipliers of self adjoint operators. We also demonstrate that the space BMO_L might coincide with or might be essentially different from the classical BMO space.

Section (3.1): Operators with Duality Theorem and Applications

The continues a line of study in[58], [70] and[71], where a class of the Hardy spaces $H_L^1(\mathbb{R}^n)$and the $BMO_L(\mathbb{R}^n)$ spaces associated with operators were introduced and developed, and they generalize the classical Hardy space $H^1(\mathbb{R}^n)$and the BMO space. For the basic facts about the classical Hardy and BMO spaces on Euclidean spaces $\mathbb{R}^n$, see, for examples,[61], [74], [75], [19], [84] and[86].

Suppose that L is a linear operator on $L^2(\mathbb{R}^n)$ which generates an analytic semi-group e^{-tL} with a kernel $p_t(x,y)$ satisfying an upper bound, that is, there exist positive constants m and ϵ such that for all $x,y\ \in \mathbb{R}^n$ and for all $t>0$,

$$|p_t(x,y)| \le \frac{ct^{\epsilon/m}}{(t^{1/m}+|x-y|)^{n+\epsilon}}. \quad (1)$$

In[58], Auscher, Duong and McIntosh defined a Hardy space $H_L^1(\mathbb{R}^n)$ associated with the operator L as the class of all functions on $\mathbb{R}^n$for which $S_L(f) \in L^1(\mathbb{R}^n)$where

$$S_L(f)(x) = \left(\int_0^\infty \int_{|y-x|<t} \left|Q_{t_m} f(y)\right|^2 \frac{dydt}{t^{n+1}}\right)^{1/2}, \quad (2)$$

and $Q_t = tLe^{-tL}$. They then obtained a molecular chfaracterization for functions in $H_L^1(\mathbb{R}^n)$ by using the theory of tent spaces developed in [62] and [63].

A new function space $BMO_L(\mathbb{R}^n)$ associated with the operator L was introduced in[70]. We say that a function f (with suitable bounds on growth) is in $BMO_L(\mathbb{R}^n)$ if

$$\sup_B \frac{1}{|B|}\int_B |f(x) - e^{-t_B L} f(x)|\, dx < \infty, \quad (3)$$

Where $t_B = r_B^m$, and r_B is the radius of the ball B. It was proved in[71] that if L has a bounded holomorphic functional calculus on L^2 and the kernel $p_t(x,y)$ of the semi-group e^{-tL}satisfies an upper bound(1), then the space $BMO_{L^*}(\mathbb{R}^n)$is the dual space of Hardy space$H_L^1(\mathbb{R}^n)$in whichL^* denotes the adjoint operator of L.

This gives a generalization of the duality of $H^1(\mathbb{R}^n)$and $BMO(\mathbb{R}^n)$of Fefferman and Stein([74]). Indeed, a valid choice of e^{-tL}in(2) and (3) is the Poisson semigroup$e^{-t\sqrt{\Delta}}$, which is defined by

$$e^{-t\sqrt{\Delta}} f(x) = \int_{\mathbb{R}^n} p_t(x-y) f(y) dy\,, t>0, \text{where } p_t(x) = \frac{c_n t}{(t^2+|x|^2)^{(n+1)/2}}.$$

For this choice of $e^{-t\sqrt{\Delta}}$, the spaces $H_{\sqrt{\Delta}}^1(\mathbb{R}^n)$ and $BMO_{\sqrt{\Delta}}(\mathbb{R}^n)$coincide with the classical Hardy $H^1(\mathbb{R}^n)$and BMO spaces, respectively ([58] and [70]).

For more properties of the space $H_L^1(\mathbb{R}^n)$and the $BMO_L(\mathbb{R}^n)$ space, see [58], [70], [71], [66] and[65].

Our concern is to introduce a class of Hardy spaces $H_L^p(\mathbb{R}^n)$ associated with L for a range of $p\ < 1$ and study their duals.

(i) To define the space $H_L^p(\mathbb{R}^n)$for $p < 1$, we use a space $\mathfrak{L}_L(\alpha, 2, s)$ of Morrey-Campanato functions introduced in [72] that plays the role of the space S of test functions on $\mathbb{R}^n$. It turns out that given an operator L with a bounded holomorphic functional calculus in $L^2(\mathbb{R}^n)$, which generates a semigroup with upper bounds (1) on its heat kernels, the Hardy space $H_L^p(\mathbb{R}^n)$can be defined as the collection of all continuous linear functionals f on $\mathfrak{L}_{L^*}\left(\frac{1}{p}-1,2,\left[\frac{n\left(\frac{1}{p}-1\right)}{m}\right]\right)$satisfying $S_L(f) \in L^p(\mathbb{R}^n)$. Note that the spaces $H^p_{\sqrt{\Delta}}(\mathbb{R}^n)$ and $\mathfrak{L}_{\sqrt{\Delta}}(\alpha, 2, s)$ coincide with the classical $H^p(\mathbb{R}^n)$and the Morrey-Campanato spaces

$L(\alpha, 2, s)$ $(= \Lambda_{n\alpha}(\mathbb{R}^n)$of Lipschitz), respectively (see [66]).

(ii) As in[58] , we give a molecular decomposition for function f in the$H_L^p(\mathbb{R}^n)$spaces by using certain estimates on area integrals and tent spaces (see Proposition(3.1.6)).

(iii) We establish a duality theorem, which says that the dual space of $H_L^p(\mathbb{R}^n)$ is $\mathfrak{L}_{L^*}\left(\frac{1}{p}-1,2,s\right)$, by applying the results, together with some estimates of the tent spaces and Carleson measures. With a choice of $e^{-t\sqrt{\Delta}}$, gives the classical result of the duality of $H^p(\mathbb{R}^n)$and $L\left(\frac{1}{p}-1,2,s\right)$for $p < 1$ (see, for example, Theorem 2.7 of [86] and [69]).

(iv) We give applications, which include the boundedness of fractional integrals on the spaces $H_L^p(\mathbb{R}^n)$and the inclusion between the classical spaces $H^p(\mathbb{R}^n)$and the $H_L^p(\mathbb{R}^n)$spaces associated with some differential operators.

In [56] and [55], the Hardy space associated with an elliptic second-order divergence operator L was introduced by using the Poisson semigroup of L. In[73], Hardy spaces associated with Schrödinger operators were studied. In comparison with the classical $H^p(\mathbb{R}^n)$spaces, an important feature of the $H_L^p(\mathbb{R}^n)$spaces is that they tightly connect the operators considered, which may be an effective
tool in the study of singular integral operators associated with the operator$L$; see [58], [65], [68], [70] and [71].

The letter "c" will denote (possibly different) constants that are independent of the essential variables.

We start with a review of some definitions of holomorphic functional calculi introduced by McIntosh [79]. Let $0 \leq\ \omega < \nu < \pi$. We define the closed sector in the complex plane$\mathbb{C}$

$$S_\omega = \{\mathcal{Z} \in \mathbb{C}: |\arg \mathcal{Z}| \leq \omega\} \cup \{0\}$$

and denote the interior of $S_\omega$by S_ω^0.

We employ the following subspaces of the space $H(S_\nu^0)$ of all holomorphic functions on S_ν^0:

$$H_\infty(S_\nu^0) = \{b = H(S_\nu^0): \|b\|_\infty < \infty\},$$

where$\|b\|_\infty = \sup\{|b(\mathcal{Z})|: \mathcal{Z} \in S_\nu^0\}$and

$$\Psi(S_\nu^0) = \{\psi \in H(S_\nu^0): \exists s > 0, |\psi(\mathcal{Z})| \leq c|\mathcal{Z}|^s(1 + |\mathcal{Z}|^{2s})^{-1}\}.$$

Given $0 \leq \omega < \pi$, a closed operator L in $L^2(\mathbb{R}^n)$is said to be of type ω if $\sigma(L) \subset S_\omega$, and for each $\nu > \omega$, there exists a constant c_ν such that

$$\|(L - \lambda I)^{-1}\|_{2,2} \leq c_\nu |\lambda|^{-1}, \lambda \notin S_\nu.$$

If L is of type ω and $\psi \in \Psi(S_\nu^0)$, we define $\psi(L) \in L(L^2, L^2)$ by

$$\psi(L) = \frac{1}{2\pi i} \int_\Gamma (L - \lambda I)^{-1} \psi(\lambda)\, d\lambda,$$

where Γ is the contour $\{\xi = re^{\pm i\theta}: r \geq 0\}$ parametrized clockwise aroundS_ω, and $\omega < \theta < \nu$. Clearly, this integral is absolutely convergent in $\mathcal{L}(L^2, L^2)$ (which is the class of all bounded linear operators on $L^2(\mathbb{R}^n)$), and it is straightforward to show, using Cauchy's theorem, that the definition is independent of the choice of$\theta \in (\omega, \nu)$. If, in addition, L is one-one and has dense range and if $b \in H_\infty(S_\nu^0)$, then $b(L)$ can be defined by

$$b(L) = [\psi(L)]^{-1}(b\psi)(L) \text{where } \psi(\mathcal{Z}) = \mathcal{Z}(1 + \mathcal{Z})^{-2}.$$

It can be shown that $b(L)$ is a well-defined linear operator in$L^2(\mathbb{R}^n)$. We say that$L$ has a bounded H_∞-calculus in $L^2(\mathbb{R}^n)$provided there exists$c_{\nu,2} > 0$such that$b(L) \in \mathcal{L}(L^2, L^2)$ and

$$\|b(L)\|_{2,2} \leq c_{\nu,2} \|b\|_\infty, \qquad \forall b \in H_\infty(S_\nu^0).$$

An important feature of this functional calculus is the following convergence lemma.

Lemma (3.1.1)[53]: (Convergence lemma)**.** *Let X be a complex Banach space. Given*$0 \leq \omega < \nu \leq \pi$, *let L be an operator of type ω on X which is one-to-one with dense domain and range. Suppose*$\{f_\alpha\}$*is a uniformly bounded net in* $H_\infty(S_\nu^0)$, *which converges to* $f \in H_\infty(S_\nu^0)$*uniformly on compact subsets of* S_ν^0, *such that* $\{f_\alpha(L)\}$ *isa uniformly bounded net in the space* $\mathcal{L}(X, X)$ *of continuous linear operators on* X.

Then $f(L) \in \mathcal{L}(X, X)$, $f_\alpha(L)u \to f(L)u$ *for all* $u \in X$ *and*$\|f(L)\| \leq \sup_\alpha \|f_\alpha(L)\|$.

For the proof of Lemma (3.1.1), see[79] and[54].

Let L be a linear operator of type ω on $L^2(\mathbb{R}^n)$with$\omega < \pi/2$, hence L generates a holomorphic semigroup $e^{-\mathcal{Z}L}$, $0 \leq |\mathrm{Arg}(\mathcal{Z})| < \pi/2 - \omega$.

Assume the following two conditions.

Assumption$(\boldsymbol{a})$. Assume that for each $t > 0$, the distribution kernel $p_t(x, y)$ of e^{-tL} belongs to $L^\infty(\mathbb{R}^n \times \mathbb{R}^n)$and satisfies the estimate

$$|p_t(x, y)| \leq h_t(x, y)$$

for $x, y \in \mathbb{R}^n$, where $h_t(x, y)$ is given by

$$h_t(x, y) = t^{-n/m} g\left(\frac{|x, y|}{t^{1/m}}\right), \tag{4}$$

in which m is a positive constant and g is a positive, bounded, decreasing function satisfying

$$\lim_{r \to \infty} r^{n+\epsilon} g(r) = 0 \tag{5}$$

Assumption$(\boldsymbol{b})$. The operator L is one-one and has dense range in $L^2(\mathbb{R}^n)$. Also,L has a bounded H_∞-calculus in $L^2(\mathbb{R}^n)$.

Now, we give some consequences of the assumptions (a) and (b) which will be used later.

First, if $\{e^{-tL}\}_{t>0}$is a bounded analytic semigroup on$L^2(\mathbb{R}^n)$whose kernel $p_t(x, y)$ satisfies the estimates(4) and(5), then for any $k \in \mathbb{N}$, the time derivatives of p_t satisfy

$$\left| t^k \frac{\partial^k p_t(x, y)}{\partial t^k} \right| \leq \frac{c}{t^{n/m}} g\left(\frac{|x - y|}{t^{1/m}}\right) \text{for all } t > 0 \text{ and almost all } x, y \in \mathbb{R}^n \tag{6}$$

For each $k \in \mathbb{N}$, the function g might depend on k but it always satisfies (5). See Theorem 6.17 of[80] and[60].

Secondly, L has a bounded H_∞-calculus in $L^2(\mathbb{R}^n)$if and only if for any nonzero function $\psi \in \Psi(S_\nu^0)$, L satisfies the square function estimate and its reverse

$$c_1\|f\|_{L^2} \leq \left(\int_0^\infty \|\psi_t(L)\|_{L^2}^2 \frac{dt}{t}\right)^{1/2} \leq c_1\|f\|_{L^2} \tag{7}$$

for some $0 < c_1 \leq c_2 < \infty$, where $\psi_t(\xi) = \psi(t\xi)$. Note that different choices of $\nu > \omega$ and $\psi \in \Psi(S_\nu^0)$ lead to equivalent quadratic norms of f.

As noted in[79], positive self-adjoint operators satisfy the quadratic estimate(7). So do normal operators with spectra in a sector and maximal accretive operators. For definitions of these classes of operators, see [91].

We now define the class of functions that the operators e^{-tL}act upon. For any $\beta > 0$, a function $f \in L_{loc}^2(\mathbb{R}^n)$is said to be a function of β-type if f satisfies

$$\left(\int_{\mathbb{R}^n} \frac{|f(x)|^2}{1+|x|^{n+\beta}} dx\right)^{1/2} \leq c < \infty. \tag{8}$$

We denote by$\mathcal{M}_\beta$the collection of all functions of β-type. If $f \in \mathcal{M}_\beta$, the norm of$f$ in$\mathcal{M}_\beta$is denoted by

$$\|f\|_{\mathcal{M}_\beta} = \inf\{c \geq 0: \text{holds}\}.$$

It is easy to see that $\mathcal{M}_\beta$ is a Banach space under the norm$\|f\|_{\mathcal{M}_\beta}$. For any given operator L, we let

$$\theta(L) = \sup\{\epsilon > 0: \text{holds}\}. \tag{9}$$

and denote by

$$\mathcal{M} = \begin{cases} \mathcal{M}_{\theta[L]}, \text{if } \theta(L) < \infty; \\ \bigcup_{\beta:0<\beta<\infty} \mathcal{M}_\beta \ \text{ if } \theta(L) < \infty. \end{cases} \tag{10}$$

Note that if $L = \Delta$ is the Laplacian on $\mathbb{R}^n$, then $\theta(\Delta) = \infty$. When $L = \sqrt{\Delta}$, we have $\theta(\sqrt{\Delta}) = 1$.

Given an integer $s \in \mathbb{Z}^+$, for any$(x,t) \in \mathbb{R}_+^{n+1}$and for $f \in \mathcal{M}$, we denote

$$P_{s,t}f(x) = f(x) - (I - e^{-tL})^{s+1}f(x) \text{ and } Q_{s,t}f(x) = t^{s+1}L^{s+1}e^{-tL}f(x). \tag{11}$$

See[59] and[77]. In particular, if $s = 0$, we denote by

$$P_t f = P_{0,t}f = e^{-tL}f \text{and } Q_t f = Q_{0,t}f = tLe^{-tL}f. \tag{12}$$

Since $f \in \mathcal{M}$, by the estimate(6) the operators $P_{s,t}f$ and $Q_{s,t}f$ are well defined.

Moreover, the kernel $p_{s,t}(x,y)$ (resp.$q_{s,t}(x,y)$) of $P_{s,t}$(resp. $Q_{s,t}$) satisfies

$$\left|p_{s,t^m}(x,y)\right| \leq c_s t^{-n} g\left(\frac{|x-y|}{t}\right)\Bigg|. \left|q_{s,t^m}(x,y)\right| \leq c_s t^{-n} g\left(\frac{|x-y|}{t}\right)\Bigg|, \tag{13}$$

where the function g satisfies the condition(5). This property is the same as the estimate(6).

The following definition was introduced in[72], which generalizes the classical Morrey-Campanato spaces $L(\alpha,q,s)$. For the basic facts about the spaces $L(\alpha,q,s)$, see[76] and[86].

Definition (3.1.2)[53]: Suppose$0 \leq \alpha < \theta(L)/n$,$1 \leq q < \infty$ and$s \geq \left[\frac{n\alpha}{m}\right]$, the integral part of$\frac{n\alpha}{m}$. We say that $f \in \mathcal{M}$ is in$\mathfrak{L}_L(\alpha,q,s)$, the spaces of Morrey-Campanato type associated with $\{e^{-tL}\}_{t>0}$, if there exists a positive constant c such that for any ball B of $\mathbb{R}^n$,

$$\left[|B|^{-1}\int_B \left|f(x) - P_{s,t_B}f(x)\right|^q dx\right]^{1/q} \le c|B|^{\alpha}, \tag{14}$$

where $t_B = r_B^m$, and r_B equals to the radius of the ball B.

The smallest bound c satisfying condition (14) is then taken to be the norm of f in this space and is denoted by $\|f\|_{\mathfrak{L}_L(\alpha,q,s)}$.

Note that $\left(\mathfrak{L}_L(\alpha,q,s), \|\cdot\|_{\mathfrak{L}_L(\alpha,q,s)}\right)$ is a seminormed vector space, with the seminorm vanishing on the space $\mathcal{K}_{(L,s)}$, defined by

$$\mathcal{K}_{(L,s)} = \{f \in \mathcal{M}: P_{s,t}f(x) = f(x) \text{for almost all } x \in \mathbb{R}^n \text{for all } t > 0\}.$$

The $\mathfrak{L}_L(\alpha,q,s)$ space is understood to be modulo $\mathcal{K}_{(L,s)}$. See [71] for a discussion of the dimensions of $\mathcal{K}_{(L,0)}$ when L is a second order elliptic operator of divergence form or a Schrödinger operator.

Now, we give some important properties of the spaces $\mathfrak{L}_L(\alpha,q,s)$ where $0 < \alpha < \theta(L)/n$, $1 \le q < \infty$ and $s \ge \left[\frac{n\alpha}{m}\right]$.

First, for each $1 \le q < \infty$, the space $\mathfrak{L}_L(0,q,0)$ is a variant of the new BMO_L space introduced in [70], and it generalizes the classical BMO space. If $\theta(L) = \infty$, then the spaces $\mathfrak{L}_L(\alpha,q,s)$ are well defined for all $0 \le \alpha < \infty$, $1 \le q < \infty$ and $s \ge \left[\frac{n\alpha}{m}\right]$. In particular, if L is the Laplacian on $\mathbb{R}^n$, then the classical Morrey-Campanato spaces $L(\alpha,q,2s)$ coincide with our spaces $\mathfrak{L}_\Delta(\alpha,q,s)$. See [66].

Secondly, if $f \in \mathfrak{L}_L(\alpha,q,s)$ for $0 < \alpha < \theta(L)/n$, $1 \le q < \infty$ and $s \ge \left[\frac{n\alpha}{m}\right]$, then

(a) for every $t > 0$ and every $K > 1$, there exists a constant $c > 0$ such that for almost all $x \in \mathbb{R}^n$,

$$\left|P_{s,t}f(x) - P_{s,Kt}f(x)\right| \le c(Kt)^{n\alpha/m}\|f\|_{\mathfrak{L}_L(\alpha,q,s)}; \tag{15}$$

(b) For any $\delta > n\alpha$ and any $x \in \mathbb{R}^n$, there exists a constant c_δ which depends on δ such that

$$\int_{\mathbb{R}^n} \frac{\left|f(y) - P_{s,t}f(y)\right|}{(t^{1/m} + |x-y|)^{n+\delta}} dy \le c_\delta t^{(m\alpha-\delta)/m}\|f\|_{\mathfrak{L}_L(\alpha,q,s)}. \tag{16}$$

For the proofs of (a) and (b), see Propositions 2.5 and 2.7 of [72], respectively.

Proposition (3.1.3)[53]: *Let* $0 < \alpha < \theta(L)/n$, $1 \le q < \infty$ *and* $s \ge \left[\frac{n\alpha}{m}\right]$.

(i) *If* $q_t(x,y)$ *denotes the kernel of the operator* Q_t, *then for each* $y \in \mathbb{R}^n$, $q_t(\cdot,y) \in \mathfrak{L}_L(\alpha,q,s)$. *Similarly, for each* $x \in \mathbb{R}^n$, $q_t(x,\cdot) \in \mathfrak{L}_L(\alpha,q,s)$.

(ii) *If* $f \in L^2(\mathbb{R}^n)$, *then for each* $t > 0$, $Q_t f \in \mathfrak{L}_L(\alpha,q,s)$, *and also,* $Q_t^* f \in \mathfrak{L}_{L^*}(\alpha,q,s)$.

Proof: In order to prove that for each $y \in \mathbb{R}^n$, $q_t(\cdot,y) \in \mathfrak{L}_L(\alpha,q,s)$, by Definition (3.1.2) it suffices to verify that for any ball B of $\mathbb{R}^n$,

$$\int_B \left|q_t(x,y) - P_{s,t_B}\big(q_t(\cdot,y)\big)(x)\right|^q dx = \int_B \left|(I - e^{-t_B L})^{s+1}\big(q_t(\cdot,y)\big)(x)\right|^q dx \le c|B|^{q\alpha+1}, \tag{17}$$

where $t_B = r_B^m$, and r_B equals the radius of the ball B.

Let us prove (17). Noting that $I - e^{-t_B L} = -\int_0^{t_B} \frac{d}{dr} e^{-rL}\, dr = \int_0^{t_B} L e^{-rL}\, dr$, we have

$$(I-e^{-t_B L})^{s+1}(tLe^{-tL}) = \int_0^{t_B}\dots\int_0^{t_B}\frac{t}{(t+r_1+\cdots+r_{s+1})^{s+2}}$$
$$\times Q_{s+1},(t+r_1+\cdots+r_{s+1})dr_1\dots dr_{s+1}.$$

From(13), the operator$(I-e^{-t_B L})^{s+1}(tLe^{-tL})$has an associated kernel $K_{s,t}(x,y)$which satisfies

$$\begin{aligned}&\left|K_{s,t}(x,y)\right|\\ &\leq c_{s+1}\int_0^{t_B}\dots\int_0^{t_B}\frac{t}{(t+r_1+\cdots+r_{s+1})^{s+2+\frac{n}{m}}}\\ &\times g\left(\frac{|x-y|}{(t+r_1+\cdots+r_{s+1})^{1/m}}\right)dr_1\dots dr_{s+1}\\ &\leq c(s,t)\int_0^{t_B}\dots\int_0^{t_B}\frac{t}{(t+r_1+\cdots+r_{s+1})^{s+2}}dr_1\dots dr_{s+1}\\ &\leq c'(s,t)\int_0^{t_B}\frac{r^{s+1}}{(t+r)^{s+1}}\frac{dr}{r}\end{aligned}$$

for some constant $c(s,t)$ dependent on s and t. We observe that$n\alpha \leq m(s+1)$, and then $\frac{r}{t+r} \leq c(t)r^{\frac{n\alpha}{m(s+1)}}$. Therefore,

$$\left|K_{s,t}(x,y)\right| \leq c'(s,t)c(t)^{s+1}\int_0^{t_B} r^{\frac{n\alpha}{m(s+1)}.(s+1)}\frac{dr}{r} \leq c\int_0^{t_B} r^{\frac{n\alpha}{m}}\frac{dr}{r} \leq ct_B^{\frac{n\alpha}{m}} \leq c|B|^{\alpha},$$

which gives the desired estimate(17), and then$q_t(\cdot,y) \in \mathfrak{L}_L(\alpha,q,s)$for each $y \in \mathbb{R}^n$. Similarly, for each $x \in \mathbb{R}^n$,$q_t(x,\cdot) \in \mathfrak{L}_L(\alpha,q,s)$. Also for$f \in L^2(\mathbb{R}^n)$, we have that $Q_t f \in \mathfrak{L}_L(\alpha,q,s)$with$\|Q_t f\|_{\mathfrak{L}_L(\alpha,q,s)} \leq c\|f\|_2andQ_t^* f\mathfrak{L}_{L^*}(\alpha,q,s) \leq c\|f\|_2$ with $\|Q_t^* f\|_{\mathfrak{L}_{L^*}(\alpha,q,s)} \leq c\|f\|_2$.

We now introduce the dual space$\left(\mathfrak{L}_L(\alpha,q,s)\right)'$with$0<\alpha<\theta(L)/n$,$1 \leq q < \infty$and$s \geq \left[\frac{n\alpha}{m}\right]$consisting of all linear functionals$\ell$from$\mathfrak{L}_L(\alpha,q,s)$to $\mathbb{C}$with the property that there exists a finite constant c such that for all $g \in \mathfrak{L}_L(\alpha,q,s)$,

$$|\ell(g)| \leq c\|g\|_{\mathfrak{L}_L(\alpha,q,s)}. \tag{18}$$

We denote by$\langle f,g\rangle$the natural pairing of elements $f \in \left(\mathfrak{L}_L(\alpha,q,s)\right)'$and $g \in \mathfrak{L}_L(\alpha,q,s)$. It follows from Proposition(3.1.3) that for all$f \in \left(\mathfrak{L}_L(\alpha,q,s)\right)'$with$0<\alpha<\theta(L)/n$,$1 \leq q < \infty$ and$s \geq \left[\frac{n\alpha}{m}\right]$,$\langle f,q_t(.,y)\rangle$ is well defined. Similarly, for all$f \in \left(\mathfrak{L}_{L^*}(\alpha,q,s)\right)'$ with $0<\alpha<\theta(L)/n$, $1 \leq q < \infty$ and $s \geq \left[\frac{n\alpha}{m}\right]$, $\langle f,q_t(x,.)\rangle$is well defined. In the following, we will denote$Q_t f(x) = \langle f,q_t(x,.)\rangle$. Also, for any $\in \left(\mathfrak{L}_{L^*}(\alpha,q,s)\right)'$, we observe that for any$g \in L^2(\mathbb{R}^n)$,

$$|\langle Q_t f,g\rangle| = |\langle f,Q_t^* g\rangle| \leq c\|Q_t^* g\|_{\mathfrak{L}_{L^*}(\alpha,q,s)} \leq c\|g\|_2,$$

and thus$Q_t f \in L^2(\mathbb{R}^n)$. These will often be used.

In what follows,$\mathbb{R}_+^{n-1}$will denote the upper half-space in $\mathbb{R}^{n+1}$. The notation $\Gamma(x) = \{(y,t) \in \mathbb{R}_+^{n+1}: |x-y|<t\}$ denotes the standard cone (of aperture 1) with vertex $x \in \mathbb{R}^n$. For any closed subset $F \subset \mathbb{R}^n$, $\mathcal{R}(F)$ will be the union of all cones with vertices in F, i.e., $\mathcal{R}(F) = \bigcup_{x \in F}\Gamma(x)$. If O is an open subset of $\mathbb{R}^n$, then the "tent" over O, denoted by $\hat{O}$, is given as $\hat{O} = [\mathcal{R}(O^c)]^c$.

We continue with the assumption that the operator L satisfies the assumptions (a) and (b). Given $a$function$f \in L^1(\mathbb{R}^n)$, the area integral function$S_L(f)$ associated with an operator L is defined by

$$S_L(f)(x) = \left(\int_{\Gamma(x)} |Q_m f(y)|^2 \frac{dydt}{t^{n+1}}\right)^{1/2}. \quad (19)$$

It follows from the assumption (b) of L that the area integral function $S_L(f)$ is bounded on $L^2(\mathbb{R}^n)$. It was proved in Theorem 6 of[58] that there exist constants c_1, c_2 such that$0 < c_1 \leq c_2 < \infty$ and

$$c_1\|f\|_p \leq \|S_L(f)\|_p \leq c_2\|f\|_p \quad (20)$$

for all$f \in L^p(\mathbb{R}^n)$,$1 < p < \infty$. See also [67] and [90].
By duality, the operator $S_{L^*}(f)$also satisfies the estimate(20), where $L^*$is the adjoint operator of L.

The following definition was introduced in [58]. We say that $f \in L^p(\mathbb{R}^n)$ belongs to a Hardy space associated with an operator L (abbreviated as H_L^1) if $S_L(f) \in L^1(\mathbb{R}^n)$, and define its norm by

$$\|f\|_{H_L^1} = \|S_L\|_{L^1}.$$

Note that if $L = \Delta$ is the Laplacian on $\mathbb{R}^n$, then the classical space $H^1(\mathbb{R}^n)$ coincides with the spaces $H_\Delta^1(\mathbb{R}^n)$and $H_{\sqrt{\Delta}}^1(\mathbb{R}^n)$and their norms are equivalent. See[74]and[19].

For a measurable function $g(y,t)$ defined on $\mathbb{R}_+^{n-1}$, we will denote

$$\mathcal{A}(g)(x) = \left(\int_{\Gamma(x)} |g(y,t)|^2 \frac{dydt}{t^{n+1}}\right)^{1/2} \quad (21)$$

and for $0 < p \leq 1$,

$$C_p(g)(x) = \sup_{x\in B} \frac{1}{|B|^{\frac{1}{p}-\frac{1}{2}}} \left(\int_{\hat{B}} |g(y,t)|^2 \frac{dydt}{t}\right)^{1/2}. \quad (22)$$

Following[63], the "tent space" $T_2^p$is defined as the space of functions g such that $\mathcal{A}(g) \in L^p(\mathbb{R}^n)$, when $0 < p < \infty$. The resulting equivalence classes are then equipped with the norm$|||g|||_{T_2^p} = \|\mathcal{A}\|_p$. When $p = \infty$, $T_2^\infty$is the class of g for which $C_1(g) \in L^\infty(\mathbb{R}^n)$, and its norm is defined by $\|C_1(g)\|_\infty$. For $0 < p \leq 1$ we denote $T_2^{p,\infty}$by

$$T_2^{p,\infty} = \left\{g: \left\|C_p(g)\right\|_\infty < \infty\right\}.$$

Obviously, $T_2^{1,\infty} = T_2^\infty$.We observe that $f \in H_L^1(\mathbb{R}^n)$if and only if$Q_{t^m} f \in T_2^1$, i.e.,$\mathcal{A}(Q_{t^m} f) = S_L(f) \in L^1$. From this point of view, we now introduce the Hardy spaces$H_L^p(\mathbb{R}^n)$for $p < 1$associated with the semigroup$\{e^{-tL}\}_{t>0}$.

Definition (3.1.4)[53]: Suppose $\frac{n}{n+\theta(L)} < p < 1$ and$s_0 = \left[\frac{n\left(\frac{1}{p}-1\right)}{m}\right]$. The generalized Hardy space$H_L^p(\mathbb{R}^n)$associated with the semigroup$\{e^{-tL}\}_{t>0}$ is the subspace of the dual space$\left(\mathfrak{L}_{L^*}\left(\frac{1}{p}-1,2,s_0\right)\right)'$ of $\mathfrak{L}_{L^*}\left(\frac{1}{p}-1,2,s_0\right)$, defined as the completion of

$$D_p := \{f \in L^2(\mathbb{R}^n) : \mathcal{A}(Q_{t^m} f) \in L^p(\mathbb{R}^n)\}, \frac{n}{n+\theta(L)} < p < 1,$$

in the quasi-norm

$$\|f\|_{H_L^P} = \|\mathcal{A}(Q_{t^m} f)\|_{L^p}.$$

We will abuse language and say $\mathcal{A}(Q_{t^m}f)(x)$is the area integral function associated with the semigroup$\{e^{-tL}\}_{t>0}$, and still denoted by$S_L(f)$.

(i) Note first that smooth functions with compact support do not necessarily belong to $H_L^p(\mathbb{R}^n)$in general. If $f \in H_L^p(\mathbb{R}^n)$, it follows from Theorem(3.1.10)below that f satisfies the cancellation condition

$$\int_{\mathbb{R}^n} f(x)g(x)dx = 0$$

for all$g \in \mathcal{K}_{(L^*,s_0)}$, where$\mathcal{K}_{(L^*,s_0)}$is given by

$$\mathcal{K}_{(L^*,s_0)} = \{g \in \mathcal{M}: P^*_{s_0,t}g(x) \text{for almost all } x \in \mathbb{R}^n \text{ for all } t > 0\}.$$

See [71] for a discussion of the dimensions of $\mathcal{K}_{(L^*,0)}$when L is a second order elliptic operator of divergence form or a Schrödinger operator.

(ii) If $\theta(L) = \infty$, then the spaces $H_L^p(\mathbb{R}^n)$are well defined for all $0 < p \le 1$. Atypical example of $\theta(L) = \infty$ is when the kernel $p_t(x,y)$ of e^{-tL} satisfies a Gaussian upper bound, that is,

$$|p_t(x,y)| \le \frac{C}{t^{n/2}} e^{-c\frac{|x-y|^2}{t}} \tag{23}$$

for $x, y \in \mathbb{R}^n$ and all $t > 0$.

(iii) We now give a list of examples of $H_L^p(\mathbb{R}^n)$in different settings.

(α) Let$p \le 1$. The classical $H^p(\mathbb{R}^n)$and the $H_\Delta^p(\mathbb{R}^n)$spaces coincide, and their quasi-norms are equivalent. See[74] and[19].

(β) Let$A = \left(\left(a_{ij}(x)\right)\right)_{1\le i,j\le n}$be an $n \times n$ matrix with entries$a_{ij} \in L^\infty(\mathbb{R}^n, \mathbb{C})$ satisfying$Re \sum a_{ij}(x)\xi_i\xi_j \ge \lambda|\xi|^2$ for all$x \in \mathbb{R}^n$,$\xi = (\xi_1, \xi_2, \dots, \xi_n) \in \mathbb{C}^n$and some$\lambda > 0$. We define a divergence form operator

$$Lf \equiv -\text{div}(A\nabla f),$$

which we interpret in the usual weak sense via a sesquilinear form.

Note that the Gaussian bound(23) on the heat kernel$e^{-tL}$is true when A has real entries, or when $n = 1, 2$ in the case of complex entries. See [57].

(γ) Let$V \in L^1_{loc}(\mathbb{R}^n)$be a nonnegative function on $\mathbb{R}^n$. The Schrödinger operator with potential V is defined by

$$L = -\Delta + V(x) \text{ on } \mathbb{R}^n, n \ge 3.$$

From the Feynman-Kac formula, it is well known that the kernels $p_t(x,y)$ of the semigroup e^{-tL}satisfy the estimate

$$0 \le p_t(x,y) \le \frac{1}{(4\pi t)^{n2}} e^{-\frac{|x-y|^2}{4t}}.$$

However, unless V satisfies additional conditions, the heat kernel can be a discontinuous function of the space variables, and the Hölder continuous estimates may fail to hold. See, [64] and [73].

Note that in [62], [63], the tent spaces give a natural and simple approach to the atomic decomposition of functions in the classical $H^p(\mathbb{R}^n)$spaces by using the area integral functions and the connection with the theory of the Carleson measures. We will adopt the same approach of tent spaces to obtain a molecular decomposition for Hardy spaces $H_L^p(\mathbb{R}^n)$. We now assume that

$$\frac{n}{n+\theta(L)} < p < 1, s_0 = \left[\frac{n\left(\frac{1}{p}-1\right)}{m}\right] \text{and } s \geq s_0. \tag{24}$$

In the following, for any given p as in(24), we let ϵ in(5) be a constant such that(i) $\epsilon > n\left(\frac{1}{p}-1\right)$ and thus $\frac{n}{n+\epsilon} < p$; (ii) $m(s_0 + 1) > \epsilon$ unless stated otherwise.

Following[63], a function $a(x,t)$ is called a T_2^p-atom, $0 < p \leq 1$, if

(i) the function $a(t,x)$ is supported in $\hat{B}$(for some ball $B \subset \mathbb{R}^n$); and

(ii) $\int_{\hat{B}} |a(t,x)|^2 \frac{dxdt}{t} \leq |B|^{1-2/p}$.

Proposition (3.1.5)[53]: (a) *Suppose* $1 < p < \infty$. *The following inequality holds, whenever* $f \in T_2^p$ *and* $g \in T_2^{p'}$ *with* $\frac{1}{p} + \frac{1}{p'} = 1$,

$$\int_{\mathbb{R}_+^{n+1}} |f(y,t)g(y,t)| \frac{dydt}{t} \leq \int_{\mathbb{R}^n} \mathcal{A}(f)(x)\mathcal{A}(g)(x)dx.$$

(b) *Assume* $0 < p \leq 1$. *Every element* $f \in T_2^p$ *can be written as* $f = \sum \lambda_j a_j$, *where* a_j *are* T_2^p *atoms,* $\lambda_j \in \mathbb{C}$, *and* $\sum |\lambda_j|^p \leq c\|f\|_{T_2^p}^p$.

(c) *If* $0 < p \leq 1$, *then the dual space of* T_2^p *is* $T_2^{p,\infty}$. *More precisely, the pairing* $\langle f, g\rangle \to \int_{\mathbb{R}_+^{n+1}} f(x,t)g(x,t)\frac{dxdt}{t}$ *realizes* $T_2^{p,\infty}$ *as equivalent to the dual of* T_2^p.

Proof: For the proofs of (a) and(b), we refer to (43) and Proposition 5 of[63], respectively.

Let us show that $\ell \in \left(T_2^p\right)'$ can be represented by a function $g \in T_2^{p,\infty}$. Following Theorem 1 of [63], we observe that if K is a compact set of $\mathbb{R}_+^{n+1}$ and f is supported in K with $f \in L^2(K)$, then $f \in T_2^p$ with $\|f\|_{T_2^p} \leq c_K\|f\|_{L^2(K)}$ for all $0 < p \leq 1$. Thus ℓ induces a bounded linear function on $L^2(K)$and is thus representable by $g_k \in L^2(K)$. Taking an increasing family of such K that exhausts $\mathbb{R}_+^{n+1}$ gives a function $g \in L_{loc}^2(\mathbb{R}_+^{n+1})$ such that $\ell(f) = \int_{\mathbb{R}_+^{n+1}} f(x,t)g(x,t)\frac{dxdt}{t}$ whenever $f \in T_2^p$ and f has compact support K. Testing ℓ against all possible atoms leads by the converse Schwartz's inequality to $|B|^{1-\frac{2}{p}} \int_{\hat{B}} |g(x,t)|^2\, dxdt/t \leq \|\ell\|^2$ for all B, i.e. $\left\|C_p(g)\right\|_\infty \leq \|\ell\|$ as desired. This representation of ℓ is then extendable to all of T_2^p, since the subspace of f with compact support is dense in T_2^p. The proof of Proposition(3.1.5) is complete.

Let m be the constant in(4). For any given $p < 1$, we choose s_0 and s the integers in(24). Let $c_{m,s}$ be a constant such that

$$c_{m,s} \int_0^\infty t^{m(s+2)} e^{-2t^m} \left(1 - e^{-t^m}\right)^{s_0+1} dt/t = 1. \tag{25}$$

We say that a function $\alpha(x)$ is a (p,s)-molecule if

$$\alpha(x) = \pi_L(a)(x) = c_{m,s} \int_0^\infty Q_{s,t^m}\left(I - P_{s_0,t^m}\right)\left(a(t,.)\right)(x)\frac{dt}{t}, \tag{26}$$

where $a(t,x)$ is a T_2^p-atom supported in the tent $\hat{B}$ of some ball $B \subset \mathbb{R}^n$, and $a(t,x)$ satisfies the condition $\int_{\hat{B}} |a(t,x)|^2 dxdt/t \leq |B|^{1-2/p}$.

Proposition (3.1.6)[53]: *Suppose* $\frac{n}{n+\theta(L)} < p \le 1$. *For any* $f \in H_L^p(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, *there exist* (p,s)*-molecules* $\alpha_k(x)$ *and numbers* $\lambda_k (k = 0,1,2,\dots)$ *such that*

$$f(x) = \sum_k \lambda_k \alpha_k(x). \quad (27)$$

The sequence λ_k *satisfies* $\sum_k |\lambda_k|^p \le c\|f\|_{H_L^p}^p$. *Conversely, every sum* (27) *satisfies* $\|f\|_{H_L^p}^p \le c\sum_k |\lambda_k|^p$.

Proof: Let $c_{m,s}$ be a constant in (25). Consider the identity:

$$1 = c_{m,s}\int_0^\infty \left(t^{m(s+1)}\mathcal{Z}^{(s+1)}e^{-t^m\mathcal{Z}}\left(1-e^{-t^m\mathcal{Z}}\right)^{s_0+1}\right)\left(t^m\mathcal{Z}e^{-t^m\mathcal{Z}}\right)\frac{dt}{t},$$

which is valid for all $\mathcal{Z} \ne 0$ in a sector S_μ^0 with $\mu \in (\omega,\pi)$. As a consequence of H_∞-functional calculus for L and the convergence Lemma$(3.1.1)$, one has

$$f(x) = c_{m,s}\int_0^\infty Q_{s,t^m}\left(I - P_{s_0,t^m}\right)Q_{t^m}f(x)\frac{dt}{t}, \quad (28)$$

where this integral converges strongly in $L^2(\mathbb{R}^n)$. See [54] and [79]. For any $f \in H_L^p(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, we let $F(x,t) = (Q_{t^m}f)(x)$. We then apply (b) of Proposition(3.1.5) to $Q_{t^m}f$ to obtain

$$\begin{aligned} f(x) &= c_{m,s}\int_0^\infty Q_{s,t^m}\left(I - P_{s_0,t^m}\right)(Q_{t^m}f)(x)\frac{dt}{t} \\ &= \sum_k \lambda_k\, c_{m,s}\int_0^\infty Q_{s,t^m}\left(I - P_{s_0,t^m}\right)\left(a_k(t,.)\right)(x)\frac{dt}{t} = \sum_k \lambda_k \alpha_k(x), \end{aligned}$$

where the sequence λ_k satisfies $\sum_k |\lambda_k|^p \le c\|Q_{t^m}(f)\|_{T_2^p}^p \le c\|f\|_{H_L^p}^p$. This proved, when $f \in H_L^p(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, possesses a (p,s)-molecule decomposition.

Conversely, by the definition $H_L^p(\mathbb{R}^n)$, it suffices to verify that for any (p,s)-molecule $\alpha(x)$, we have

$$\|S_L(\alpha)\|_p \le c, \quad (29)$$

where c is a positive constant independent of α.

Assume that $\alpha(x) = \pi_L(a)$ where $a = a(t,x)$ is a usual T_2^p-atom supported in $\hat{B}$ (for some $B = B(\mathcal{Z}_0, r_B) \subset \mathbb{R}^n$). One writes

$$\|S_L(\alpha)\|_{L^p}^p = \int_{4B} |S_L(\alpha)(x)|^p\,dx + \int_{(4B)^c} |S_L(\alpha)(x)|^p\,dx = I + II.$$

Using Hölder's inequality and estimate(20), one obtains

$$\begin{aligned} \int_{4B} |S_L(\alpha)(x)|^p\,dx &\le |4B|^{1-\frac{p}{2}}\left\|S_L\left(\pi_L(a)\right)\right\|_{L^2}^p \le c|B|^{1-\frac{p}{2}}\|\pi_L(a)\|_{L^2}^p \\ &\le c|B|^{1-\frac{p}{2}}\||a|\|_{T_2^2}^p \le c. \end{aligned}$$

We now estimate the term II. Firstly, we will show that there exists a constant $c > 0$ such that for any $x \notin 4B$,

$$S_L^2(\alpha)(x) \le c r_B^{2\epsilon}|B|^{2-\frac{2}{p}}|x - \mathcal{Z}_0|^{-2(n+\epsilon)}. \quad (30)$$

Let us prove(30). For $k = 0,\dots,s_0+1$, we denote

$$\Psi_{t,v}^k(L,s)f(y) = \left(t^m + (k+1)v^m\right)^{s+2}\left(\frac{d^{s+2}P_r}{dr^{s+2}}\bigg|_{r=t^m+(k+1)v^m} f\right)(y)$$

By (6), the kerne $\Psi_{t,\nu}^{k}(L,s)(x,y)$ of $\Psi_{t,\nu}^{k}(L,s)$satisfies

$$\left|\Psi_{t,\nu}^{k}(L,s)(x,y)\right| \leq c_k \frac{(t+\nu)^{\epsilon}}{(t+\nu+|y-\mathcal{Z}|)^{n+\epsilon}}.$$

Since$\left(I-P_{s_0,t^m}\right)=\sum_{k=0}^{s_0+1}(-1)^k C_{s_0+1}^{k} e^{-kt^m L}$, we obtain

$$S_L^2(\alpha)(x)=\left(S_L\left(\int_0^{\infty} Q_{s,\nu^m}\left(I-P_{s_0,\nu^m}\right)\left(a(\nu,.)\right)\frac{d\nu}{\nu}\right)\right)^2(x)$$

$$=\int_{\Gamma(x)}\left[\int_0^{r_B}\sum_{k=0}^{s_0+1}(-1)^k C_{s_0+1}^{k}\frac{t^m\nu^{m(s+1)}}{(t^m+(k+1)\nu^m)^{s+2}}\times\Psi_{t,\nu}^{k}(L,s)\left(a(\nu,.)\right)(y)\frac{d\nu}{\nu}\right]^2\frac{dydt}{t^{n+1}}$$

$$\leq c\sum_{k=0}^{s_0+1}\int_{\Gamma(x)}\left[\int_0^{r_B}\int_B\frac{t^m\nu^{m(s+1)}}{(t^m+(k+1)\nu^m)^{s+2}}\left|\Psi_{t,\nu}^{k}(L,s)(y,\mathcal{Z})a(\nu,\mathcal{Z})\right|\frac{d\mathcal{Z}d\nu}{\nu}\right]^2\frac{dydt}{t^{n+1}}$$

$$\leq c\int_0^{r_B}\int_{|x-y|<t}\left[\iint_{\hat{B}}\frac{t^m\nu^{m(s+1)}}{(t+\nu)^{m(s+2)}}\frac{(t+\nu)^{\epsilon}|a(\nu,\mathcal{Z})|}{(t+\nu+|y-\mathcal{Z}|)^{n+\epsilon}}\frac{d\mathcal{Z}d\nu}{\nu}\right]^2\frac{dydt}{t^{n+1}}$$

$$+c\int_{r_B}^{\infty}\int_{|x-y|<t}\left[\iint_{\hat{B}}\frac{t^m\nu^{m(s+1)}}{(t+\nu)^{m(s+2)}}\frac{(t+\nu)^{\epsilon}|a(\nu,\mathcal{Z})|}{(t+\nu+|y-\mathcal{Z}|)^{n+\epsilon}}\frac{d\mathcal{Z}d\nu}{\nu}\right]^2\frac{dydt}{t^{n+1}}$$

$$=II_1+II_2,\text{respectively.}$$

We only consider term II_2 since the estimate of term $II_1$is much simpler. For $x\notin 4B$ and $t\geq r_B$, we set $B=B_1\cup B_2$, where$B_1=B\cap\left\{\mathcal{Z}:|y-\mathcal{Z}|\leq\frac{|x-\mathcal{Z}_0|}{2}\right\}$. For any $\mathcal{Z}\in B_1$and$|y-x|<t$, we have

$$|x-\mathcal{Z}_0||y-x|+|y-\mathcal{Z}|+|\mathcal{Z}-\mathcal{Z}_0|\leq t+\frac{|x-\mathcal{Z}_0|}{2}+r_B\leq 2t+\frac{|x-\mathcal{Z}_0|}{2}, \qquad (31)$$

which impliest$t\geq|x-\mathcal{Z}_0|/4$, and then $(t+s+|y-\mathcal{Z}|)\geq|x-\mathcal{Z}_0|/4$. Obviously, for any$\mathcal{Z}\in B_2$ and$|y-x|<t$, we also have$(t+s+|y-\mathcal{Z}|)\geq|x-\mathcal{Z}_0|/2$. Those, together with

$$(t+\nu)^{\epsilon}.\frac{t^m\nu^{m(s+1)}}{(t+\nu)^{m(s+2)}}\leq t^{\epsilon}\left(\frac{\nu}{t}\right)^{m(s+1)}$$

and $m(s+1)>\epsilon$, give

$$II_2\leq c\int_{r_B}^{\infty}\int_{|x-y|<t}\left[\iint_{\hat{B}}\nu^{m(s+1)}t^{\epsilon-m(s+1)}|a(\nu,\mathcal{Z})|\frac{d\mathcal{Z}d\nu}{\nu}\right]^2\frac{dydt}{t^{n+1}}|x-\mathcal{Z}_0|^{-2(n+\epsilon)}$$

$$\leq c|B|\int_{r_B}^{\infty}\int_0^{r_B}\nu^{2m(s+1)}t^{2\epsilon-2m(s+1)}\frac{d\nu dt}{\nu}\int_{\hat{B}}|a(\nu,\mathcal{Z})|^2\frac{d\mathcal{Z}d\nu}{\nu}|x-\mathcal{Z}_0|^{-2(n+\epsilon)}$$

$$\leq c|B|^{2-\frac{2}{p}}\int_{r_B}^{\infty}\int_0^{r_B}\nu^{2m(s+1)}t^{2\epsilon-2m(s+1)}\frac{d\nu dt}{\nu}|x-\mathcal{Z}_0|^{-2(n+\epsilon)}$$

$$\leq c r_B^{2\epsilon+2n-\frac{2n}{p}}|x-\mathcal{Z}_0|^{-2(n+\epsilon)}.$$

Estimate(30) then follows readily. Since$p(n+\epsilon)>n$, we obtain

$$\int_{(4B)^c}|S_L(\alpha)(x)|^p\,dx\leq c r_B^{p\epsilon+np-n}\int_{(4B)^c}|x-\mathcal{Z}_0|^{-(n+\epsilon)p}\,dx\leq c.$$

Combining estimates of I andII, we obtain(29), and then the proof of Proposition (3.1.6) is complete.

Now, let $T_{2,c}^{p}$ be the set of all $f \in T_2^p$ with compact support in $\mathbb{R}_+^{n+1}$. Consider the operator π_L initially defined on $T_{2,c}^{p}$ by

$$\pi_L(f)(x) = c_{m,s} \int_0^\infty Q_{s,t^m}\big(I - P_{s_0,t^m}\big)\big(f(.,t)\big)(x)\frac{dt}{t}, \tag{32}$$

where $c_{m,s}$ is a constant in (25). Note that for any compact set K in $\mathbb{R}_+^{n+1}$,

$$\int_K |f(x,t)|^2 dxdt \le c(K,p)\|\mathcal{A}(f)\|_p^2$$

for all $p > 0$. See page 306 of [63]. This and the estimate (7) imply that the integral (32) is well defined, and $\pi_L(f) \in L^2(\mathbb{R}^n)$ for all $f \in T_{c,2}^{p}$.

Lemma (3.1.7)[53]: *The operator π_L, initially defined on $T_{c,2}^{p}$, extends to a bounded linear operator from*

(a) T_2^p to L^p if $1 < p < \infty$;

(b) T_2^p to H_L^p if $\frac{n}{n+\theta(L)} < p \le 1$.

Proof. The property (b) is contained in the second part of Proposition (3.1.6). We now verify the property (a). From Proposition (3.2.5) and estimate (20), we have

$$\begin{aligned}\left|\int_{\mathbb{R}^n} \pi_L(f)(x)g(x)dx\right| &\le c\left|\int_{\mathbb{R}_+^{n+1}} f(x,t)\left(Q_{s,t^m}\big(I - P_{s_0,t^m}\big)\right)^* g(x)\frac{dxdt}{t}\right| \\ &\le c\left|\int_{\mathbb{R}^n} \mathcal{A}(f)(x)\mathcal{A}\left(\left(Q_{s,t^m}\big(I - P_{s_0,t^m}\big)\right)^* g\right)(x)\,dx\right| \\ &\le c\|\mathcal{A}(f)\|_p \left\|\mathcal{A}\left(\left(Q_{s,t^m}\big(I - P_{s_0,t^m}\big)\right)^* g\right)\right\|_{p'} \\ &\le c|||f|||_{T_2^p}\|g\|_{p'}\end{aligned}$$

for any $g \in L^{p'}, \frac{1}{p} + \frac{1}{p'} = 1$. Hence, we obtain $\|\pi_L(f)\|_p \le c|||f|||_{T_2^p}$.

We next state the following H_L^p-estimate for functions in the space $H_L^p(\mathbb{R}^n)$. For its proof, it is similar to that of the second part of Proposition (3.1.6).

Proposition (3.1.8)[53]: *Suppose $\frac{n}{n+\theta(L)} < p \le 1$ and $s_0 = \left[\frac{n\left(\frac{1}{p}-1\right)}{m}\right]$. For any L^2-function f supported on B, the function $\big(I - P_{s_0,r_B^m}\big)f$ belongs to $H_L^p(\mathbb{R}^n)$, and there exists a positive constant c such that*

$$\left\|\big(I - P_{s_0,r_B^m}\big)f\right\|_{H_L^p} \le c|B|^{\frac{1}{p}-\frac{1}{2}}\|f\|_{L^2},$$

where r_B is the radius of the ball B.

We assume that the operator L satisfies the assumptions (a) and (b). It was proved in Theorem3.1 of [71] that the dual space of $H_L^1(\mathbb{R}^n)$ is the space $BMO_{L^*}(\mathbb{R}^n)$ in which L^* is the adjoint operator of L. The aim is to prove the following theorem.

Recall that a measure μ in $\mathbb{R}_+^{n+1}$ is a Carleson measure V^β of order of $\beta \ge 1$ if there is a positive constant c such that for each ball B on $\mathbb{R}^n$,

$$\mu\big(\hat{B}\big) \le c|B|^\beta, \tag{33}$$

where $\hat{B}$ is the tent over B. The smallest bound c in (33) is defined to be the norm of μ, and denoted by $|||\mu|||_{V^\beta}$. See, for example, page338, Chapter XV, of [85].

Proposition (3.1.9)[53]: *Suppose* $0 < \alpha < \theta(L)/n$ *and* $s \geq s_0 = \left[\frac{n\alpha}{m}\right]$. *If* $f \in \mathfrak{L}_L(\alpha, 2, s_0)$, *then the measure*

$$\mu_f(x,t) = \left|Q_{s,t^m}\left(\mathcal{I} - P_{s_0,t^m}\right)f(x)\right|^2 \frac{dxdt}{t}$$

is a Carleson measure $V^{2\alpha+1}$ *with* $\left|\left\|\mu_f\right\|\right|_{V^{(2\alpha+1)}} \leq c\|f\|^2_{\mathfrak{L}_L(\alpha,2,s_0)}$.

Proof: Given $0 < \alpha < \theta(L)/n$, we let ϵ in (5) be the constant such that $n\alpha < \epsilon < \theta(L)$ and $m(s+1) > \epsilon$. In order to prove Proposition(3.1.9), it suffices to prove that there exists a positive constant $c > 0$ such that for any ball $B = B(x_B, r_B)$ on $\mathbb{R}^n$,

$$\int\int_{\hat{B}} \left|Q_{s,t^m}\left(I - P_{s_0,t^m}\right)f(x)\right|^2 \frac{dxdt}{t} \leq c|B|^{2\alpha+1}\|f\|^2_{\mathfrak{L}_L(\alpha,2,s_0)}. \quad (34)$$

Note that $\left(I - P_{s_0,t^m}\right) = \left(I - P_{s_0,t^m}\right)\left(I - P_{s_0,r_B^m}\right) + P_{s_0,r_B^m}\left(I - P_{s_0,t^m}\right)$. Hence, estimate (34) will follow from the following estimates (35) and (36):

$$\int\int_{\hat{B}} \left|Q_{s,t^m}\left(I - P_{s_0,t^m}\right)\left(I - P_{s_0,r_B^m}\right)f(x)\right|^2 \frac{dxdt}{t} \leq c|B|^{2\alpha+1}\|f\|^2_{\mathfrak{L}_L(\alpha,2,s_0)} \quad (35)$$

and

$$\int\int_{\hat{B}} \left|Q_{s,t^m}P_{s_0,r_B^m}\left(I - P_{s_0,t^m}\right)f(x)\right|^2 \frac{dxdt}{t} \leq c|B|^{2\alpha+1}\|f\|^2_{\mathfrak{L}_L(\alpha,2,s_0)}. \quad (36)$$

To prove(35), let us introduce the square function $\mathcal{G}f$, given by

$$\mathcal{G}(f)(x) = \left(\int_0^\infty \left|Q_{s,t^m}\left(I - P_{s_0,t^m}\right)f(x)\right|^2 \frac{dt}{t}\right)^{1/2}.$$

From (7), the function $\mathcal{G}(f)$ is bounded on $L^2(\mathbb{R}^n)$. Let $b_1 = \left(I - P_{s_0,r_B^m}\right)f\chi_{2B}$, and $b_2 = \left(I - P_{s_0,r_B^m}\right)f\chi_{(2B)^c}$. Using property(15)of $\mathfrak{L}_L(\alpha, 2, s_0)$, we obtain

$$\begin{aligned}\int\int_{\hat{B}} \left|Q_{s,t^m}\left(I - P_{s_0,t^m}\right)b_1(x)\right|^2 \frac{dxdt}{t} &\leq \|\mathcal{G}(b_1)\|_2^2 \leq c\|b_1\|^2_{L^2(\mathbb{R}^n)} \\ &= c\int_{2B} \left|\left(I - P_{s_0,r_B^m}\right)f(x)\right|^2 dx \\ &\leq c\left(\int_{2B} \left|\left(I - P_{s_0,r_{2B}^m}\right)f(x)\right|^2 dx + |B|. \sup_{x\in 2B}\left|P_{s_0,r_B^m}f(x) - P_{s_0,r_{2B}^m}f(x)\right|^2\right) \\ &\leq c|B|^{2\alpha+1}\|f\|^2_{\mathfrak{L}_L(\alpha,2,s_0)}. \end{aligned} \quad (37)$$

On the other hand, for any $x \in B$ and $y \in (2B)^c$, one has $|x-y| \geq r_B$. And then by estimate (13) and property (iv) of $\mathfrak{L}_L(\alpha, 2, s_0)$,

$$\begin{aligned}\left|Q_{s,t^m}\left(I - P_{s_0,t^m}\right)b_2(x)\right| &\leq c\int_{\mathbb{R}^n\backslash 2B} \frac{t^\epsilon}{(t + |x-y|)^{n+\epsilon}}\left|\left(\mathcal{I} - P_{s_0,r_B^m}\right)f(y)\right|dy \\ &\leq c\left(\frac{t}{r_B}\right)^\epsilon \int_{\mathbb{R}^n} \frac{r_B^\epsilon}{(r_B + |x-y|)^{n+\epsilon}}\left|\left(I - P_{s_0,r_B^m}\right)f(y)\right|dy \\ &\leq c|B|^{2\alpha}\left(\frac{t}{r_B}\right)^\epsilon \|f\|_{\mathfrak{L}_L(\alpha,2,s_0)}\end{aligned}$$

since $n\alpha < \epsilon$. Therefore,

$$\begin{aligned}\int\int_{\hat{B}} \left|Q_{s_0,t^m}\left(I - P_{s_0,t^m}\right)b_2(x)\right|^2 \frac{dxdt}{t} &\leq c|B|^{2\alpha}\frac{1}{r_B^{2\epsilon}}\int\int_{\hat{B}} t^{2\epsilon}\frac{dxdt}{t}\|f\|^2_{\mathfrak{L}_L(\alpha,2,s_0)} \\ &\leq c|B|^{2\alpha+1}\|f\|^2_{\mathfrak{L}_L(\alpha,2,s_0)}.\end{aligned}$$

This, together with the estimate (37), gives the estimate (35).

Let us prove(36). For $k = 1, 2, \dots, s_0 + 1$, we denote by

$$\psi_{t,r_B}^k f(x) = (kr_B^m + t^m)^{s+1} \left(\frac{d^{s+1} P_\nu}{d\nu^{s+1}} \mid_{\nu = kr_B^\epsilon + t^m f} \right)(x).$$

From (13), the kernel $\psi_{t,r_B}^k(x,y)$ of ψ_{t,r_B}^ksatisfies

$$\left|\psi_{t,r_B}^k(x,y)\right| \le c_k \frac{r_B^\epsilon}{(r_B + t + |x-y|)^{n+\epsilon}}.$$

Since $P_{s_0,r_B^m} f(x) = \sum_{k=1}^{s_0+1} (-1)^{k+1} C_{s_0+1}^k e^{-kr_B^m L} f(x)$, from property (16) of $\mathfrak{L}_L(\alpha, 2, s_0)$ together with $n\alpha < \epsilon$, we have

$$\begin{aligned}
&\left|Q_{s,t^m} P_{s_0,r_B^m}\left(I - P_{s_0,t^m}\right) f(x)\right| \\
&= \left|\sum_{k=1}^{s_0+1} (-1)^{k+1} C_{s_0+1}^k \frac{t^{m(s+1)}}{(kr_B^m + t^m)^{s+1}} \psi_{t,r_B}^k \left(I - P_{s_0,t^m}\right) f(x)\right| \\
&\le c \left(\frac{t}{r_B}\right)^{m(s+1)-\epsilon} \int_{\mathbb{R}^n} \frac{t^\epsilon}{(t + |x-y|)^{n+\epsilon}} \left|\left(I - P_{s_0,t^m}\right) f(y)\right| dy \\
&\le c t^{n\alpha/m} \left(\frac{t}{r_B}\right)^{m(s+1)-\epsilon} \|f\|_{\mathfrak{L}_L(\alpha,2,s_0)}, \quad \forall 0 < t \le t_B.
\end{aligned}$$

Therefore, by using the condition $m(s+1) > \epsilon$,

$$\begin{aligned}
\int \int_{\hat{B}} \left|Q_{s,t^m}\left(I - P_{s_0,t^m}\right) f(x)\right|^2 \frac{dxdt}{t} &\le c \int \int_{\hat{B}} t^{2n\alpha/m - 1} \left(\frac{t}{r_B}\right)^{m(s+1)-\epsilon} dxdt \|f\|_{\mathfrak{L}_L(\alpha,2,s_0)}^2 \\
&\le c|B|^{2\alpha+1} \|f\|_{\mathfrak{L}_L(\alpha,2,s_0)}^2,
\end{aligned}$$

which gives the proof of (34), and therefore the proof of Proposition (3.1.9).

Theorem (3.1.10)[53]: *Suppose* $\frac{n}{n+\theta(L)} < p < 1$ *and* $s_0 = \left[\frac{n\left(\frac{1}{p}-1\right)}{m}\right]$. *Then, the dual space of the* $H_L^p(\mathbb{R}^n)$ *space is the* $\mathfrak{L}_{L^*}\left(\frac{1}{p} - 1, 2, s_0\right)$ *space, in the following sense.*

(i) *Suppose* $f \in \mathfrak{L}_{L^*}\left(\frac{1}{p} - 1, 2, s_0\right)$. *Then the linear functional* ℓ *given by*

$$\ell(g) = \int_{\mathbb{R}^n} f(x) g(x) dx, \tag{38}$$

initially defined on the dense subspace $H_L^p(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, *has a unique extension to* $H_L^p(\mathbb{R}^n)$.

(ii) *Conversely, every continuous linear functional* ℓ *on the* $H_L^p(\mathbb{R}^n)$ *space can be realized as above, i.e., there exists* $f \in \mathfrak{L}_{L^*}\left(\frac{1}{p} - 1, 2, s_0\right)$. *such that* (38) *holds and* $\|f\|_{\mathfrak{L}_{L^*}\left(\frac{1}{p}-1,2,s_0\right)} \le c\|\ell\|$.

Proof: Suppose $f \in \mathfrak{L}_{L^*}\left(\frac{1}{p} - 1, 2, s_0\right)$ and b is a (p,s)-molecule of $H_L^p(\mathbb{R}^n)$. Without loss of generality, we assume that

$$b(x) = c_{m,s} \int_0^\infty Q_{s,t^m}\left(I - P_{s_0,t^m}\right)\left(a(t,.)\right)(x) \frac{dt}{t},$$

where $a(t,x)$ is a usual T_2^p-atom supported in $\hat{B}$ (for some $B = B(\mathcal{Z}_0, r_B) \subset \mathbb{R}^n$), and $c_{m,s}$ is the constant in (25). We can apply the same argument as in Theorem 5.1 of [71] to obtain the following identity:

$$\int_{\mathbb{R}^n} b(x)f(x)dx = c_{m,s}\int_{\mathbb{R}^{n+1}_+} a(t,x)Q^*_{s,t^m}\big(I-P^*_{s,t^m}\big)f(x)\frac{dxdt}{t}.$$

The details are omitted here. This, together with Proposition(3.1.9), shows that

$$\left|\int_{\mathbb{R}^n} b(x)f(x)dx\right| \le c\|\|a\|\|_{T_2^2}\left(\int_{\hat{B}}\left|Q^*_{s,t^m}\big(I-P^*_{s_0,t^m}\big)f(x)\right|^2\frac{dxdt}{t}\right)^{1/2} \le c|B|^{\frac{1}{2}-\frac{1}{p}}|B|^{\frac{1}{p}-1+\frac{1}{2}}$$
$$\le c. \qquad (39)$$

For any $g \in H^p_L(\mathbb{R}^n)\cap L^2(\mathbb{R}^n)$, it follows from Proposition (3.1.6) that there exist (p,s)-molecules $\alpha_k(x)$ and numbers $\lambda_k(k=0,1,2,\dots.)$ such that $g(x)=\sum_k\lambda_k\alpha_k(x)$. The sequence λ_k satisfies $\sum_k|\lambda_k|^p \le c\|g\|^p_{H^p_L}$. Hence by (39) we obtain

$$\left|\int_{\mathbb{R}^n} f(x)g(x)dx\right| \le \sum_k|\lambda_k|\left|\int_{\mathbb{R}^n}\alpha_k(x)f(x)dx\right| \le c\sum_k|\lambda_k| \le c\left(\sum_k|\lambda_k|^p\right)^{1p}$$
$$\le c\|g\|_{H^p_L}.$$

This proves (i).

Let us prove(ii). Note that by (b), for every $h_t(x)\in T^p_2$,

$$\mathcal{R}(h_t)(x) = c_{m,s}\int_0^\infty Q_{st^m}\big(I-P_{s_0,t^m}\big)(h_t)(x)\frac{dt}{t} \in H^p_L,$$

where $c_{m,s}$ is a constant in (25). Therefore, for each continuous linear functional ℓ on the $H^p_L(\mathbb{R}^n)$space, we obtain

$$|(\ell\circ\mathcal{R})(h_t)| = |\ell o\mathcal{R}(h_t)| \le \|\ell\|_{H^p_L\to\mathbb{C}}\|\mathcal{R}(h_t)\|_{H^p_L} \le \|\ell\|_{H^p_L\to\mathbb{C}}\|\mathcal{R}\|_{T^p_2\to H^p_L}\|h_t\|_{T^p_2}$$

for all $h_t(x)\in T^p_2$. It then follows from (c) of Proposition (3.1.5) that there exists a function $\mathcal{Z}_t(x)\in T^{p,\infty}_2$ such that

$$(\ell o\mathcal{R})(h_t) = \int_{\mathbb{R}^{n+1}_+}\mathcal{Z}_t(x)h_t(x)\frac{dxdt}{t}. \qquad (40)$$

On the other hand, by(28) we have that for any $g\in H^p_L(\mathbb{R}^n)\cap L^2(\mathbb{R}^n)$,

$$g(x) = c_{m,s}\int_0^\infty Q_{st^m}\big(I-P_{s_0,t^m}\big)Q_{t^m}g(x)\frac{dt}{t}.$$

This shows that for each continuous linear functional ℓ on the $H^p_L(\mathbb{R}^n)$space, we have that for all $g\in H^p_L(\mathbb{R}^n)$,

$$\ell(g) = \lim_{k\to\infty}\ell(g_k) = \lim_{k\to\infty}\ell o\mathcal{R}oQ_{t^m}(g_k) = \ell o\mathcal{R}oQ_{t^m}(g), \qquad (41)$$

where $\{g_k\}_k$ is a family of functions satisfying $g_k\in H^p_L(\mathbb{R}^n)\cap L^2(\mathbb{R}^n)$ and $\lim_{k\to\infty} g_k = g$. From(40) and(41), we have that for all $g\in H^p_L(\mathbb{R}^n)$, there exists a function $\mathcal{Z}_t(x)\in T^{p,\infty}_2$ such that

$$\ell(g) = lo\mathcal{R}oQ_{t^m}(g) = \int_{\mathbb{R}^{n+1}_+}\mathcal{Z}_t(x)Q_{t^m}g(x)\frac{dxdt}{t} = \int_{\mathbb{R}^n}\left(\int_0^\infty Q^*_{t^m}\mathcal{Z}_t(x)\frac{dt}{t}\right)g(x)dx$$
$$\stackrel{\text{def}}{=}\int_{\mathbb{R}^n} f(x)g(x)dx, \qquad (42)$$

where $f(x)=\int_0^\infty Q^*_{t^m}\mathcal{Z}_t(x)\frac{dt}{t}$.

We now prove that $f\in\mathfrak{L}_{L^*}\left(\frac{1}{p}-1,2,s_0\right)$. For any ball $B=B(x_B,r_B)$, it follows from (42) and Proposition (3.1.8) that we obtain

$$\left(\int_B \left|f - P^*_{s_0,r_B^m} f\right|^2 dx\right)^{1/2} = \sup_{\|g\|_{L^2(B)}\le 1} \left|\int_{\mathbb{R}^n} \left(f(x) - P^*_{s_0,r_B^m} f(x)\right) g(x)dx\right|$$
$$= \sup_{\|g\|_{L^2(B)}\le 1} \left|\int_{\mathbb{R}^n} f(x)\left(I - P_{s_0,r_B^m}\right) g(x)dx\right|$$
$$\le \sup_{\|g\|_{L^2(B)}\le 1} \left|\ell\left(\left(I - P_{s_0,r_B^m}\right)g\right)\right|$$
$$\le \|\ell\| \sup_{\|g\|_{L^2(B)}\le 1} \left\|\ell\left(\left(I - P_{s_0,r_B^m}\right)g\right)\right\|_{H_L^p}$$
$$\le c\|\ell\||B|^{\frac{1}{2}+\left(\frac{1}{p}-1\right)}.$$

This proves $f \in \mathfrak{L}_{L^*}\left(\frac{1}{p} - 1, 2, s_0\right)$ with $\|f\|_{f\in\mathfrak{L}_{L^*}\left(\frac{1}{p}-1,2,s_0\right)} \le c\|\ell\|$, and then the proof of Theorem (3.1.10) is complete.

As a consequence of Theorem(3.1.10), we have the following corollary.

Corollary (3.1.11)[53]: *Suppose that the operator L satisfies the assumptions (a) and (b). Let $0 < \alpha < \theta(L)/n$ and $s \ge s_0 = \left[\frac{n\alpha}{m}\right]$. Then the spaces $\mathfrak{L}_L(\alpha, 2, s)$ and $\mathfrak{L}_L(\alpha, 2, s_0)$ coincide, and their norms are equivalent.*

Proof: Suppose $\frac{n}{n+\theta(L)} < p < 1$ and $s \ge s_0 = \left[\frac{n\left(\frac{1}{p}-1\right)}{m}\right]$. As in Definition(3.1.4), we define $H_{L^*}^{p,s}(\mathbb{R}^n)$ as the collection of all continuous linear functionals on $\mathfrak{L}_L(\frac{1}{p} - 1, 2, s)$ satisfying $Q^*_{t^m} f \in T_2^p$, and thus $H_{L^*}^p(\mathbb{R}^n) = H_{L^*}^{p,s_0}(\mathbb{R}^n)$. Similarly to the proof of Proposition(3.1.6), a molecular characterization(27) also holds for functions in $H_{L^*}^{p,s}(\mathbb{R}^n)$. Hence, the spaces $H_{L^*}^{p,s}(\mathbb{R}^n)$ and $H_{L^*}^p(\mathbb{R}^n)$ coincide, and their quasi-norms are equivalent. On the other hand, the same argument as in the proof of Theorem (3.1.10) shows that the dual space of $H_{L^*}^{p,s}(\mathbb{R}^n)$ is the space $\mathfrak{L}_L(\frac{1}{p} - 1, 2, s)$ For the proof, we omit details here. This, together with Theorem(3.1.10), gives Corollary(3.1.11).

We continue with the assumptions that the operator L satisfies the assumptions (a) and (b). For $0 < \alpha < \frac{n}{m}$, we consider the generalized fractional integrals $L^{-\alpha}$ associated with the operator L, defined by

$$L^{-\alpha} f(x) = \frac{1}{\Gamma(\alpha)} \int_0^\infty t^{\alpha-1} e^{-tL} f(x)dt, \tag{43}$$

where $\Gamma(\alpha)$ is an appropriate constant. Note that if L is the Laplacian $-\Delta$ on $\mathbb{R}^n$, then L^α is the classical fractional integral. See, for example, Chapter 5 in[82].

For $1 < p < \infty$, we let $H_L^p(\mathbb{R}^n) = L^p(\mathbb{R}^n)$(by (20)). The following theorem generalizes Theorem (3.1.10) of Taibleson-Weiss([86]).

Theorem (3.1.12)[53]: *Suppose $\frac{n}{n+\theta(L)} < p_1 < \infty, 0 < \alpha < \frac{n}{m}$ and $\frac{1}{p_2} = \frac{1}{p_1} - \frac{m\alpha}{n}$, then the fractional integral L^α maps $H_L^{p_1}(\mathbb{R}^n)$ continuously into $H_L^{p_2}(\mathbb{R}^n)$. If we replace $H_L^\infty(\mathbb{R}^n)$ with $BMO_L(\mathbb{R}^n)$, the result holds for $\alpha = \frac{n}{mp_1}$.*

Proof: For any given p_1, we let ϵ in (5) be the constant such that $\epsilon < \theta(L)$ and $\frac{n}{n+\epsilon} < p_1$. The result will follow from the repeated application four cases below.

Case I. $1 < p_1 < p_2 < \infty$. This is a well-known result of TheoremII. 2.7, page 12 of[88].

Case II. $p_1 \leq 1 < p_2$ and $0 < \alpha < \frac{n}{m}$. Choose $s_0 = \left[\frac{n\left(\frac{1}{p_1}-1\right)}{m}\right]$ and $s \geq \left[\alpha + \frac{\epsilon}{m}\right]$. If $b(x) = \int_0^\infty Q_{s,t^m}\left(\mathcal{J} - P_{s_0,t^m}\right)\left(a(t,.)\right)(x)\, dt/t$, and $a(t,x)$ is a usual $T_2^{p_1}$-atom supported in $\hat{B}$(for some $B = B(\mathcal{Z}_0, r_B) \subset \mathbb{R}^n$), we will show that $\|L^{-\alpha}(b)\|_{p_2} \leq c$.

Case III. $p_1 < p_2 \leq 1$ and $0 < \alpha < \frac{n}{m}$. If b is a (p_1, s)-molecule as in Case II, we will show that $\left\|S_L\left(L^{-\alpha}(b)\right)\right\|_{p_2} \leq c$.

Case IV. $p > 1$ and $\alpha = \frac{n}{m}$. We need to show that $L^{-\frac{n}{mp}}: L^p(\mathbb{R}^n) \to BMO_L(\mathbb{R}^n)$ continuously. By Theorem 3.1 of [71], $BMO_L(\mathbb{R}^n) = \left(H^1_{L^*}(\mathbb{R}^n)\right)'$ in which L^* is the adjoint operator of L, and from Cases I and II we have the desired result by duality. We give the details for Cases II and III.

Proof of Case II. Let $2 < q_2 < \infty$ such that $\frac{1}{p_1} - \frac{1}{p_2} = \frac{1}{2} - \frac{1}{q_2} = \frac{m\alpha}{n}$. We write

$$\|L^{-\alpha}(b)\|_{p_2} \leq \left(\int_{4B} |L^{-\alpha}(b)(x)|^{p_2} dx\right)^{1/p_2} + \left(\int_{(4B)^c} |L^{-\alpha}(b)(x)|^{p_2} dx\right)^{1/p_2} = I + II.$$

Note that $\frac{1}{p_1} - \frac{1}{p_2} = \frac{1}{2} - \frac{1}{q_2} = \frac{m\alpha}{n}$. Using Hölder's inequality, Case I and (a) of Lemma(3.1.7), we obtain

$$I \leq |B|^{\frac{1}{p_2} - \frac{1}{q_2}} \|L^{-\alpha}(b)\|_{q_2} \leq c|B|^{\frac{1}{p_1} - \frac{1}{2}} \|b\|_2 \leq c|B|^{\frac{1}{p_1} - \frac{1}{2}} \|a\|_{T_2^2} \leq c.$$

We now estimate the term II. We will show that there exists a constant $c > 0$ such that for any $x \notin 4B$,

$$|L^{-\alpha}(b)(x)| \leq c r_B^{\epsilon + m\alpha + n - \frac{n}{p_1}} |x - \mathcal{Z}_0|^{-(n+\epsilon)}. \tag{44}$$

Let us prove(44). For any $k = 0, \ldots, s_0 + 1$, we denote by

$$\psi^k_{t,v}(L,s) f(x) = (-1)^{s+1} (t^m + (k+1)v^m)^{s+1} \left(\frac{d^{s+1} P_r}{dr^{s+1}} \mid_{r = t^m + (k+1)v^m} f\right)(x).$$

Note that $\left(I - P_{s_0,t^m}\right) f(x) = \sum_{k=0}^{s_0+1} (-1)^k C^k_{s_0+1} e^{-kt^m L} f(x)$. This, together with property (13) and the fact that for $x \notin 4B$ and $y \in B$,$|x - y| > 2|x - \mathcal{Z}_0|$, yields

$$|L^{-\alpha}(b)(x)| \leq c \left|\int_0^\infty \int_0^\infty t^\alpha e^{-tL} Q_{s,v^m}\left(\mathcal{J} - P_{s_0,t^m}\right) a(v,.)(x) \frac{dv}{v}\frac{dt}{t}\right|$$

$$\leq c \sum_{k=0}^{s_0+1} \left|\int_0^\infty \int_0^\infty \int_B \frac{v^{m(s+1)} t^{m\alpha}}{\left((k+1)v^m + t^m\right)^{s+1}} \times \psi^k_{t,v}(L,s)(x,y) a(v,y) dy \frac{dv}{v}\frac{dt}{t}\right|$$

$$\leq c \int_0^\infty \int_0^\infty \frac{v^{m(s+1)} t^{m\alpha}}{(v+t)^{m(s+1)}} . \frac{(t+v)^\epsilon}{(t + v + |x,y|)^{n+\epsilon}} |a(v,y)| dy \frac{dv}{v}\frac{dt}{t}$$

$$\leq c \int_0^{r_B} \int_{\hat{B}} \frac{v^{m(s+1)} t^{m\alpha}}{(v+t)^{m(s+1)-\epsilon}} |a(v,y)| \frac{dy dv}{v}\frac{dt}{t} |x - \mathcal{Z}_0|^{-(n+\epsilon)}$$

$$\leq c \int_{r_B}^\infty \int_{\hat{B}} \frac{v^{m(s+1)} t^{m\alpha}}{(v+t)^{m(s+1)-\epsilon}} |a(v,y)| \frac{dy dv}{v}\frac{dt}{t} |x - \mathcal{Z}_0|^{-(n+\epsilon)}$$

$$= II_1 + II_2.$$

We only estimate the term II_2 since the estimate of the term II_1 is much simpler.
Using Hölder's inequality and the condition $m(s+1) > \epsilon + m\alpha$, we obtain

$$II_2 \leq c|B|^{1/2} \int_{r_B}^{\infty} \left(\int_0^{r_B} \nu^{2(\epsilon+m\alpha)} \left(\frac{\nu}{t}\right)^{2(m(s+1)-\epsilon-m\alpha)} \frac{d\nu}{\nu} \right)^{1/2} \frac{dt}{t}$$
$$\times \||a\||_{T_2^2} |x - \mathcal{Z}_0|^{-(n+\epsilon)}$$
$$\leq c r_B^{m(s+1)+n-\frac{n}{p_1}} |x - \mathcal{Z}_0|^{-(n+\epsilon)} \int_{r_B}^{\infty} t^{\epsilon+m\alpha-m(s+1)} \frac{dt}{t}$$
$$\leq c r_B^{\epsilon+m\alpha+n-\frac{n}{p_1}} |x - \mathcal{Z}_0|^{-(n+\epsilon)}.$$

Similarly, we have that $II_1 \leq c r_B^{\epsilon+m\alpha+n-\frac{n}{p_1}} |x - \mathcal{Z}_0|^{-(n+\epsilon)}$, and then the estimate (44) is obtained. Hence,

$$\int_{(4B)^c} |L^{-\alpha}(b)(x)|^{p_2} dx \leq c r_B^{\left(\epsilon+m\alpha+n-\frac{n}{p_1}\right)p_2} \int_{(4B)^c} |x - \mathcal{Z}_0|^{-(n+\epsilon)p_2} dx$$
$$\leq c r_B^{m\alpha p_2+n-\frac{p_2 n}{p_1}} \leq c.$$

This completes the proof of Case II.

Proof of Case III. We write

$$\left\|S_L\left(L^{-\alpha}(b)\right)\right\|_{p_2} \leq c \left(\int_{4B} \left|S_L\left(L^{-\alpha}(b)\right)(x)\right|^{p_2} dx \right)^{1/p_2}$$
$$+c \left(\int_{(4B)^c} \left|S_L\left(L^{-\alpha}(b)\right)(x)\right|^{p_2} dx \right)^{1/p_2}$$
$$= I + II.$$

Since the area integral function $S_L$is bounded onL^r for all $1 < r < \infty$, by Case I we have $I \leq c|B|^{\frac{1}{p_2}\frac{1}{q_2}} \|L^{-\alpha}(b)\|_{L^{q_2}} \leq c$.

We now estimate the term II. As in CaseII, it suffices to show that there exists a constant $c > 0$ such that for any $x \notin 2c_1$,

$$S_L\left(L^{-\alpha}(b)\right)(x) \leq c r_B^{\epsilon+m\alpha+n-\frac{n}{p_1}} |x - \mathcal{Z}_0|^{-(n+\epsilon)}. \quad (45)$$

Let us prove(45). For any $k = 0, \dots, s_0 + 1$, we denote by

$$\psi_{t,\nu,\gamma}^k(L, s)f(x) = (t^m + \nu^m + (k+1)\gamma^m)^{s+2} \left(\frac{d^{s+2}P_r}{dr^{s+2}} \,|_{r=t^m+\nu^m+(k+1)\gamma^m} f \right)(x)$$

and

$$\mathrm{h}_{t,\nu,\gamma,k}^{m,\alpha,s} = \frac{t^m \nu^{m\alpha} \gamma^{m(s+1)}}{(t^m + \nu^m + (k+1)\gamma^m)^{s+2}} \leq c \frac{t^m \nu^{m\alpha} \gamma^{m(s+1)}}{(t + \nu + \gamma)^{m(s+2)}}.$$

From estimate(13), we obtain that the kernel$\psi_{t,\nu,\gamma}(L, s)(y, \mathcal{Z})$of the operator $\psi_{t,\nu,\gamma}(L, s)$satisfies

$$\left|\psi_{t,\nu,\gamma}^k(L, s)(y, \mathcal{Z})\right| \leq c \frac{(t + \nu + \gamma)^{\epsilon}}{(t + \nu + \gamma + |y - \mathcal{Z}|)^{n+\epsilon}}.$$

We then obtain

$$\mathrm{S}_L^2\left(L^{-\alpha}(b)\right)(x)$$
$$\leq c \left(S_L \left(\int_0^{\infty} \int_0^{\infty} \nu^{\alpha} e^{-\nu L} Q_{s,\gamma^m} \left(I - P_{s_0,\gamma^m}\right)^{s_0+1} \left(a(\gamma, .)\right) \frac{d\nu}{\nu} \frac{d\gamma}{\gamma} \right) \right)^2 (x)$$

$$\leq \sum_{k=0}^{s_0+1} c_k \int_{\Gamma(x)} \left[\int_0^\infty \int_0^{r_B} \mathrm{h}_{t,\nu,\gamma,k}^{m,\alpha,s} \psi_{t,\nu,\gamma}^k (L,s)\big(a(\gamma,.)\big)(y)\, \frac{d\nu d\gamma}{\nu\gamma}\right]^2 \frac{dydt}{t^{n+1}}$$

$$\leq c \int_0^{r_B} \int_{|y-x|<t}$$

$$\times \left[\int_0^{r_B} \int_{\hat{B}} \frac{t^m \nu^{m\alpha} \gamma^{m(s+1)} (t+\nu+\gamma)^\epsilon}{(t+\nu+\gamma)^{m(s+2)} (t+\nu+\gamma+|y-\mathcal{Z}|)^{n+\epsilon}} |a(\gamma,\mathcal{Z})| \frac{d\mathcal{Z} d\nu d\gamma}{\nu\gamma}\right]^2 \frac{dydt}{t^{n+1}}$$

$$+c \int_0^{r_B} \int_{|y-x|<t}$$

$$\times \left[\int_{r_B}^{\infty} \int_{\hat{B}} \frac{t^m \nu^{m\alpha} \gamma^{m(s+1)} (t+\nu+\gamma)^\epsilon}{(t+\nu+\gamma)^{m(s+2)} (t+\nu+\gamma+|y-\mathcal{Z}|)^{n+\epsilon}} |a(\gamma,\mathcal{Z})| \frac{d\mathcal{Z} d\nu d\gamma}{\nu\gamma}\right]^2 \frac{dydt}{t^{n+1}}$$

$$+c \int_{r_B}^{\infty} \int_{|y-x|<t}$$

$$\times \left[\int_0^{r_B} \int_{\hat{B}} \frac{t^m \nu^{m\alpha} \gamma^{m(s+1)} (t+\nu+\gamma)^\epsilon}{(t+\nu+\gamma)^{m(s+2)} (t+\nu+\gamma+|y-\mathcal{Z}|)^{n+\epsilon}} |a(\gamma,\mathcal{Z})| \frac{d\mathcal{Z} d\nu d\gamma}{\nu\gamma}\right]^2 \frac{dydt}{t^{n+1}}$$

$$+c \int_{r_B}^{\infty} \int_{|y-x|<t}$$

$$\times \left[\int_{r_B}^{\infty} \int_{\hat{B}} \frac{t^m \nu^{m\alpha} \gamma^{m(s+1)} (t+\nu+\gamma)^\epsilon}{(t+\nu+\gamma)^{m(s+2)} (t+\nu+\gamma+|y-\mathcal{Z}|)^{n+\epsilon}} |a(\gamma,\mathcal{Z})| \frac{d\mathcal{Z} d\nu d\gamma}{\nu\gamma}\right]^2 \frac{dydt}{t^{n+1}}$$

$$= II_1 + II_2 + II_3 + II_4.$$

Let us estimate the term II_4 . The same argument as in(31) shows that for $x \notin 4B, t \geq r_B, \mathcal{Z} \in B_2$ and $|y-x| < t$, we have $(t+s+|y-\mathcal{Z}|) \geq c|x-\mathcal{Z}_0|$. Those, together with the fact that

$$\frac{t^m \nu^{m\alpha} \gamma^{m(s+1)} (t+\nu+\gamma)^\epsilon}{(t+\nu+\gamma)^{m(s+2)}} \leq (t\nu)^{-\beta} \gamma^{\epsilon+m\alpha+2\beta},$$

where $\beta = (m(s+1) - m\alpha - \epsilon)/2 > 0$, show that

$$II_4 \leq c \int_{r_B}^{\infty} \left[\int_{r_B}^{\infty} \int_{\hat{B}} \frac{1}{\nu^\beta} \gamma^{\epsilon+m\alpha+2\beta} |a(\gamma,\mathcal{Z})| \frac{d\mathcal{Z} d\nu d\gamma}{\nu\gamma}\right]^2 \frac{dydt}{t^{1+2\beta}} |x-\mathcal{Z}_0|^{-2(n+\epsilon)}$$

$$\leq c|B| \int_{r_B}^{\infty} \left[\int_{r_B}^{\infty} \left(\int_0^{r_B} \gamma^{2(\epsilon+m\alpha+2\beta)} \frac{d\gamma}{\gamma}\right)^{1/2} \frac{d\nu}{\nu^{1+\beta}}\right]^2 \frac{dt}{t^{1+2\beta}} |||a|||_{T_2^2}^2$$

$$\times |x-\mathcal{Z}_0|^{-2(n+\epsilon)}$$

$$\leq c r_B^{2\left(n+\epsilon+m\alpha-\frac{n}{p_1}+2\beta\right)} \int_{r_B}^{\infty} \left[\int_{r_B}^{\infty} \frac{d\nu}{\nu^{1+\beta}}\right]^2 \frac{dt}{t^{1+2\beta}} |x-\mathcal{Z}_0|^{-2(n+\epsilon)}$$

$$\leq c r_B^{2(\epsilon+m\alpha)+2\beta-\frac{2n}{p_1}} |x-\mathcal{Z}_0|^{-2(n+\epsilon)}.$$

The same argument as above shows that $I_1 + II_2 + II_3 \leq c r_B^{2(\epsilon+m\alpha)+2n-\frac{2n}{p_1}} \times |x-z_0|^{-2(n+\epsilon)}$. This proves (45), and gives the desired estimate

$$\int_{(4B)^c} \left| S_L\big(L^{-\alpha}(b)\big)(x)\right|^{p_2} dx \le c.$$

The proof of Case III is obtained. Hence, the proof of Theorem (3.1.12) is complete.

Assume that L is a linear operator of type ω on $L^2(\mathbb{R}^n)$ with $\omega < \pi/2$, hence $(-L)$ generates an analytic semigroup e^{-zL}, $0 \le |\mathrm{Arg}(z)| < \pi/2 - \omega$. We assume that for each $t > 0$, the kernel $p_t(x,y)$ of e^{-tL} is a Hölder continuous function in x, y and there exist positive constants m and $0 < \gamma \le 1$ such that for all $t > 0$, and $x, y, h \in \mathbb{R}^n$,

$$|p_t(x,y)| \le c\frac{t^{1/m}}{(t^{1/m} + |x-y|)^{n+1}}, \tag{46}$$

$$|p_t(x+h,y) - p_t(x,y)| + |p_t(x,y+h) - p_t(x,y)| \le c|h|^\gamma \frac{t^{1/m}}{(t^{1/m} + |x-y|)^{n+1+\gamma}} \tag{47}$$

whenever $2|h| \le t^{1/m} + |x-y|$; and

$$\int_{\mathbb{R}^n} p_t(x,y)dx = \int_{\mathbb{R}^n} p_t(x,y)dy = 1, \qquad \forall t > 0. \tag{48}$$

We have the following equivalence between the $H^p(\mathbb{R}^n)$ spaces and the $H_L^p(\mathbb{R}^n)$ spaces associated with operators.

Theorem (3.1.13): *Assume that L satisfies the assumptions* (46), (47) *and* (48). *Then for* $\frac{n}{n+\gamma} < p \le 1$, *the spaces* $H^p(\mathbb{R}^n)$ *and* $H_L^p(\mathbb{R}^n)$ *coincide, and their quasinorms are equivalent.*

As a consequence, for $0 < \alpha < \frac{\gamma}{n}$, *the classical Morrey-Campanato spaces* $L(\alpha, 2, 0)$ *and the spaces* $\mathfrak{L}_L(\alpha, 2, 0)$ *coincide, and their norms are equivalent.*

Proof: We remark that for L satisfying (46), (47) and (48), our proof below shows that L has a bounded holomorphic functional calculus on L^2 because the area integral functions S_L and S_{L^*} are bounded on L^2 in which L^* is the adjoint operator of L. See Theorem 3 of [81].

Let $q_t(x,y)$ denote the kernel of the operator $Q_t = t\frac{d}{dt}e^{-tL}$. From Lemma 6.10 of [71], we have the following estimates: for any $0 < \gamma_1 < \gamma$ and $0 < \beta_1 < 1$, there exist constant $c > 0$ such that for all $t > 0$, and $x, y, h \in \mathbb{R}^n$,

$$|q_t(x,y)| \le c\frac{t^{\beta_1/m}}{(t^{1/m} + |x-y|)^{n+\beta_1}},$$

$$|q_t(x+h,y) - q_t(x,y)| + |q_t(x,y+h) - q_t(x,y)| \le c|h|^{\gamma_1} \frac{t^{\beta_1/m}}{(t^{1/m} + |x-y|)^{n+\beta_1+\gamma_1}}$$

whenever $2|h| \le t^{1/m} + |x-y|$; and

$$\int_{\mathbb{R}^n} q_t(x,y)dx = \int_{\mathbb{R}^n} q_t(x,y)dy = 0, \qquad \forall t > 0.$$

It then follows from a standard harmonic analysis argument that $H_L^p(\mathbb{R}^n) = H^p(\mathbb{R}^n)$ for $\frac{n}{n+\gamma} < p \le 1$. See, for example, Chapter XIV, in [85]. Hence, the proof of Theorem (3.1.13) is complete.

To begin with, let us recall some basic facts about the Neumann Laplacian Δ_N on $\mathbb{R}^n$, which was studied in [65]. In what follows, $\mathbb{R}^n_+$ denotes the upper-half space in $\mathbb{R}^n$, i.e.,

$$\mathbb{R}^n_+ = \{(x', x_n) | \in \mathbb{R}^n : x' = (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}, x_n > 0\}.$$

Similarly, $\mathbb{R}^n_-$ denotes the lower-half space in $\mathbb{R}^n$.

We denote by Δ_{N_+} (resp. Δ_{N_-}) the Neumann Laplacian on $\mathbb{R}^n_+$ (resp. on $\mathbb{R}^n_-$). See page 57 of [83]. The Neumann Laplacians are self-adjoint and positive definite operators. Using the spectral theory one can define the semigroup $\{\exp(-t\Delta_{N_+})\}_{t\geq 0}$ (resp. $\{\exp(-t\Delta_{N_-})\}_{t\geq 0}$) generated by the operator Δ_{N_+} (resp. Δ_{N_-}). For any f defined on $\mathbb{R}^n$, we set

$$f_- = f|_{\mathbb{R}^n_-} \text{ and } f_+ = f|_{\mathbb{R}^n_+},$$

where $f|_{\mathbb{R}^n_+}$ and $f|_{\mathbb{R}^n_-}$ are restrictions of the function f to $\mathbb{R}^n_+$ and $\mathbb{R}^n_-$, respectively.

We let Δ_N be the uniquely determined unbounded operator acting on $L^2(\mathbb{R}^n)$ such that

$$(\Delta_N f)_+ = \Delta_{N_+} f_+ \text{ and } (\Delta_N f)_- = \Delta_{N_-} f_- \quad (49)$$

for all $f: \mathbb{R}^n \mapsto \mathbb{R}$ such that $f_+ \in W^{1,2}(\mathbb{R}^n_+)$ and $f_- \in W^{1,2}(\mathbb{R}^n_-)$. Then, Δ_N is a positive definite self-adjoint operator and let $p_t(x, y)$ be the heat kernel of the semigroup $\exp(-t\Delta_N)$. By (49), we have

$$(\exp(-t\Delta_N) f)_+ = \exp(-t\Delta_{N_+}) f_+ \text{ and } (\exp(-t\Delta_N) f)_- = \exp(-t\Delta_{N_-}) f_- \quad (50)$$

Moreover, we have

$$p_t(x, y) = \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x'-y'|^2}{4t}} \left(e^{-\frac{|x_n - y_n|^2}{4t}} + e^{-\frac{|x_n + y_n|^2}{4t}} \right) H(x_n, y_n), \quad (51)$$

where $H: \mathbb{R} \to \{0, 1\}$ is the Heaviside function, given by $H(t) = 0$ if $t < 0$; 1 if $t \geq 0$. See [65].

We show the following proposition.

Proposition (3.1.14)[53]: *Suppose* $\frac{n}{n+1} < p \leq 1$. *The operator* Δ_N *satisfies the assumptions* (a) *and* (b). *Moreover, we have* $H^p_{\Delta_N}(\mathbb{R}^n) \underset{\neq}{\subset} H^p(\mathbb{R}^n)$. *That is,* $H^p_{\Delta_N}(\mathbb{R}^n)$ *is a proper subspace of the classical* $H^p(\mathbb{R}^n)$ *space.*

Proof: Since Δ_N is a self-adjoint positive definite operator, hence it has a bounded H_∞-calculus in $L^2(\mathbb{R}^n)$. From the equation (50), Δ_N generates the conservative semigroup $e^{-t\Delta_N}$ that is $e^{-t\Delta_N}(1) = 1$ for all $t > 0$, which satisfies the assumptions (a) and (b). This gives that

$$H^p_{\Delta_N}(\mathbb{R}^n) \subseteq H^p(\mathbb{R}^n), \qquad \frac{n}{n+1} < p \leq 1,$$

On the other hand, from Theorem 4.1 of [65] and Proposition 5.3 of [72], this operator Δ_N generates the spaces $\mathfrak{L}_{\Delta_N}(\alpha, 2, 0)$ with $0 \leq \alpha < n^{-1}$ such that

(i) $L(\alpha, 2, 0) \subseteq \mathfrak{L}_{\Delta_N}(\alpha, 2, 0)$.

(ii) We have $f(x) = \log|x| \, \chi_{\{x: x \in \mathbb{R}^n_+\}}(x) \in BMO_{\Delta_N}(\mathbb{R}^n)$, however, $f \notin BMO(\mathbb{R}^n)$.

(iii) For $0 \leq \alpha < n^{-1}$, we have $f(x) = e^{-|x|^2} \chi_{\{x: x \in \mathbb{R}^n_+\}}(x) \in \mathfrak{L}_{\Delta_N}(\alpha, 2, 0)$, however, $f \notin L(\alpha, 2, 0)$.

From the properties (i), (ii) and (iii) above, we have that $\mathfrak{L}(\alpha, 2, 0)$ is a proper subspace of $\mathfrak{L}_{\Delta_N}(\alpha, 2, 0)$. Proposition (3.1.14) then follows from Theorem (3.1.10) and the fact that $(H^p)' = L\left(\frac{1}{p} - 1, 2, 0\right)$ (see, for example, Theorem 2.7 of [86]) (see [89]).

Section (3.2): The BMO Spaces Associated with Operators and Applications

The classical space of functions of bounded mean oscillation (BMO) plays a crucial role in modern harmonic analysis. See for examples [74], [75], [82] and [19]. In the case of the Euclidean space $\mathbb{R}^n$, a function f is said to in $BMO(\mathbb{R}^n)$ if

$$\|f\|_{BMO(\mathbb{R}^n)} = \sup_Q \frac{1}{|Q|}\int_Q |f(x) - f_Q|\, dx < \infty. \qquad (52)$$

where f_Q denotes the average value of f on the cube Q and the supremum is taken over all cubes Q in $\mathbb{R}^n$.

An important application of the theory of BMO spaces is the following interpolation result.

Proposition (3.2.1)[65]: If T is a bounded sublinear operator from $L^2(\mathbb{R}^n)$to $L^2(\mathbb{R}^n)$, and T is bounded from $L^\infty(\mathbb{R}^n)$ to $BMO(\mathbb{R}^n)$, then T is bounded from $L^p(\mathbb{R}^n)$to$L^p(\mathbb{R}^n)$ for all $2 < p < \infty$.

It is well known that Caldero´n-Zygmund operators (such as the Hilbert transform on the real line, the Riesz transforms on $\mathbb{R}^n$, or the purely imaginary powers of the Laplacian on $\mathbb{R}^n$) do not map the space L^∞into L^∞, but the standard conditions on their kernels ensure that they map L^∞ into the BMO space boundedly, hence we can apply Proposition(3.2.1) to obtain L^p boundedness of these operators for $p > 2$. The BMO space is a natural substitute of the space L^∞in the theory of Calder´on-Zygmund singular integrals.

We study of singular integral operators corresponding to spectral multiplier of an operator L which generates a semigroup with appropriate kernel bounds, see[70]. Such multipliers do not always map L^∞ or appropriate L^p spaces into the classical BMO space, see Example (3.2.18) below. Hence the classical BMO space is not necessarily a suitable space to study such singular integrals. To study these rough operators, we introduced a newBMO_Lspace associated with an operator L. To explain our approach to BMO_L space associated with an operator let us recall that the space of BMO functions can be characterized by the Carleson measure estimate as follows:

Proposition (3.2.2)[65]: A function f is in BMO if and only if f satisfies

$$\int_{\mathbb{R}^n} \frac{|f(x)|}{1 + |x|^{n+1}}\, dx < \infty,$$

and

$$\mu_f(x,t) = \left| t\frac{\partial}{\partial t} e^{-t\sqrt{\Delta}} f(x)\right|^2 \frac{dxdt}{t}$$

is a Carleson measure.

One can see from the characterization in Proposition(3.2.2) that the BMO space is associated with the Laplace operator on $\mathbb{R}^n$and it seems to be natural idea to replace the Laplace operator Δ by more general operators operator L, see also [74] and[29].

We use equivalent approach, see Definition(3.2.5) below. In this definition the BMO_L space associated with L is defined by using the function $e^{-t_Q L}f$to replace the average f_Q in Definition(3.2.4) of BMO where the value t_Q is scaled to the length of the sides of Q. We discuss various examples which shows that Definition(3.2.5) is an effective tool in study of singular integrals operators associated with the operator L. See [58], [95] and[99] for other ideas related to generalization of the BMO space and BMO spaces associated with an operator$L$.

Many important features of the classical BMO space are retained by the new BMO_Lspaces such as the John-Nirenberg inequality and duality between the Hardy space and the BMO_L

space. See[70] and[71]. One of these important features is that the interpolation property in Proposition(3.2.1) is still valid if the classical space BMO is replaced by theBMO_Lspace associated with an operator$L$. Indeed, the following result is proved in[70] (Theorem(3.2.19)).

Proposition (3.2.3)[65]: Let χ be a space of homogeneous type. If T is a bounded sublinear operator from $L^2(\chi)$ to $L^2(\chi)$, and T is bounded from $L^\infty(\chi)$into $BMO_L(\chi)$, then T is bounded from $L^p(\chi)$ to $L^p(\chi)$for all $2 < p < \infty$.

A natural question arising from Proposition(3.2.3) is to compare the classical BMO space and the BMO_L space associated with an operator L. We study this question systematically and we show that depending on the choice of the operator L, all the following cases are possible

Case1: $BMO \cong BMO_L$;

Case2: $BMO \subseteq BMO_L$ and$BMO \neq BMO_L$;

Case3: $BMO_L \subseteq BMO$and$BMO_L \neq BMO$;

Case4: $BMO \nsubseteq BMO_L$ and $BMO_L \nsubseteq BMO$.

For other results related to Cases 1 and 2 see Proposition 2.5 of [70], [71] and Proposition(3.2.8) of[78]. We show that if $f \in L^{n/\alpha}(\mathbb{R}^n)$ and $L^{-\alpha}f < \infty$almost everywhere then$L^{-\alpha}f \in BMO_L$. We construct an example of a function $f \in L^p(\mathbb{R})$ and an operator L such that $L^{-\frac{1}{2p}}f \in BMO_L$ but $L^{-\frac{1}{2p}}f \notin BMO$. This shows that the new BMO_L space does make a difference in estimates of singular integrals. We obtain sharp estimates of the L^∞ to BMO_L norm of the purely imaginary powers L^{is} of a self adjoint operator L. We also obtain the BMO type estimates for spectral multipliers of a self adjoint operator L and for maximal operators $\sup_{t>0}|F(tL)|$ corresponding to L and appropriate functions F. L^pboundedness of these operators,$2 < p < \infty$, then follows from Proposition(3.2.3).

We begin by recalling the definitions of various BMO spaces on the usual upper-half space in $\mathbb{R}^n$. For any subset $A \subset \mathbb{R}^n$ and a function $f: \mathbb{R}^n \to \mathbb{C}$ by $f|_A$we denote the restriction of f to the set A. Next we set

$$\mathbb{R}^n_+ = \{(x', x_n) \in \mathbb{R}^n : x' = (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}, x_n > 0\}.$$

Definition (3.2.4)[65]: A function f on $\mathbb{R}^n_+$is said to be in $BMO_r(\mathbb{R}^n_+)$ if there exists $F \in BMO(\mathbb{R}^n)$ such that$F|_{\mathbb{R}^n_+} = f$. If $f \in BMO_r(\mathbb{R}^n_+)$, we set

$$\|f\|_{BMO_r(\mathbb{R}^n_+)} = \inf\{\|F\|_{BMO(\mathbb{R}^n_+)} : F|_{\mathbb{R}^n_+} = f\}.$$

A function f on $\mathbb{R}^n_+$belongs to $BMO_{\mathcal{Z}}(\mathbb{R}^n_+)$ if the function F defined by

$$F(x) = \begin{cases} f(x) \text{if} x \in \mathbb{R}^n_+; \\ 0 \quad \text{if} x \notin \mathbb{R}^n_+, \end{cases} \tag{53}$$

belongs to $BMO(\mathbb{R}^n)$. If $f \in BMO_{\mathcal{Z}}(\mathbb{R}^n_+)$, we set$\|f\|_{BMO_{\mathcal{Z}}(\mathbb{R}^n_+)} = \|F\|_{BMO_{\mathcal{Z}}(\mathbb{R}^n_+)}$. Compare Section4.5.1, page221 of [105] and Section5.4 of[93]. In order to analyze the spaces $BMO_r(\mathbb{R}^n_+)$ and $BMO_{\mathcal{Z}}(\mathbb{R}^n_+)$, let us introduce the following notations, see[61]. For any $x = (x', x_n) \in \mathbb{R}^n$, we set$\tilde{x}(x', x_n)$. If f is any function defined on $\mathbb{R}^n_+$, its even extension f_e is defined on$\mathbb{R}^n$by

$$f_e(x) = \begin{cases} f(x) \text{if} x \in \mathbb{R}^n_+; \\ f(\tilde{x}) \text{if} x \in \mathbb{R}^n_-, \end{cases}$$

and its odd extension$f_o$is defined by

$$f_o(x) = \begin{cases} f(x) \text{ if} x \in \mathbb{R}^n_+; \\ -f(\tilde{x}) \text{if} x \in \mathbb{R}^n_-, \end{cases}$$

Where

$$\mathbb{R}^n_- = \{(x', x_n) \in \mathbb{R}^n : x' = (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}, x_n < 0\}.$$

For any function $f \in L^1_{loc}(\mathbb{R}^n_+)$, we define

$$\|f\|_{BMO_e(\mathbb{R}^n_+)} = \|f_e\|_{BMO(\mathbb{R}^n)} \text{ and } \|f\|_{BMO_o(\mathbb{R}^n_+)} = \|f_o\|_{BMO(\mathbb{R}^n)}$$

and we denote by $BMO_e(\mathbb{R}^n_+)$ and $BMO_o(\mathbb{R}^n_+)$ the corresponding Banach spaces.

We will see that $BMO_e(\mathbb{R}^n_+)$ is suitable for the analysis of the Neumann Laplacian on $\mathbb{R}^n_+$ whereas $BMO_o(\mathbb{R}^n_+)$ is suitable for the study of the Dirichlet Laplacian on $\mathbb{R}^n_+$. See Proposition(3.2.9) below.

In what follows, $Q = Q[x_Q, l_Q]$ denotes a cube of $\mathbb{R}^n$ centered at x_Q and of the side length l_Q. Given any cube Q, we denote the reflection of Q across $\partial\mathbb{R}^n_+$ by

$$\tilde{Q} = \{(x', x_n) \in \mathbb{R}^n, (x', -x_n) \in Q\}. \tag{54}$$

Let $Q_+ = Q \cap \mathbb{R}^n_+$ and $Q_- = Q \cap \overline{\mathbb{R}^n_-}$ where $\overline{\mathbb{R}^n_-} = \{(x', x_n) \in \mathbb{R}^n : x' = (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}, x_n < 0\}$. If both Q_- and Q_+ are not empty, we then define

$$\begin{cases} \hat{Q}_- = \{(x', x_n) : x' \in Q \cap \mathbb{R}^{n-1}, -l_Q < x_n \le 0\}, \\ \hat{Q}_+ = \{(x', x_n) : x' \in Q \cap \mathbb{R}^{n-1}, 0 < x_n \le l_Q\}. \end{cases} \tag{55}$$

Obviously, we have the following properties:

(i) $Q_- \subseteq \hat{Q}_-$, $Q_+ \subseteq \hat{Q}_+$ and thus $Q \subseteq (\hat{Q}_- \cup \hat{Q}_+)$;

(ii) $|Q| = |\hat{Q}_-| = |\hat{Q}_+|$.

These will be often used in the sequel.

By Δ_{n,N_+} (and Δ_{n,N_-}) we denote the Neumann Laplacian on $\mathbb{R}^n_+$ (and on $\mathbb{R}^n_-$ respectively). Similarly by Δ_{n,D_+} (and Δ_{n,D_-}) we denote the Dirichlet Laplacian on $\mathbb{R}^n_+$ (and on $\mathbb{R}^n_-$ respectively).

The Dirichlet and Neumann Laplacian are positive definite self-adjoint operators. By the spectral theorem one can define the semigroups generated by these operators $\{\exp(-t\Delta_{n,D_+}) : t \ge 0\}$ and $\{\exp(-t\Delta_{n,N_+}) : t \ge 0\}$. By $p_{t,\Delta_{n,D_+}}(x, y)$ and $p_{t,\Delta_{n,N_+}}(x, y)$ we denote the heat kernels corresponding to the semigroups generated by Δ_{n,D_+} and Δ_{n,N_+} respectively.

For $n = 1$ by the reflection method (see for example [83]) we obtain

$$p_{t,\Delta_{n,D_+}}(x, y) = \frac{1}{(4\pi t)^{1/2}}\left(e^{-\frac{|x_1 - y_1|^2}{4t}} - e^{-\frac{|x_1 - y_1|^2}{4t}}\right).$$

Then for $n \ge 2$

$$\begin{aligned} p_{t,\Delta_{n,D_+}}(x, y) &= \left(p_{t,\Delta_{1,D_+}}(x_n, y_n)\right)\left(p_{t,\Delta_{n-1}}(x', y')\right) \\ &= \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x', y'|^2}{4t}}\left(e^{-\frac{|x_n - y_n|^2}{4t}} - e^{-\frac{|x_n - y_n|^2}{4t}}\right), \end{aligned} \tag{56}$$

where $p_{t,\Delta_{n-1}}(x, y)$ is the heat kernel corresponding to the standard Laplace operator acting on $\mathbb{R}^{n-1}$. Applying the reflection method also to the Neumann Laplacian we obtain (see [83])

$$\begin{aligned} p_{t,\Delta_{n,N_+}}(x, y) &= \left(p_{t,\Delta_{1,N_+}}(x_n, y_n)\right)\left(p_{t,\Delta_{n-1}}(x', y')\right) \\ &= \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x', y'|^2}{4t}}\left(e^{-\frac{|x_n - y_n|^2}{4t}} - e^{-\frac{|x_n - y_n|^2}{4t}}\right). \end{aligned} \tag{57}$$

We skip the index n and we denote the Dirichlet and Neumann Laplacian byΔ_{D_+}andΔ_{N_+}. Note that by(56)

$$\exp\left(-t\Delta_{D_+}\right)f(x) = \int_{\mathbb{R}^n_+} p_{t,\Delta_{D_+}}(x,y)f(y)\,dy$$
$$= \frac{1}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} f_o(y)dy$$
$$= \exp(-t\Delta)f_o(x)\ (58)$$

for $x \in \mathbb{R}^n_+$and all $t > 0$.Similarly

$$\exp\left(-t\Delta_{N_+}\right)f(x) = \int_{\mathbb{R}^n_+} p_{t,\Delta_{N_+}}(x,y)f(y)\,dy = \frac{1}{(4\pi t)^{\frac{n}{2}}}\int_{\mathbb{R}^n} e^{-\frac{|x-y|^2}{4t}} f_e(y)dy$$
$$= \exp(-t\Delta)f_e(x) \qquad (59)$$

for $x \in \mathbb{R}^n_+$and all $t > 0$.
Next for any function f on $\mathbb{R}^n$, we set

$$f_- = f|_{\mathbb{R}^n_-}\text{and } f_+ = f|_{\mathbb{R}^n_+}.$$

Now let Δ_N be the uniquely determined unbounded operator acting on $L^2(\mathbb{R}^n)$such that

$$(\Delta_N f)_+ = \Delta_{N_+}f_+\text{and}(\Delta_N f)_- = \Delta_{N_-}f_- \qquad (60)$$

for all $f: \mathbb{R}^n \to \mathbb{R}$ such that $f_+ \in W^{1,2}(\mathbb{R}^n_+)andf_- \in W^{1,2}(\mathbb{R}^n_-)$. Then, $\Delta_N$is a positive definite self-adjoint operator. By(60)

$$(\exp(-t\Delta_N)f)_+ = \exp\left(-t\Delta_{N_+}\right)f_+$$
$$\text{and } (\exp(-t\Delta_N)f)_- = \exp\left(-t\Delta_{N_-}\right)f_-. \qquad (61)$$

Let $p_{t,\Delta_N}(x,y)$be the heat kernel of $\exp(-t\Delta_N)$. By(61) and(57) we obtain

$$p_{t,\Delta_N}(x,y) = \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x',y'|^2}{4t}}\left(e^{-\frac{|x_n-y_n|^2}{4t}} + e^{-\frac{|x_n-y_n|^2}{4t}}\right)H(x_n-y_n), \qquad (62)$$

where $H: \mathbb{R} \to \{0,1\}$ is the Heaviside function given by

$$H(t) = \begin{cases} 0 \text{ if } t < 0; \\ 0 \text{ if } t \geq 0. \end{cases} \qquad (63)$$

Similarly we define the Dirichlet Laplacian on $\mathbb{R}^n$by the formula

$$(\Delta_D f)_+ = \Delta_{D_+}f_+\text{and}(\Delta_D f)_- = \Delta_{D_-}f_- \qquad (64)$$

for all $f: \mathbb{R}^n \to \mathbb{R}$ such that $f_+ \in W_0^{1,2}(\mathbb{R}^n_+)$ and $f_- \in W_0^{1,2}(\mathbb{R}^n_-)$. Then, $\Delta_D$is a positive definite self-adjoint operator. By(64)

$$(\exp(-t\Delta_D)f)_+ = \exp\left(-t\Delta_{D_+}\right)f_+$$
$$\text{and } (\exp(-t\Delta_D)f)_- = \exp\left(-t\Delta_{D_-}\right)f_-. \qquad (65)$$

Hence by (56) the kernel $p_{t,\Delta_D}(x,y)$of the operator $\exp(-t\Delta_D)$is given by

$$p_{t,\Delta_D}(x,y) = \frac{1}{(4\pi t)^{n/2}} e^{-\frac{|x',y'|^2}{4t}}\left(e^{-\frac{|x_n-y_n|^2}{4t}} + e^{-\frac{|x_n+y_n|^2}{4t}}\right)H(x_n-y_n). \qquad (66)$$

Finally we define the Dirichlet-Neumann Laplacian by the formula

$$(\Delta_{DN} f)_+ = \Delta_{N_+}f_+\text{and}(\Delta_{DN} f)_- = \Delta_{D_-}f_- \qquad (67)$$

for all $f: \mathbb{R}^n \to \mathbb{R}$ such that $f_+ \in W^{1,2}(\mathbb{R}^n_+)$and $f_- \in W_0^{1,2}(\mathbb{R}^n_-)$. By (67)

$$(\exp(-t\Delta_{DN})f)_+ = \exp\left(-t\Delta_{N_+}\right)f_+$$
$$\text{and } (\exp(-t\Delta_{DN})f)_- = \exp\left(-t\Delta_{D_-}\right)f_-. \qquad (68)$$

Hence by (56) and(57), the kernel $p_{t,\Delta_{DN}}(x,y)$of$\exp(-t\Delta_{DN})$is given by

$$p_{t,\Delta_{DN}}(x,y)=\frac{1}{(4\pi t)^{n/2}}e^{-\frac{|x',y'|^2}{4t}}\left(e^{-\frac{|x_n-y_n|^2}{4t}}+(2H(x_n)-1)e^{-\frac{|x_n+y_n|^2}{4t}}\right)H(x_n-y_n).\ (69)$$

Let us note that

(α) All the operators $\Delta,\Delta_{N_+},\Delta_{D_+},\Delta_{N_-},\Delta_{D_-}$ and $\Delta_D,\Delta_N,\Delta_{DN}$ are self-adjoint and they generate bounded analytic positive semigroups acting on all L^p spaces for $1\le p\le\infty$;

(β) Suppose that $p_{t,L}(x,y)$ is the kernel corresponding to the semigroup generated by L and that L is one of the operators listed in (α). Then the kernel $p_{t,L}(x,y)$ satisfies Gaussian bounds, that is

$$\left|p_{t,L}(x,y)\right|\le\frac{C}{t^{n/2}}e^{-c\frac{|x-y|^2}{t}} \tag{70}$$

for all $x,y\in\Omega$, where $\Omega=\mathbb{R}^n$ for $\Delta,\Delta_D,\Delta_N,\Delta_{DN}$; $\Omega=\mathbb{R}^n_+$ for $\Delta_{N_+},\Delta_{D_+}$ and $\Omega=\mathbb{R}^n_-$ for $\Delta_{N_-},\Delta_{D_-}$.

(γ) If L is one of the operators $\Delta,\Delta_{N_+},\Delta_{N_-}$ and Δ_N, then L conserves probability, that is

$$\exp(-tL)\,\mathbb{1}=\mathbb{1}.$$

This conservative property does not hold for $\Delta_D,\Delta_{D_+},\Delta_{D_-}$ and Δ_{DN}.

Suppose that $\Omega\subset\mathbb{R}^n$ is an open subset of $\mathbb{R}^n$. Suppose that L is a linear operator on $L^2(\Omega)$ which generates an analytic semigroup e^{-tL} with a kernel $p_t(x,y)$ satisfying Gaussian upper bound (70).

We define

$$\mathcal{M}(\Omega)=\left\{f\in L^1_{loc}(\Omega):\exists d>0,\int_\Omega\frac{|f(x)|^2}{1+|x|^{n+d}}dx<\infty\right\}.$$

Note that in virtue of the Gaussian bounds (70) we can extend the action of the semigroup operators $\exp(-tL)$ to the space $\mathcal{M}(\Omega)$, that is we can define $\exp(-tL)f$ for all $f\in\mathcal{M}(\Omega)$. By $B(x,r)$ we denote the ball in Ω with respect to the Euclidean distance restricted to Ω that is

$$B(x,r)=\{y\in\Omega:|x-y|<r\}.$$

The following $BMO_L(\Omega)$ space associated with an operator L was introduced in [70].

Definition (3.2.5)[65]: We say that $f\in\mathcal{M}(\Omega)$ is of bounded mean oscillation associated with an operator L (abbreviated as $BMO_L(\Omega)$) if

$$\|f\|_{BMO_L(\Omega)}=\sup_{B(y,r)}\frac{1}{|B(y,r)|}\int_{B(y,r)}|f(x)-\exp(-r^2L)f(x)|\,dx<\infty, \tag{71}$$

where the supremum is taken over all balls $B(y,r)$ in Ω. The smallest bound for which (71) is satisfied is then taken to be the norm of f in this space, and is denoted by $\|f\|_{BMO_L(\Omega)}$.

Remark (3.2.6)[65]: (i) Note that $\left(BMO_L(\Omega),\|\cdot\|_{BMO_L(\Omega)}\right)$ is a semi-normed vector space, with the semi-norm vanishing on the kernel space $\mathcal{K}_L$ defined by

$$\mathcal{K}_L=\{f\in\mathcal{M}(\Omega);\exp(-tL)f=f,\forall t>0\}.$$

The class of functions of $BMO_L(\Omega)$ (modulo $\mathcal{K}_L$) is a Banach space. See Section 6 of [71] for a discussion on the dimension of the space $\mathcal{K}_L$ of $BMO_L(\mathbb{R}^n)$ when L is a second order divergence form elliptic operator or a Schrödinger operator. In the sequel by $BMO_L(\Omega)$ we always denote the space $BMO_L(\Omega)$(modulo $\mathcal{K}_L$) and we skip (modulo $\mathcal{K}_L$) to simplify notation.

(ii) Similarly to the classical BMO space, it is easy to check that $L^\infty(\Omega) \subset BMO_L(\Omega)$ with $\|f\|_{BMO_L(\Omega)} \le 2\|f\|_{L^\infty}$.
(iii) The classical BMO space (modulo all constant functions) and the $BMO_\Delta(\mathbb{R}^n)$ space (modulo all harmonic functions) coincide, and their norms are equivalent. See [15].
(iv) Note that the Euclidean distance in Definition(3.2.5) can be replaced by any equivalent distance. That is if there exists $c > 0$ such that $c^{-1}|x-y| \le d(x,y) \le c|x-y|$then one can take in(71) the supremum over all balls$B^d(x,r)$ with respect to the metric$d$. In particular if$\Omega = \mathbb{R}^n$,$\Omega = \mathbb{R}^n_+$or$\Omega = \mathbb{R}^n_-$, one can take the supremum over all cubes Q such that$Q \subset \Omega$ in(71), i.e., we can define equivalent norm in$BMO_L(\Omega)$by the formula

$$\|f\|_{BMO_L(\mathbb{R}^n)} = \sup_Q \frac{1}{|Q|}\int_Q \left|f(x) - \exp\left(-l_Q^2 L\right) f(x)\right| dx < \infty, \tag{72}$$

where$l_Q$is the side length of Q and the supremum is taken over all cubes $Q \subset \Omega$. The following proposition is essentially equivalent to Proposition(3.2.8) of [78].

Proposition (3.2.7)[65]: Assume that for every $t > 0$, $e^{-tL}(\mathbb{1}) = \mathbb{1}$almost every-where, that is,$\int_{\mathbb{R}^n} p_{t,L}(x,y)\,dy = 1$for almost all $x \in \mathbb{R}^n$. Then, we have$BMO(\mathbb{R}^n) \subset BMO_L(\mathbb{R}^n)$, and there exists a positive constant $c > 0$ such that

$$\|f\|_{BMO_L(\mathbb{R}^n)} \le c\|f\|_{BMO(\mathbb{R}^n)}. \tag{73}$$

However, the converse inequality does not hold in general.
We remark that condition $e^{-tL}(\mathbb{1}) = \mathbb{1}$, is necessary for(73). Indeed, (73) implies $\|\mathbb{1}\|_{BMO_L(\mathbb{R}^n)} = 0$. Hence $e^{-tL}(\mathbb{1}) = \mathbb{1}$ almost everywhere for all$t > 0$,

We describe the equivalence between the BMO spaces on the half space and BMO spaces corresponding to the Neumann and Dirichlet Laplacian.

Proposition (3.2.8)[65]: (i) The spaces $BMO_r(\mathbb{R}^n_+)$and $BMO_e(\mathbb{R}^n_+)$coincide, and their norms are equivalent.
(ii) The spaces $BMO_{\mathcal{Z}}(\mathbb{R}^n_+)$ and $BMO_o(\mathbb{R}^n_+)$coincide, and their norms are equivalent.

Proof: Following[61], for any function $f \in L^1(\mathbb{R}^n_+)$we set

$$\|f\|_{H^1_e(\mathbb{R}^n_+)} = \|f_e\|_{H^1(\mathbb{R}^n)} \text{ and} \|f\|_{H^1_o(\mathbb{R}^n_+)} = \|f_o\|_{H^1(\mathbb{R}^n)} \tag{74}$$

and by$H^1_e(\mathbb{R}^n_+)$and$H^1_o(\mathbb{R}^n_+)$we denote the corresponding Banach spaces. It follows from Corollaries1.6, 1.8 of[61] and Proposition32 of[93] that the dual space of $H^1_e(\mathbb{R}^n_+)$is the space$BMO_r(\mathbb{R}^n_+)$and the dual space of $H^1_o(\mathbb{R}^n_+)$is the space $BMO_{\mathcal{Z}}(\mathbb{R}^n_+)$. See also[55].
The inclusion $BMO_e(\mathbb{R}^n_+) \subseteq BMO_r(\mathbb{R}^n_+)$is obvious. Hence to prove (i) it is enough to show that$BMO_r(\mathbb{R}^n_+) \subseteq BMO_e(\mathbb{R}^n_+)$. Let$f \in BMO_r(\mathbb{R}^n_+)$. To see that$f \in BMO_e(\mathbb{R}^n_+)$, by the definition it reduces to proving $f_e \in BMO(\mathbb{R}^n)$ where f_e is the even extension of f. For any $g(x) \in H^1(\mathbb{R}^n)$n), we denote by $\tilde{g}(x) = g(\tilde{x})$ where$\tilde{x} = (x', -x_n)$. Since$\left(H^1_e(\mathbb{R}^n_+)\right)' = BMO_r(\mathbb{R}^n_+)$, we have

$$\begin{aligned}\left|\int_{\mathbb{R}^n} f_e(x)g(x)dx\right| &= \left|\int_{\mathbb{R}^n_-} f_e(x)g(x)dx + \int_{\mathbb{R}^n_+} f_e(x)g(x)dx\right| \\ &= \left|\int_{\mathbb{R}^n_+} f(x)\big(g(\tilde{x}) + g(x)\big)dx\right| \le c\|f\|_{BMO_r(\mathbb{R}^n_+)}\|\tilde{g} + g\|_{H^1_e(\mathbb{R}^n_+)} \\ &\le c\|f\|_{BMO_r(\mathbb{R}^n_+)}\|g\|_{H^1(\mathbb{R}^n)}\end{aligned}$$

This shows that$BMO_r(\mathbb{R}^n_+) \subset BMO_e(\mathbb{R}^n_+)$, and proves(i).

We now prove(ii). The inclusion $BMO_{\mathcal{Z}}(\mathbb{R}^n_+) \subseteq BMO_o(\mathbb{R}^n_+)$is obvious. Let $f \in BMO_o(\mathbb{R}^n_+)$and thus $f_o \in BMO(\mathbb{R}^n)$. To see that $f \in BMO_{\mathcal{Z}}(\mathbb{R}^n_+)$, it reduces to proving $f \in \left(H^1_o(\mathbb{R}^n_+)\right)'$since $BMO_{\mathcal{Z}}(\mathbb{R}^n_+) = \left(H^1_o(\mathbb{R}^n_+)\right)'$. If$g \in H^1_o(\mathbb{R}^n_+)$, then$g_o \in H^1(\mathbb{R}^n)$. Hence

$$\left|\int_{\mathbb{R}^n_+} f(x)g(x)dx\right| = \frac{1}{2}\left|\int_{\mathbb{R}^n_-} f_o(x)g_o(x)dx\right|$$
$$\leq c\|f_o\|_{BMO(\mathbb{R}^n)}\|g_o\|_{H^1(\mathbb{R}^n)} \leq c\|f\|_{BMO(\mathbb{R}^n_+)}\|g\|_{H^1(\mathbb{R}^n_+)}.$$

This shows that $BMO_o(\mathbb{R}^n_+) \subset BMO_{\mathcal{Z}}(\mathbb{R}^n_+)$, and proves(ii).

We use Proposition(3.2.8) to obtain the following result.

Proposition (3.2.9)[65]: (i) The spaces $BMO_{\Delta_{D_+}}(\mathbb{R}^n_+)$, $BMO_{\mathcal{Z}}(\mathbb{R}^n_+)$ and $BMO_o(\mathbb{R}^n_+)$ coincide, and their norms are equivalent.

(ii) The spaces$BMO_{\Delta_{N_+}}(\mathbb{R}^n_+)$, $BMO_r(\mathbb{R}^n_+)$and$BMO_e(\mathbb{R}^n_+)$coincide, and their norms are equivalent.

Proof: We first prove(i). Let$f \in BMO_{\mathcal{Z}}(\mathbb{R}^n_+)$. By Proposition (3.2.8) we have that $f \in BMO_o(\mathbb{R}^n_+)$and then $f_o \in BMO(\mathbb{R}^n)$. To prove$f \in BMO_{\Delta_{D_+}}(\mathbb{R}^n_+)$, it suffices to show that for any cube $Q \subseteq \mathbb{R}^n_+$,

$$\int_Q \left|f(x) - e^{-l_Q^2\Delta_{D+}}f(x)\right| dx \leq c|Q|\|f\|_{BMO_{\mathcal{Z}}(\mathbb{R}^n_+)}. \quad (75)$$

By(73) and Propositions(3.2.8)

$$\frac{1}{|Q|}\int_Q \left|f(x) - e^{-l_Q^2\Delta_{D+}}f(x)\right| dx = \frac{1}{|Q|}\int_Q \left|f(x) - e^{-l_Q^2\Delta}f_o(x)\right| dx \leq c\|f_o\|_{BMO(\mathbb{R}^n)}$$
$$\leq c\|f\|_{BMO_o(\mathbb{R}^n_+)} \leq c\|f\|_{BMO_{\mathcal{Z}}(\mathbb{R}^n_+)}.$$

This proves(75).

Next assume that$f \in BMO_{\Delta_{D_+}}(\mathbb{R}^n_+)$. By Proposition(3.2.8), $f \in BMO_{\mathcal{Z}}(\mathbb{R}^n_+)$or equivalently $f_o \in BMO(\mathbb{R}^n)$. Note that by(58) it is enough to prove that for any cube$Q \subseteq \mathbb{R}^n$,

$$\int_Q \left|f_o(x) - e^{-l_Q^2\Delta}f_o(x)\right| dy \leq c|Q|\|f\|_{BMO_{\Delta_{D_+}}(\mathbb{R}^n_+)}. \quad (76)$$

We now verify(76). Let us examine the cubes Q.

Case1: If$Q \subseteq \mathbb{R}^n_-$, then for any $x \in Q$,

$$-\exp\left(-l_Q^2\Delta_{D_+}\right)f(\tilde{x}) = \exp\left(-l_Q^2\Delta\right)f_o(x)$$

and$\tilde{x} \in \hat{Q} \subseteq \mathbb{R}^n_+$(here $\tilde{Q}$ is a cube defined in(54)). Note also that$\left|\tilde{Q}\right| = |Q|$.Hence

$$\int_Q \left|f_o(x) - e^{-l_Q^2\Delta}f_o(x)\right| dy = \int_{\tilde{Q}} \left|f_o(\tilde{x}) - e^{-l_Q^2\Delta_{D+}}f(\tilde{x})\right| dx$$
$$\leq c|Q|\|f\|_{BMO_{\Delta_{D_+}}(\mathbb{R}^n_+)}.$$

Case2: If $Q \cap \mathbb{R}^n_- \neq \emptyset$ and $Q \cap \mathbb{R}^n_+ \neq \emptyset$, then let $\hat{Q}_-$ and $\hat{Q}_+$ be the two cubes as in (55). By(58) and Proposition(3.2.8),

$$\int_Q \left|f_o(x) - e^{-l_Q^2\Delta}f_o(x)\right| dy = \int_{Q_-\cup Q_+} \left|f_o(x) - e^{-l_Q^2\Delta}f_o(x)\right| dx$$
$$\leq 2\int_{\hat{Q}_+} \left|f(x) - e^{-l_Q^2\Delta_{D+}}f(x)\right| dx \leq 2|Q|\|f\|_{BMO_{\Delta_{D_+}}(\mathbb{R}^n_+)}.$$

Case3: If $Q \subseteq \mathbb{R}^n_+$, then$e^{-l_Q^2\Delta}f_o(x) = e^{-l_Q^2\Delta_{D+}}f(x)$for any$x \in Q$. Hence

$$\int_Q \left|f_o(x) - e^{-l_Q^2\Delta} f_o(x)\right| dx \le |Q| \|f\|_{BMO_{\Delta_{D_+}}(\mathbb{R}^n_+)}.$$

The estimate(76) follows readily. This shows that $f_o \in BMO(\mathbb{R}^n)$ so $f \in BMO_{\mathcal{Z}}(\mathbb{R}^n_+)$.
The proof of (ii) is similar to the proof of (i) so we skip it.
In a similar way as for the upper-half space, we can define the space $BMO_{\Delta_{D_-}}(\mathbb{R}^n_-)$ and $BMO_{\Delta_{N_-}}(\mathbb{R}^n_-)$ associated with the Dirichlet and Neumann Laplacian Δ_{D_-}, Δ_{N_-} on the lower-half space $\mathbb{R}^n_-$.
The same argument as in Proposition (3.2.9) gives the following proposition.

Proposition (3.2.10)[65]: (i) The spaces $BMO_{\Delta_{D_-}}(\mathbb{R}^n_-)$, $BMO_{\mathcal{Z}}(\mathbb{R}^n_-)$ and $BMO_o(\mathbb{R}^n_-)$ coincide, and their norms are equivalent.
(ii) The spaces $BMO_{\Delta_{D_-}}(\mathbb{R}^n_-)$, $BMO_r(\mathbb{R}^n_-)$ and $BMO_e(\mathbb{R}^n_-)$ coincide, and their norms are equivalent.

We mention that all cases of relation between the classical BMO and the new BMO spaces are possible. The following theorem provides simple example to prove this statement.

Proposition (3.2.11)[65]: The BMO spaces corresponding to the operators Δ_N, Δ_D and Δ_{ND} can be described in the following way:

$$BMO_{\Delta_N}(\mathbb{R}^n) = \{f \in \mathcal{M}(\mathbb{R}^n): f_+ \in BMO_r(\mathbb{R}^n_+) \text{and } f_- \in BMO_r(\mathbb{R}^n_-)\};$$
$$BMO_{\Delta_D}(\mathbb{R}^n) = \{f \in \mathcal{M}(\mathbb{R}^n): f_+ \in BMO_{\mathcal{Z}}(\mathbb{R}^n_+) \text{and } f_- \in BMO_{\mathcal{Z}}(\mathbb{R}^n_-)\};$$
$$BMO_{\Delta_{DN}}(\mathbb{R}^n) = \{f \in \mathcal{M}(\mathbb{R}^n): f_+ \in BMO_r(\mathbb{R}^n_+) \text{and } f_- \in BMO_{\mathcal{Z}}(\mathbb{R}^n_-)\}.$$

Proof: In the following proof L is one of the operators Δ_N, Δ_D or Δ_{DN}. If $L = \Delta_N$, then we denote by $L_+ = \Delta_{D_+}$ and $L_- = \Delta_{N_-}$. Similarly if $L = \Delta_D$ then $L_+ = \Delta_{D_+}$ and $L_- = \Delta_{D_-}$. Finally for $L = \Delta_{DN}$ we let $L_+ = \Delta_{N_+}$ and $L_- = \Delta_{D_-}$.
By (61), (65) and(68)

$$(\exp(-tL)f)_+ = \exp(tL_+) f_+ \text{and} (\exp(-tL)f)_- = \exp(tL_-) f_- \tag{77}$$

for any of the three considered operators. Hence for any cube $Q \subset \mathbb{R}^n$ we have

$$\int_Q \left|f - e^{-l_Q^2 L} f(x)\right| dx = \int_{Q\cap\mathbb{R}^n_-} \left|f_- - e^{-l_Q^2 L_-} f_-(x)\right| dx$$
$$+ \int_{Q\cap\mathbb{R}^n_+} \left|f_+ - e^{-l_Q^2 L_+} f_+(x)\right| dx. \tag{78}$$

In virtue of Propositions |(3.2.9) and (3.2.10) it is enough to show that

$$BMO_L(\mathbb{R}^n) = \left\{f \in \mathcal{M}(\mathbb{R}^n): f_+ \in BMO_{L_+}(\mathbb{R}^n_+) \text{and} f_- \in BMO_{L_-}(\mathbb{R}^n_-)\right\}$$

Assume now that $f \in \mathcal{M}(\mathbb{R}^n)$ such that $f_- \in BMO_{L_-}(\mathbb{R}^n_-)$ and $f_+ \in BMO_{L_+}(\mathbb{R}^n_+)$. In order to prove $f \in BMO_L(\mathbb{R}^n)$, it suffices to prove that for any cube $Q \subseteq \mathbb{R}^n$,

$$\int_Q \left|f(x) - e^{-l_Q^2 L} f(x)\right| dx \le c\left(\|f_-\|_{BMO_{L_-}(\mathbb{R}^n_-)} + \|f_+\|_{BMO_{L_+}(\mathbb{R}^n_+)}\right).$$

As in the proof of Proposition(3.2.9), we consider the following three cases of Q.
Case1: If $Q \subseteq \mathbb{R}^n_-$, then by(78)

$$\int_Q \left|f(x) - e^{-l_Q^2 L} f(x)\right| dx = \int_Q \left|f_-(x) - e^{-l_Q^2 L_-} f_-(x)\right| dx \le c|Q| \|f_-\|_{BMO_{L_-}(\mathbb{R}^n_-)}.$$

Case2: If $Q \cap \mathbb{R}^n_- = \emptyset$ and $Q \cap \mathbb{R}^n_+ = \emptyset$, then let $\tilde{Q}_-$ and $\tilde{Q}_+$ be the cubes as in(55). By(78)

$$\int_Q \left|f(x) - e^{-l_Q^2 L} f(x)\right| dx = \int_{Q_-\cup Q_+} \left|f(x) - e^{-l_Q^2 L} f(x)\right| dx$$

$$\leq \int_{\tilde{Q}_-} \left| f_- - e^{-l_Q^2 L_-} f_-(x) \right| dx + \int_{\tilde{Q}_+} \left| f_+ - e^{-l_Q^2 L_+} f_+(x) \right| dx$$
$$\leq c|Q| \left(\|f_-\|_{BMO_{L_-}(\mathbb{R}^n_-)} + \|f_+\|_{BMO_{L_+}(\mathbb{R}^n_+)} \right).$$
Case3: If $Q \subseteq \mathbb{R}^n_-$, then by(80)
$$\int_Q \left| f(x) - e^{-l_Q^2 L} f(x) \right| dx = \int_Q \left| f_+ - e^{-l_Q^2 L_+} f_+(x) \right| dx \leq c|Q| \|f_+\|_{BMO_{\Delta_{N_+}}(\mathbb{R}^n_+)}.$$
Hence $f \in BMO_L(\mathbb{R}^n)$.

We now assume that $f \in BMO_L(\mathbb{R}^n)$. By (78), we have that
$$f_- \in BMO_{L_-}(\mathbb{R}^n_-) \text{and } f_+ \in BMO_{L_+}(\mathbb{R}^n_+)$$
Now Proposition(3.2.11) is a straightforward consequence of Propositions(3.2.9) and(3.2.10).

The logarithmic function is a simple example that typifies some of the essential properties of the classical space BMO. For example if we define function $\log: \mathbb{R}^n \to \mathbb{R}$ by the formula $\log^e(x) = \log|x_n|$for all $x \in \mathbb{R}^n$ and $\log(x) = H(x_n) \log|x_n|$, where H is the Heaviside function then
$$\log^e \in BMO(\mathbb{R}^n)$$
$$\log \notin BMO(\mathbb{R}^n). \tag{79}$$
See, for examples, Chapter IV of [19] and page217 of[85] .

Theorem (3.2.12)[65]: In the notation described above the following inclusions hold
$$BMO_{\Delta_D}(\mathbb{R}^n) \subsetneq BMO(\mathbb{R}^n) \subsetneq BMO_{\Delta_N}(\mathbb{R}^n). \tag{80}$$
That is, the classical BMO space is a proper subspace of $BMO_{\Delta_N}(\mathbb{R}^n)$, and $BMO_{\Delta_D}(\mathbb{R}^n)$is a proper subspace of BMO.

Moreover, we have
$$BMO(\mathbb{R}^n) \nsubseteq BMO_{\Delta_D}(\mathbb{R}^n) \; and \; BMO_{\Delta_D}(\mathbb{R}^n) \nsubseteq BMO(\mathbb{R}^n). \tag{81}$$
Proof: It is a straight for ward consequence of Definition(3.2.4) that if $f_+ \in BMO_{\mathcal{Z}}(\mathbb{R}^n_+)$ and $f_- \in BMO_{\mathcal{Z}}(\mathbb{R}^n_-)$then $f \in BMO$. It also follows from Definition(3.2.4)that if $f \in BMO$ then $f_+ \in BMO_r(\mathbb{R}^n_+)$ and $f_- \in BMO_r(\mathbb{R}^n_-)$. Hence it follows from Theorem(3.2.12) and Propositions (3.2.9) and (3.2.10) that
$$BMO_{\Delta_D}(\mathbb{R}^n) \subset BMO(\mathbb{R}^n) \subset BMO_{\Delta_N}(\mathbb{R}^n).$$
To prove that the above inclusions are proper we note that by(79) and Definition(3.2.4)
$$\log_+ \notin BMO_{\mathcal{Z}}(\mathbb{R}^n_+) \text{and } \log_+ \in BMO_r(\mathbb{R}^n_+),$$
where $\log_+$ is the restriction of$\log^e$to$\mathbb{R}^n_+$. Next if $\log_-$is the restriction of$\log^e$to$\mathbb{R}^n_-$ then
$$\log_- \notin BMO_{\mathcal{Z}}(\mathbb{R}^n_-) \text{and } \log_- \in BMO_r(\mathbb{R}^n_-).$$
Hence
$$\log^e \in BMO \text{ and } \log^e \notin BMO_{\Delta_N}(\mathbb{R}^n).$$
Similarly
$$\log \notin BMO \text{and } \log \in BMO_{\Delta_N}(\mathbb{R}^n)$$
This ends the proof of(77). Finally to prove (78) we note that $BMO_{\Delta_{DN}}(\mathbb{R}^n)$and $\log \notin BMO_{\Delta_{DN}}(\mathbb{R}^n)$.

Corollary (3.2.13)[65]: (i) The dual space of $H^1_\Delta(\mathbb{R}^n)$is the space$BMO_\Delta(\mathbb{R}^n)$.

(ii) The dual spaces of $H^1_{\Delta_D}(\mathbb{R}^n)$, $H^1_{\Delta_N}(\mathbb{R}^n)orH^1_{\Delta_{DN}}(\mathbb{R}^n)$are the spaces $BMO_{\Delta_D}(\mathbb{R}^n)$, $BMO_{\Delta_N}(\mathbb{R}^n)$ or $BMO_{\Delta_{DN}}(\mathbb{R}^n)$, respectively.

(iii) For the Neumann Laplacian Δ_N on $\mathbb{R}^n$, we have that $H^1_{\Delta_N}(\mathbb{R}^n) \subsetneq H^1(\mathbb{R}^n)$ and $H^1_{\Delta_N}(\mathbb{R}^n) \neq \emptyset$. That is, $H^1_{\Delta_N}(\mathbb{R}^n)$ is a proper subspace of the classical Hardy space $H^1(\mathbb{R}^n)$.

For any $0 < \alpha < n$, the fractional powers $L^{-\alpha/2}$ of L is defined by

$$L^{-\alpha/2} f(x) = \frac{1}{\Gamma(\alpha/2)} \int_0^\infty L^{-\alpha/2-1} e^{-tL} f(x)\, dt. \tag{82}$$

We assume that the semigroup e^{-tL} has a kernel $p_t(x,y)$ which satisfies the upper bound(70) so $\left|L^{-\alpha/2} f(x)\right| \leq c\mathcal{I}_\alpha(|f|)(x)$ for all $x \in \mathbb{R}^n$, where

$$\mathcal{I}_\alpha f(x) = \int_{\mathbb{R}^n} \frac{f(y)}{|x-y|^{n-\alpha}} dy, 0 < \alpha < n.$$

is the classical fractional powers of the Laplacian Δ on $\mathbb{R}^n$. Let us recall that the semigroup $\{\exp(-tL): t > 0\}$ acting on $L^p(\mathbb{R}^n)$ is equicontinuous on $L^p(\mathbb{R}^n)$ if $\sup_{t>0} \|e^{-tL}\|_{L^p \to L^p} < \infty$. Note that all the semigroups which we consider here are equicontinuous on all $L^p(\mathbb{R}^n)$ for $1 \leq p \leq \infty$. We need the following Hardy-Littlewood-Sobolev theorem. See [88].

Proposition (3.2.14)[65]: Suppose that e^{-tL} is a semigroup which is equicontinuous on $L^1(\mathbb{R}^n)$ and $L^\infty(\mathbb{R}^n)$. Also suppose that

$$p_t(x,x) \leq t^{-n/2}.$$

Then for $0 < \alpha < n$,

(i) for $1 < p < \frac{n}{\alpha}$ and $\frac{1}{q} = \frac{1}{p} - \frac{n}{\alpha}$, we have

$$\left\|L^{-\alpha/2} f\right\|_{L^q} \leq c_{p,q} \|f\|_{L^p};$$

(ii) $L^{-\alpha/2}$ is of weak-type $(1,q)$, that is, for any $\lambda > 0$, we have

$$\left|\left\{x: \left|L^{-\alpha/2} f(x)\right| > \lambda\right\}\right| \leq c\left(\frac{\|f\|_{L^1}}{\lambda}\right)^q,$$

where $q = \left(1 - \frac{\alpha}{n}\right)^{-1}$.

Let us consider the limiting case $q = \infty$ in Proposition(3.2.14). It is wellknown that for every $f \in L^{-\alpha/2}$ $(\mathbb{R}^n)$, either $\mathcal{I}_\alpha f \equiv \infty$ or $\mathcal{I}_\alpha f \in BMO(\mathbb{R}^n)$ with

$$\|\mathcal{I}_\alpha f\|_{BMO(\mathbb{R}^n)} \leq c\|f\|_{L^{-\frac{\alpha}{2}}} \tag{83}$$

see [85].

An example of $\mathcal{I}_\alpha f \equiv \infty$ is given by $f(x) = |x|^{-\alpha} \log^{-1}|x| \chi_{\{x:|x|\geq 2\}}$. The following result generalizes estimates (83).

Lemma (3.2.15)[65]: Assume that the semigroup e^{-tL} has a kernel $p_t(x,y)$ which satisfies upper bound(70). Then for $0 < \alpha < n$, the difference operator $(I - e^{-tL})L^{-\alpha/2}$ has an associated kernel $K_{\alpha,t}(x,y)$ which satisfies

$$\left|K_{\alpha,t}(x,y)\right| \frac{c}{|x-y|^{n-\alpha}} \frac{t}{|x-y|^2} \text{ for some constant } c > 0. \tag{84}$$

Proof: Note that

$$I - e^{-tL} = \int_0^t \frac{d}{dr} e^{-rL} = -\int_0^t L e^{-rL}\, dr.$$

Hence by (82)

$$(I - e^{-tL})L^{-\alpha/2} = \frac{1}{\Gamma(\alpha/2)} \int_0^t \int_0^\infty \left(v \frac{d}{du} e^{-vL} \right) \Big|_{v=r+s} \frac{1}{r+s} \frac{dsdr}{s^{-\alpha/2+1}}.$$

By Lemma2.5 of [60], the kernel of the operator $v \frac{d}{dv} e^{-vL}$has Gaussian upper bound(70). Hence, the operator$(I - e^{-tL})L^{-\alpha/2}$ has an associated kernel $K_{\alpha,t}(x,y)$which satisfies

$$\begin{aligned}
|K_{\alpha,t}(x,y)| &\le c \int_0^t \int_0^\infty \frac{1}{(r+s)^{n/2}} e^{-c_1 \frac{|x-y|^2}{r+s}} \frac{1}{r+s} \frac{dsdr}{s^{-\alpha/2+1}} \\
&\le c \int_0^t \int_0^r \frac{1}{(r+s)^{n/2}} e^{-c_1 \frac{|x-y|^2}{r+s}} \frac{1}{r+s} \frac{dsdr}{s^{-\alpha/2+1}} \\
&+ c \int_0^t \int_r^\infty \frac{1}{(r+s)^{n/2}} e^{-c_1 \frac{|x-y|^2}{r+s}} \frac{1}{r+s} \frac{dsdr}{s^{-\alpha/2+1}} \\
&= \mathrm{I} + \mathrm{II}.
\end{aligned}$$

Let us estimate term I. Note that$0 < s < r$. We have

$$\mathrm{I} \le c \int_0^t \int_0^r r^{-n/2} e^{-c_2 \frac{|x-y|^2}{r}} \frac{dsdr}{rs^{-\alpha/2+1}}$$

$$= \frac{c}{|x-y|^{n-\alpha}} \int_0^{t|x-y|^2} r^{(\alpha-n-2)} e^{-c_2 r^{-1}}\, dr \le \frac{c}{|x-y|^{n-\alpha}} \frac{t}{|x-y|^2},$$

where the last inequality follows from $r^{(\alpha-n-2)/2} e^{-cr^{-1}} \le c$for some positive constant c. On the other hand, using the condition $0 < \alpha < n$ we obtain

$$\mathrm{II} \le c \int_0^t \int_r^\infty s^{-\frac{n}{2}} e^{-c_2 \frac{|x-y|^2}{s}} \frac{dsdr}{s^{-\alpha/2+2}} \le \frac{ct}{|x-y|^{n+2-\alpha}} \int_0^\infty s^{(\alpha-n-4)/2} e^{-c_2 s^{-1}} ds$$

$$\le \frac{c}{|x-y|^{n-\alpha}} \frac{t}{|x-y|^2}.$$

Therefore, condition(84) is satisfied and the proof of Lemma(3.2.15) is complete.

Theorem (3.2.16)[65]: Assume that the semigroupe^{-tL} has a kernel $p_t(x,y)$ which satisfies the upper bound(70). If $f \in L^{-\alpha/2}\,(\mathbb{R}^n)$ and $L^{-\alpha/2} f < \infty$ almost everywhere, then $L^{-\alpha/2}\, f \in BMO_L(\mathbb{R}^n)$ with

$$\|L^{-\alpha/2}\, f\|_{BMO_L(\mathbb{R}^n)} \le c\|f\|_{n/\alpha}$$

for $0 < \alpha < n$, where the positive constant c depends only on α and n.

Suppose that T is a bounded operator on$L^2(\Omega)$. We say that a measurable function$K_T\colon \Omega^2 \to \mathbb{C}$is the (singular) kernel of T if

$$\langle Tf_1, f_2 \rangle = \int_\Omega Tf_1(x)\overline{f_2}(x)dx = \int_\Omega \int_\Omega K_T(x,y) f_1(y)\overline{f_2(x)}dxdy. \tag{85}$$

for all $f_1, f_2 \in C_c(\Omega)$(for all$f_1, f_2 \in C_c(\Omega)$such thatsupp $f_1 \cap$ supp $f_2 = \emptyset$ respectively).

In order to prove Theorem (3.2.16), we need the following estimate on the kernel $K_{\alpha,t}(x,y)$of the operator$(I - e^{-tL})L^{-\alpha/2}$(see also [98]).

Proof: In virtue of the definition of $BMO_L(\mathbb{R}^n)$, it suffices to prove there exists a constant $C > 0$ such that that for any ball $B(x,r)$ with radius r centered at x

$$\frac{1}{|B(x,r)|} \int_{B(x,r)} \left| \left(I - e^{-r^2 L}\right) L^{-\alpha/2} f(y) \right| dy \le C\|f\|_{L^{-n/\alpha}} \tag{86}$$

for all $f \in L^{n/\alpha}\,(\mathbb{R}^n)$. Set$f_1(y) = f(y)$if $|x - y| \le 2r$ and$f_1(y) = 0$ otherwise. Next, put$f_2 = f - f_1$. Note that

$$\frac{1}{|B|}\int_{B(x,r)}\left|\left(I-e^{-r^2L}\right)L^{-\alpha/2}f(y)\right|dy$$

$$\leq \frac{1}{|B|}\int_{B(x,r)}\left|\left(I-e^{-r^2L}\right)L^{-\alpha/2}f_1(y)\right|dy+\frac{1}{|B|}\int_{B(x,r)}\left|\left(I-e^{-r^2L}\right)L^{-\alpha/2}f_2(y)\right|dy$$

$$= \mathrm{I}+\mathrm{II},$$

Where$|B| = |B(x,r)|$. To estimate the first term note that, by Hölder's inequality$\|f_1\|_{L^p} \leq c|B(x,r)|^{1/p-\alpha/n}\|f\|_{L^{-n/\alpha}}$. for all $1 < p < n/\alpha$. Next, set$1/q = 1/p - \alpha/n$. By Proposition(3.2.14)

$$\mathrm{I} \leq \frac{1}{|B|^{1/q}}\left\|\left(I-e^{-r^2L}\right)L^{-\alpha/2}f_1\right\|_{L^q} \leq c\frac{1}{|B|^{1/q}}\left\|L^{-\alpha/2}f_1\right\|_{L^q}$$

$$\leq c\frac{1}{|B|^{1/q}}\|f_1\|_{L^p} \leq c\frac{1}{|B|^{1/q}}\|f\|_{L^{n/\alpha}}.$$

To estimate the second term note that if$y \in B(x,r)$, then by Lemma(3.2.15)

$$\left|\left(I-e^{-r^2L}\right)L^{-\alpha/2}f_2(y)\right| \leq \int_{B(x,2r)^c}\left|K_{\alpha,r^2}(x,\mathcal{Z})\right||f(\mathcal{Z})|d\mathcal{Z}$$

$$\leq c\sum_{k=1}^{\infty}\int_{2^k r\leq|x-\mathcal{Z}|<2^{k+1}r}\frac{1}{|x-\mathcal{Z}|^{n-\alpha}}\frac{r^2}{|x-\mathcal{Z}|^2}|f(\mathcal{Z})|d\mathcal{Z}$$

$$\leq c\sum_{k=1}^{\infty}2^{2k}\frac{1}{|B(x,r2^{k+1})|^{1-\alpha/n}}\int_{B(x,r2^{k+1})}|f(\mathcal{Z})|d\mathcal{Z}$$

$$\leq c\sum_{k=1}^{\infty}2^{2k}\|f\|_{L^{n/\alpha}} \leq c\|f\|_{L^{n/\alpha}}.$$

Combining the above estimates, we obtain(86).

Remark(3.2.17)[65]: (i) Under the extra assumption that for each$t > 0$, the kernel $p_t(x,y)$ of $e^{-tL}$is a Hölder continuous function in$x$, it can be proved that for $f \in L^{n/\alpha}$ $(\mathbb{R}^n)$, either $L^{\alpha/2}f \equiv \infty$or $L^{n/\alpha}f \in BMO_L$ $(\mathbb{R}^n)$with

$$\left\|L^{n/\alpha}f\right\|_{BMO_L(\mathbb{R}^n)}\|f\|_{L^{n/\alpha}}.$$

(ii) We now give a list of examples of operators L satisfying the assumptions in Proposition (3.2.14) and Theorem(3.2.16).

(α) The operator Δ_N,Δ_D or Δ_{DN};

(β) Let $V \in L^1_{\text{loc}}$ $(\mathbb{R}^n)$be a nonnegative function on $\mathbb{R}^n (n \geq 3)$. The Schrödinger operator with potential V is defined by

$$L = -\Delta + V(x) \text{ on } \mathbb{R}^n. \quad (87)$$

From the Feynman-Kac formula, it is well-known that the kernels $p_t(x,y)$ of the semigroup e^{-tL}satisfy the estimate

$$0 \leq p_t(x,y) \leq \frac{1}{(4\pi t)^{n2}}e^{-\frac{|x-y|^2}{4t}}. \quad (88)$$

However, unless V satisfies additional conditions, the heat kernel can be a discontinuous function of the space variables and the Hölder continuous estimates may fail to hold. See, for example,[11].

We note that the corresponding result in Theorem 1 of [99] is a special case of Theorem(3.2.16).

(γ) Let $A = \left(a_{ij}(x)\right)_{1\le i,j\le n}$ be an $n\times n$ matrix with complex entries $a_{ij} \in L^\infty(\mathbb{R}^n)$ satisfying $\lambda|\xi|^2 \le Re\sum a_{ij}(x)\,\xi_i\xi_j$ for all $x \in \mathbb{R}^n$, $\xi = (\xi_1\xi_2, \dots, \xi_n) \in \mathbb{C}^n$ and some $\lambda > 0$. Let T be the divergence form operator

$$Lf \equiv -\operatorname{div}(A\nabla f),$$

which we interpret in the usual weak sense via a sesquilinear form. It is known that Gaussian bound(70) on the heat kernel e^{-tL} is true when A has real entries, or when $n = 1, 2$ in the case of complex entries. See, for example,[57].

The following example complements Theorems(3.2.12) and(3.2.16). It also provides a convincing justification of introduction of the BMO_L spaces.

Example (3.2.18)[65]: Let Δ_N be the Neumann Laplacian on $\mathbb{R}$. Then, there exists a function $f \in L^{1/\alpha}(\mathbb{R})$ such that $\Delta_N^{-\alpha/2}f(x) < \infty$ for almost every $x \in \mathbb{R}$, $\Delta_N^{-\alpha/2}f \in BMO_{\Delta_N}(\mathbb{R})$ and

$$\left\|\Delta_N^{-\alpha/2}f\right\|_{BMO_{\Delta_N}(\mathbb{R})} \le c\|f\|_{L^{n/\alpha}}. \tag{89}$$

However, $\Delta_N^{-\alpha/2}f \notin BMO(\mathbb{R})$.

Proof: For any $0 < \alpha < 1$, we let

$$f(x) = -\frac{1}{x^\alpha \log x}\chi_{\{0<x\le 1/2\}}(x). \tag{90}$$

Then

$$\int_{\mathbb{R}} |f(y)|^{1/\alpha}\,dy = \int_0^{1/2} \frac{1}{y(\log y^{-1})^{1/\alpha}}\,dy = (1-\alpha)\alpha^{-1}(\log 2)^{1/\alpha-1} < \infty.$$

This proves that $f \in L^{1/\alpha}(\mathbb{R})$. It can be verified that $\mathcal{I}_\alpha f(x) < \infty$ a.e.. Also, we have that $\Delta_N^{-\alpha/2}f(x) < \infty$ a.e.. Hence,

(a) $\mathcal{I}_\alpha f(x) \in BMO(\mathbb{R})$ with $\|\mathcal{I}_\alpha f\|_{BMO(\mathbb{R})} \le c\|f\|_{L^{n/\alpha}}$. See [31, page221].

(b) By Theorem(3.2.16), we have that $\Delta_N^{-\alpha/2}f \in BMO_{\Delta_N}(\mathbb{R})$ with estimate(89). We now prove $\Delta_N^{-\alpha/2}f \notin BMO(\mathbb{R})$. Denote by $k_\alpha^N(x,y)$ the kernel of the fractional power $\Delta_N^{-\alpha/2}$ of Δ_N. By(62) and(82)

$$k_\alpha^N(x,y) = \frac{1}{\gamma(\alpha)}\left(\frac{1}{|x-y|^{1-\alpha}} + \frac{1}{|x+y|^{1-\alpha}}\right)H(xy), \tag{91}$$

where H is the Heaviside function (63). By(91)

$$\Delta_N^{-\alpha/2}f(x) = \begin{cases} 0 & \text{if } x \le 0; \\ \mathcal{I}_\alpha(f_e)(x) & \text{if } x > 0, \end{cases} \tag{92}$$

where $f_e \in L^{1/\alpha}(\mathbb{R})$ is given by the formula $f_e(x) = -\frac{1}{|x|^\alpha \log|x|}\chi_{\{|x|\le 1/2\}}(x)$.

For any $k \ge 5$, we denote $Q_k = [-1/k, 1/k]$. Next if $0 < x < y < 1/2$, then $|x-y| < |y|$. Hence

$$\begin{aligned}\Delta_N^{-\alpha/2}f(x) &= \frac{1}{\gamma(\alpha)}\int_{-1/2}^{1/2} \frac{1}{|x-y|^{1-\alpha}}f_e(y)dy \\ &\ge -\frac{1}{\gamma(\alpha)}\int_x^{1/2} \frac{1}{|x-y|^{1-\alpha}}\frac{1}{y^\alpha \log y}\,dy\end{aligned}$$

$$\geq -\frac{1}{\gamma(\alpha)}\int_x^{1/2}\frac{1}{y\log y}dy$$
$$\geq \frac{1}{\gamma(\alpha)}\left(\log\left(\log\frac{1}{x}\right)-\log(\log 2)\right),$$
which yields
$$m_{Q_k}\left(\Delta_N^{-\alpha/2}f\right)=\frac{1}{|Q_k|}\int_{Q_k}\Delta_N^{-\alpha/2}f(y)dy$$
$$\geq \frac{1}{2\gamma(\alpha)}\int_0^{1/k}\left(\log\left(\log\frac{1}{y}\right)-\log(\log 2)\right)dy$$
$$\geq \frac{1}{2\gamma(\alpha)}(\log(\log k)-\log(\log 2)).$$
Therefore, from(92) we obtain
$$\frac{1}{|Q_k|}\int_{Q_k}\left|\Delta_N^{-\alpha/2}f(y)-m_{Q_k}\left(\Delta_N^{-\alpha/2}f\right)\right|dx$$
$$=\frac{k}{2}\int_0^{1/k}\left|\Delta_N^{-\alpha/2}f(x)-m_{Q_k}\left(\Delta_N^{-\alpha/2}f\right)\right|dx+\frac{k}{2}\int_{-1/k}^{0}\left|m_{Q_k}\left(\Delta_N^{-\alpha/2}f\right)\right|dx$$
$$\geq \frac{1}{2}\left|m_{Q_k}\left(\Delta_N^{-\alpha/2}f\right)\right|$$
$$\geq \frac{1}{4\gamma(\alpha)}(\log(\log k)-\log(\log 2)).$$
Note that the last term in the above inequality tends to ∞ as $k\to\infty$. Hence
$$\sup_Q\frac{1}{Q}\int_Q\left|\Delta_N^{-\alpha/2}f(x)-m_Q\left(\Delta_N^{-\alpha/2}f\right)\right|dx=\infty,$$
where the supremum is taken over all cubes Q of $\mathbb{R}$. Therefore $\Delta_N^{-\alpha/2}f\notin BMO(\mathbb{R})$.

We apply the technique of BMO_L spaces to discuss optimal L^p estimates for the imaginary powers of the operator L. We refer readers to [96], [100] for related results concerning imaginary powers of self-adjoint operators.

Let us recall that if L is a self-adjoint positive definite operator on$L^2(\mathbb{R}^n)$. Then L admits the spectral resolution:
$$L=\int_0^\infty \lambda dE_L(\lambda),$$
where the$E_L(\lambda)$are spectral projectors. For any bounded Borel function $F:[0,\infty)\to\mathbb{C}$, we define the operator $F(L)$ by the formula
$$F(L)=\int_0^\infty F(\lambda)dE_L(\lambda). \tag{93}$$
In particular
$$L^{is}=\int_0^\infty L^{is}dE(t).$$
By spectral theory$\left\|L^{is}\right\|_{L^2\to L^2}=1$ for all $s\in\mathbb{R}$. In the following theorem we obtain sharp estimates for the $L^\infty\to BMO_L$ norm of the operators L^{is}.

Theorem (3.2.19)[65]: Assume that the heat kernel $p_t(x,y)$ corresponding to the self-adjoint operator L satisfies upper bound(70). Then

$$\left\|L^{is}f\right\|_{BMO_L(\mathbb{R}^n)} \leq c(1+|s|)^{n/2}\|f\|_{L^\infty}$$

for all $s \in \mathbb{R}$.

Proof: It is enough to show that for any ball $B(x,r)$ with radius r centered at x, there exists a constant $C > 0$ such that

$$\frac{1}{|B(x,r)|}\int_{B(x,r)}\left|\left(I-e^{-r^2L}\right)L^{is}f(y)\right|dy \leq C(1+|s|)^{n/2}\|f\|_{L^\infty}. \tag{94}$$

To prove(94), for any $f \in L^\infty(\mathbb{R}^n)$, we set $\theta = (1+|s|)^{n/2}$, $f_1(y) = f(y)$ if $|x-y| \leq \theta^1 r$ and $f_1(y) = 0$ otherwise. Next, we put $f_2 = f - f_1$. Note that

$$\frac{1}{|B|}\int_{B(x,r)}\left|\left(I-e^{-r^2L}\right)L^{is}f(y)\right|dy \leq \frac{1}{|B|}\int_{B(x,r)}\left|\left(I-e^{-r^2L}\right)L^{is}f_1(y)\right|dy$$
$$+\frac{1}{|B|}\int_{B(x,r)}\left|\left(I-e^{-r^2L}\right)L^{is}f_2(y)\right|dy$$
$$\mathrm{I}+\mathrm{II},$$

where $|B| = |B(x,r)|$. To estimate the term I we note that, by Hölder's inequality

$$\|f_1\|_{L^2} \leq |B(x,\theta^1 r)|^{1/2}\|f\|_{L^\infty}$$
$$\leq \frac{|B(x,r)|^{\frac{1}{2}}}{\theta^{\frac{n}{2}}}\|f\|_{L^\infty} = |B|^{1/2}(1+|s|)^{n/2}\|f\|_{L^\infty}.$$

Then

$$\mathrm{I} \leq |B|^{1/2}\left\|\left(I-e^{-r^2L}\right)L^{is}f_1\right\|_{L^2} \leq \mathrm{c}|B|^{1/2}\left\|L^{is}f_1\right\|_{L^2}$$
$$\leq c|B|^{\frac{1}{2}}\|f_1\|_{L^2} \leq c(1+|s|)^{n/2}\|f\|_{L^\infty}.$$

To estimate the term II we note that if $y \in B(x,r)$, then

$$\left|\left(I-e^{-r^2L}\right)L^{is}f_2(y)\right| \leq \int_{B(x,\theta^{-1}r)^c}\left|K_{is,r^2}(y,\mathcal{Z})\right||f(\mathcal{Z})|d\mathcal{Z}$$
$$\leq \|f\|_{L^\infty}\sup_{x\in\Omega,r>0}\int_{B(x,\theta^{-1}r)^c}\left|K_{is,r^2}(y,\mathcal{Z})\right||f(\mathcal{Z})|d\mathcal{Z},$$

where $K_{is,r^2}(y,\mathcal{Z})$ is the kernel of the operator $\left(I-e^{-r^2L}\right)L^{is}$. Hence the proof of Theorem (3.2.19) reduces to the following Lemma.

Lemma (3.2.20)[65]: Assume that L is a self-adjoint operator and its heat kernel $p_t(x,y)$ satisfies the Gaussian bound(70). Then the associated kernel $K_{is,r^2}(y,\mathcal{Z})$ of the operator $\left(I-e^{-r^2L}\right)L^{is}$ satisfies

$$\int_{B(x,\theta^{-1}r)^c}\left|K_{is,r^2}(y,\mathcal{Z})\right|d\mathcal{Z} \leq c(1+|s|)^{n/2}$$

for all $s \in \mathbb{R}$ and $r > 0$.

The proof of Lemma(3.2.20) is a minor modification of the proof of estimates(68) of [104].

Theorem (3.2.19) applied to the standard Laplace operator gives the following estimates.

Corollary (3.2.21)[65]: If Δ is the standard Laplace operator acting on $\mathbb{R}^n$ then

$$\left\|\Delta^{is}f\right\|_{BMO(\mathbb{R}^n)} \leq c(1+|s|)^{n/2}\|f\|_{L^\infty} \tag{95}$$

for all $s \in \mathbb{R}$.

Proof: Corollary(3.2.21) is a straightforward consequence of Theorem(3.2.19) and the equivalence of the classical BMO space andBMO_{Δ}.

Remark(3.2.22)[65]: For the standard Laplace operator one can explicitly compute the kernel $|K_{is,r^2}(y,\mathcal{Z})|$and check that

$$\int_{B(x,\theta^{-1}r)^c} |K_{is,r^2}(y,\mathcal{Z})| d\mathcal{Z} \geq c(1+|s|)^{n/2}\log(1+|s|).$$

See[104]. Hence one has to replace $B(x,2r)$ by$B(x,\theta^{-1}r)^c$to obtain estimates without the additional logarithmic term. As in[104] (Theorem1) one can show that the norm of$\|\Delta^{is}f\|_{L^\infty\to BMO\,(\mathbb{R}^n)} \geq c(1+|s|)^{n/2}$. Hence the estimates in Theorem (3.2.19) and Corollary(3.2.21) are sharp. Even for the Laplace operator, our estimate(95) is stronger than any other known estimates of $L^\infty \to BMO$ norm of the imaginary powers of the Laplace operator.

Theorem 2 of[104] says that if L satisfies assumption of Theorem(3.2.19) then the following estimates of the weak type(1,1) norm of the imaginary powers of$L$ holds

$$\|L^{is}\|_{L^1\to L^{1,\infty}} \leq c(1+|s|)^{n/2} \tag{96}$$

Note, however, that the week type(1,1) norm is not subadditive so despite its name is not a norm. Whereas$\|\cdot\|_{L^\infty\to BMO_L}$, the norm of linear operators form$L^\infty$to BMO_L, is a proper norm. This difference is crucial for the results which we discuss next.

Suppose that$F:\mathbb{R}\to\mathbb{C}$. Let us recall that the Mellin transform of the function F is defined by

$$m(u) = \frac{1}{2\pi}\int_0^\infty F(\lambda)\lambda^{-1-iu}d\lambda, u\in\mathbb{R}.$$

Moreover the inverse transform is given by the following formula

$$F(\lambda) = \int_{\mathbb{R}} m(u)\lambda^{iu}du,\ \lambda\ \in [0,\infty).$$

Next we define the maximal operator $F^*(L)$ by the formula

$$F^*(L)f(L) = \sup_{t>0}|F(tL)f(x)|,$$

where $f\in L^p\ (\Omega)$ for some$1\leq\ p\leq\infty$.

Corollary (3.2.23)[65]: Assume that L is a self-adjoint operator acting on $L^2\ (\mathbb{R}^n)$ and that the heat kernel $p_t(x,y)$ of the operator L satisfies upper bound(70). Suppose also that$F:\mathbb{R}\to\mathbb{C}$is a bounded Borel function such that

$$\int_{\mathbb{R}} |m(u)|(1+|s|)^{n/2}du = C_{F,n} \leq \infty$$

where m is the Mellin transform of F. Then$F(L)$ and$F^*(L)$ are bounded operators from$L^\infty$to$BMO_L$and

$$\|\cdot\|_{L^\infty\to BMO_L}\|\cdot\|_{L^\infty\to BMO_L} \leq cC_{F,n}.$$

Proof: Note that

$$F(tL) = \int_0^\infty F(t\lambda)dE_L(\lambda)\int_0^\infty\int_{\mathbb{R}} m(u)(t\lambda)^{iu}\,dudE_L(\lambda)$$

$$\int_{\mathbb{R}}\int_0^\infty m(u)(t\lambda)^{iu}\,dE_L(\lambda)du = \int_{\mathbb{R}} m(u)t^{iu}L^{iu}\,du.$$

Hence

$$\sup_{t>0}|F(tL)f(x)| \le \int_{\mathbb{R}} |m(u)|\left|L^{iu}f(x)\right| du$$

And

$$\|F^*(L)f\|_{BMO_L} \le \int_{\mathbb{R}} |m(u)|\|f\|_{L^\infty}\left\|L^{iu}\right\|_{L^\infty\to BMO_L} du$$
$$\le c\|f\|_{L^\infty}\int_{\mathbb{R}} |m(u)|(1+|u|)^{n/2}\, du.$$

The inequality$\|F(L)\|_{L^\infty\to BMO_L} \le \|F^*(L)\|_{L^\infty\to BMO_L}$is an obvious conse-quence of the definition of$F^*(L)$.

We discuss an application of $BMO_L(\Omega)$ technique to the theory of Hörmander spectral multipliers. If $F(L)$ is the operator defined by (93) then by $K_F(L)$we denote the kernel associated with$F(L)$. See(84) of [70].

Theorem (3.2.24)[65]: Suppose that$\|F(L)\|_{L^\infty} \le C_1$, and that

$$\sup_{r>0}\sup_{y\in\Omega}\int_{B(x,r)^c}\left|K_{F(L)\left(I-e^{-r^2L}\right)}(x,y)\right| dx \le C_1. \tag{97}$$

Then

$$\|F(L)\|_{L^\infty\to BMO_L} \le cC_1.$$

Proof: We note again that it is enough to show that for any ball $B(x,r)$ with radius r centered at$x$, there exists a constant $C>0$ such that

$$\frac{1}{|B(x,r)|}\int_{B(x,r)}\left|\left(I-e^{-r^2L}\right)F(L)f(y)\right| dy \le cC_1\|f\|_{L^\infty}. \tag{98}$$

To prove(98) for any $f\in L^\infty(\mathbb{R}^n)$we set$f_1(y)=f(y)if|x-y|\le 2r$ and$f_1(y)=0$ otherwise. Next, we put$f_2=f-f_1$. Note that

$$\frac{1}{|B|}\int_{B(x,r)}\left|\left(I-e^{-r^2L}\right)F(L)(y)\right| dy \le$$
$$\le \frac{1}{|B|}\int_{B(x,r)}\left|\left(I-e^{-r^2L}\right)F(L)f_1(y)\right| dy$$
$$+\frac{1}{|B|}\int_{B(x,r)}\left|\left(I-e^{-r^2L}\right)F(L)f_2(y)\right| dy$$
$$= \mathrm{I}+\mathrm{II},$$

where$|B|=|B(x,r)|$. To estimate the term I we note that, by Hölder's inequality

$$\|f_1\|_{L^2} \le |B(x,2r)|^{1/2}\|f\|_{L^\infty} \le c|B(x,2r)|^{1/2}\|f\|_{L^\infty}.$$

Then

$$1 \le |B|^{-1/2}\left\|\left(I-e^{-r^2L}\right)F(L)f_1\right\|_{L^2}$$
$$\le c|B|^{-1/2}\|F(L)f_1\|_{L^2}$$
$$\le c|B|^{-1/2}C_1\|f_1\|_{L^2}$$
$$\le cC_1\|f_1\|_{L^\infty}.$$

To estimate the term II we note that if$y\in B(x,r)$, then

$$\left|\left(I-e^{-r^2L}\right)F(L)f_2(y)\right| \le \int_{B(y,r)^c}\left|K_{\left(I-e^{-r^2L}\right)F(L)}(y,\mathcal{Z})\right| |f(\mathcal{Z})|d\mathcal{Z}$$
$$\le \|f\|_{L^\infty}\sup_{x\in\Omega,\mathrm{r}>0}\int_{B(x,r)^c}\left|K_{\left(I-e^{-r^2L}\right)F(L)}(y,\mathcal{Z})\right| d\mathcal{Z} \le cC_1\|f_1\|_{L^\infty}$$

In the standard theory of Hörmander spectral multipliers one usually begins with proving weak type $(1,1)$ estimates for a spectral multiplier$F(L)$. Next $F(L)$ is bounded onL^2 by the spectral theorem so continuity of the operator $F(L)$ onL^p spaces for $1 < p < \infty$ follows from the Marcinkiewicz interpolation theorem. One can use Theorem$(3.2.24)$ and Proposition$(3.2.3)$ to obtain an alternative proof of boundedness of $F(L)$ on anL^p space for$1 < p < \infty$. Of course continuity of $F(L)$ as an operator fromL^∞toBMO_L is of independent interest even if we already know that $F(L)$ is of weak type$(1,1)$.

The Hörmander type spectral multipliers is a very broad subject. For example such multipliers were studied in [92], [94], [97], [101], [102], [103]. One can use Theorem$(3.2.24)$ to show that all spectral multipliers of weak type$(1,1)$ which are discussed in [92], [94], [97], [101], [102], [103] are also bounded fromL^∞toBMO_L. As an example we discuss the followingBMO_L versions of Theorem3.1 of[97].Let us recall that if $F: \mathbb{R} \to \mathbb{C}$ then

$$\|F\|_{W_s^p} = \left\|(I + \Delta)^{n/2} F\right\|_{L^p(\mathbb{R})}.$$

Theorem (3.2.25)[65]: Suppose that L is a self-adjoint operator acting on$L^2(\Omega)$, $\Omega \subset \mathbb{R}^n$ and that the heat kernel $p_t(x,y)$of L satisfies the Gaussian bound(19) and that $\eta \in C_c^\infty(\mathbb{R}_+)$. Then for every $s > n/2$ and for all Borel bounded function F such that$\sup_{t>0}\|\eta\delta_t F\|_{W_s^\infty} < \infty$the operator $F(L)$ is bounded on $L^p(\Omega)$ for all $1 < p < \infty$. Moreover

$$\|F(L)\|_{L^\infty \to BMO_L} \le C_s \left(\sup_{t>0}\|\eta\delta_t F\|_{W_s^\infty}\right) for\ all s > \frac{n}{2}. \quad (99)$$

Proof: Note that by[97], we have

$$\sup_{r>0}\sup_{y\in\Omega}\int_{B(x,r)^c}\left|K_{F(L)\left(I-e^{-r^2L}\right)}(x,y)\right| dx \le C_s\left(\sup_{t>0}\|\eta\delta_t F\|_{W_s^\infty}\right).$$

Hence Theorem (3.2.25) is a straightforward consequence of Theorem (3.2.24).

Chapter 4
Upper Bound and Higher Order Journé Commutators

The result follows from a new proof of boundedness of iterated commutators in terms of the BMO norm of their symbol functions, using Hytönen's representation theorem of Calder´on– Zygmund operators as averages of dyadic shifts. The tensor products of Riesz transforms are a representative testing class for Journé operators. Previous results in this direction do not apply to tensor products and only to Journé operators which can be reduced to Calderón–Zygmund operators. Upper norm estimates of Journé commutators are new even in the case of no iterations. Lower norm estimates for iterated commutators only existed when no tensor products were present. In the case of one dimension, lower estimates were known for products of two Hilbert transforms, and without iterations. New methods using Journé operators are developed to obtain these lower norm estimates in the multi-parameter real variable setting.

Section (4.1): Multi-Parameter Iterated Commutators

In [111] the product BMO space on $\mathbb{R}^{d_1} \otimes \cdots \otimes \mathbb{R}^{d_t}$ was characterized by the multi-parameter iterated commutators of Riesz transforms. This extended to the product setting the classical results of R. R. Coifman, R. Rochberg, and G. Weiss [3], a characterization of classical BMO in terms of boundedness on $L^2(\mathbb{R}^d)$of the commutator of a singular integral operator with a multiplication operator, which by duality also implies a weak factorization result of $H^1(\mathbb{R}^d)$.

In the multi-parameter setting, let M_b be the operator of pointwise multiplication by $b \in BMO_{\text{prod}}(\mathbb{R}^{\vec{d}})$. Let T_i be the Calder´on–Zygmund operators on $\mathbb{R}^{d_i}$. One seeks to characterize product BMO in terms of commutators in the sense that

$$\|b\|_{BMO_{\text{prod}}} \leq \left\|\left[\dots\left[[M_b, T_1]T_2\right]\dots, T_t\right]\right\|_{L^2 \to L^2} \lesssim \|b\|_{BMO_{\text{prod}}},$$

where the first and second inequality will be referred to as lower bound and upper bound, respectively.

In the case of Hilbert transform, the above result in bi-parameter setting was proved by S.H. Ferguson and M. T. Lacey in [1], where the upper bound was first shown by S. H. Ferguson and C. Sadosky [5].

M. Lacey and E. Terwilleger [113] then extended the result to the multi-parameter setting. The Riesz transform result was proved by M. T. Lacey, S. Petermichl,$J.C.$ Pipher, and B. D. Wick in [111], where they obtained a more general upper bound result for any Calder´on– Zygmund operators of convolution type with high degree of smoothness. Later on in[112]they simplified the proof of the upper bound for Riesz transforms by means of dyadic shifts. S. Petermichl [108] proved the lower bound for a larger class of Calder´on– Zygmund operators satisfying certain criteria.

We show the upper bound for any given collection of Calder´on–Zygmund operators. As a corollary, we prove new characterizations of product BMO in terms of commutators of Calder´on–Zygmund operators.

The main theorem is the following:

Theorem (4.1.1)[106]: Let $b \in BMO_{\text{prod}}(\mathbb{R}^{\vec{d}})$ and $(T_i)_{1\leq i\leq t}$ be a collection of Calder´on–Zygmund operators, with each T_i acting on parameter i of $\mathbb{R}^{\vec{d}} = \mathbb{R}^{d_1} \otimes \cdots \otimes \mathbb{R}^{d_t}$. Then,

$$\left\|\left[\dots\left[[M_b, T_1]T_2\right]\dots, T_t\right]\right\|_{L^2 \to L^2} \lesssim \|b\|_{BMO_{\text{prod}}},$$

where C depends only on $\vec{d}$ and $\prod_{i=1}^{t}\|T_i\|_{CZ}$.
One of the interesting results implied directly by the theorem is that a perturbation of a collection of operators characterizing product BMO still characterizes product BMO. In other words, characterizing families such as the Riesz transforms are stable under small perturbations in the sense that the Calder´on–Zygmund operator norm of the perturbation terms are small. We organize this observation into the following corollary.

Corollary (4.1.2)[106]: Let $\left(T_{i,s_i}\right)_{1\le i\le t,1\le s_i\le n_i}$ be a family of Calder´on– Zygmund operators characterizing the space $BMO_{\text{prod}}(\mathbb{R}^{\vec{d}})$, that is,$\exists C_1, C_2 > 0$, such that

$$C_1\|b\|_{BMO_{\text{prod}}} \le \sup_{1\le i\le t,1\le s_i\le n_i} \left\|\left[\dots\left[\left[M_b, T_{1,s_1}\right]T_{2,s_2}\right]\dots, T_{t,s_t}\right]\right\|_{L^2\to L^2} \le C_2\|b\|_{BMO_{\text{prod}}}.$$

Then, $\exists\epsilon > 0$ such that for any family of Calder´on–Zygmund operators $\left(T'_{i,s_i}\right)_{1\le i\le t,1\le s_i\le n_i}$ satisfying $\left\|T'_{i,s_i}\right\|_{CZ} \le \epsilon$, the family $\left(T_{i,s_i} + T'_{i,s_i}\right)_{1\le i\le t,1\le s_i\le n_i}$ still characterizes $BMO_{\text{prod}}\left(\mathbb{R}^{\vec{d}}\right)$.

In particular, since Calder´on–Zygmund operators form a linear space, whose norm can be made arbitrarily small by multiplying a small constant, it means that once we have a collection of operators characterizing BMO, we automatically obtain infinitely many collections of operators which also characterize BMO. More specifically, let$\left(T_{i,s_i}\right)_{1\le i\le t,1\le s_i\le n_i}$ be a family as in the corollary above, for any arbitrary family of Calder´on–Zygmund operators$\left(T'_{i,s_i}\right)_{1\le i\le t,1\le s_i\le n_i}$, there exist $\epsilon_1, \dots, \epsilon_t > 0$ such that for any $0 < c_i < \epsilon_i$, $1 \le i \le t$, the family $\left(T_{i,s_i} + c_iT'_{i,s_i}\right)_{1\le i\le t,1\le s_i\le n_i}$ characterizes $BMO_{\text{prod}}\left(\mathbb{R}^{\vec{d}}\right)$.

We proof of the main theorem is the representation theorem by $T.P.$ Hytönen[109], which states that any Calder´on–Zygmund operator can be represented as an average of dyadic shift operators with respect to a probabilistic measure on a collection of dyadic grids. While the earliest version of this theorem appeared in[110], here we choose to apply a slightly diﬀerent one given in[109]. In our proof, we will reduce the problem to the upper bound for commutators with dyadic shifts. This is the first use of Hytönen's representation theorem to commutator theory. The novelty of this approach to the upper bound is twofold. First, the commutators with dyadic shifts which have infinite complexity in our case, are carefully studied and eﬀectively reduced to paraproducts and another class of bounded operators. In contrast to typical methods dealing with multi-parameter theory, this allows our argument to be iterated. Second, new paraproducts and a similar type of operators are introduced, and this is where the delicate estimates in product theory are required.

We recall several preliminary results on dyadic shifts, representation theorem, and multiparameter paraproducts. A full proof of the main theorem in its one-parameter case is introduced, while the proof of the main theorem in arbitrarily many parameters is presented.

We give some essential background for the proof of the main theorem.

Recall that while the standard dyadic grid is defined as

$$\mathfrak{D}^0 := \{2^{-k}([0,1)^d + m): k \in \mathbb{Z}, m\mathbb{Z}^d\},$$

for any parameter $\omega\left(\omega_j\right)_{j\in\mathbb{Z}} \in (\{0,1\}^d)^{\mathbb{Z}}$, one can define an associated shifted dyadic grid as

$$\mathfrak{D}^{\omega} := \{I\dot{+}\omega : I \in \mathfrak{D}^0\},$$

Where

$$I\dot{+}\omega := I + \sum_{j:2^{-j}<\ell(I)} 2^{-j}\omega_j.$$

For a fixed shifted grid $\mathfrak{D}^{\omega}$ and $i,j \in \mathbb{Z}_+$, a dyadic shift operator S_{ω}^{ij} is defined to be bounded on L^2 with operator norm less than 1. Specifically,

$$S_{\omega}^{ij}f := \sum_{K\in\mathfrak{D}^{\omega}} \sum_{\substack{I\in\mathfrak{D}^{\omega}, I\subset K \\ \ell(I)=2^{-i}\ell(K)}} \sum_{\substack{J\in\mathfrak{D}^{\omega}, J\subset K \\ \ell(J)=2^{-j}\ell(K)}} a_{IJK}\langle f, h_I\rangle h_J$$

$$=: \sum_K \sum_{I,J\subset K}^{(i,j)} a_{IJK}\langle f, h_I\rangle h_J,$$

with $\left|a_{IJK}\right| \le |I|^{1/2}|J|^{1/2}/|K|$. S_{ω}^{ij} is called cancellative if all the Haar functions in the definition are cancellative, otherwise, it is called non- cancellative.

Recall that in one dimension, any dyadic interval I is associated with a cancellative Haar function $h_I^0 = |I|^{-1/2}\left(\chi_{I_l} - \chi_{I_r}\right)$ and a noncancellative one $h_I^1 = |I|^{-1/2}\chi_I$. While in d dimensions, each cube $I = I_1 \times\cdots\times I_d$ is associated with 2^d Haar functions:

$$h_I^{\epsilon}(x) = h_{I_1\times\ldots\times I_d}^{(\epsilon_1,\ldots,\epsilon_d)}(x_1,\ldots,x_d) = \prod_{i=1}^{d} h_{I_i}^{\epsilon_i}(x_i), \epsilon \in \{0,1\}^d,$$

where h_I^1 is called noncancellative, while all the other $2^d - 1$ Haar functions h_I^{ϵ} for $\epsilon \in \{0,1\}^d\backslash\{1\}$ are cancellative. Note that all the cancellative Haar functions for a fixed grid form an orthonormal basis of $L^2(\mathbb{R}^d)$. We usually suppress the parameter ϵ to abbreviate the notation.

We now introduce T. P. Hytönen's representation theorem, a key tool in our proof. Interested readers can find its proof and a more detailed discussion in[109] and[110]. The operator T mentioned in the following will denote a Calder´on–Zygmund operator associated with a δ-standard kernel K. T. P. Hytönen[109] proved the following theorem.

Theorem (4.1.3)[106]: Let T be a Calder´on–Zygmund operator, then it has an expansion, say for $f, g \in C_0^{\infty}(\mathbb{R}^d)$,

$$\langle g, Tf\rangle = c.\|T\|_{CZ}.\mathbb{E}_{\omega} \sum_{i,j=0}^{\infty} 2^{-\max(i,j)\delta/2}\langle g, S_{\omega}^{ij}f\rangle,$$

where c is a dimensional constant and S_{ω}^{ij} is a dyadic shift of parameter (i,j) on the dyadic grid $\mathcal{D}^{\omega}$; all of them except possibly S_{ω}^{00} are cancellative.

According to the proof of Theorem (4.1.3), in the representation of any T, only S_{ω}^{00} may be noncancellative, and if this is the case, only one of $\{h_I\}, \{h_J\}$ in its definition is noncancellative, i.e. S_{ω}^{00} is a paraproduct with some BMO symbol a satisfying $\|a\|_{BMO} \le 1$ and $a_I = \langle a, h_I\rangle|I|^{-1/2}$, $\forall\, I \in \mathcal{D}$.

Recall that a multi-parameter paraproduct associated with function b can be viewed as a bilinear operator which is defined as

$$B_0(b,f)=\sum_{R\in\mathfrak{D}_{\vec{d}}}\beta_R\,\langle b,h_R^{\epsilon_1}\rangle\langle f,h_R^{\epsilon_2}\rangle h_R^{\epsilon_3}|R|^{-\frac{1}{2}},$$

where $\epsilon_j\in\{0,1\}^{\vec{d}}$, $\mathfrak{D}_{\vec{d}}$denotes the tensor product of dyadic grids, and$\{\beta_R\}_R$is a sequence satisfying$|\beta_R|\le 1$. Note that $h_R^{\epsilon_j}$is cancellative if and only if$\epsilon_j\neq\vec{1}$. According to Journ´e[16] and later on improved by$C$. Muscalu,J. Pipher, T. Tao, and C. Thiele [18], [12], one has the following boundedness result.

Theorem (4.1.4)[106]: Let $\vec{d}=(d_1,\dots,d_t)$ and $\epsilon_j(\epsilon_{j,1},\dots,\epsilon_{j,t})$. If$\epsilon_j\neq\vec{1}and\forall\ 1\le s\le t$, there is at most one of $j=2,3$ such that $\epsilon_{j,s}=\vec{1}$, then the operatorB_0satisfies

$$B_0\colon BMO_{\text{prod}}\left(\mathbb{R}^{\vec{d}}\right)\times L^2\left(\mathbb{R}^{\vec{d}}\right)\to L^2\left(\mathbb{R}^{\vec{d}}\right).$$

We present a detailed proof of the main theorem in the one-parameter setting, which will later on be utilized to prove the multi-parameter result. As an essential part of the proof, delicate estimates of new paraproducts and a new operator P will be introduced.

Given a BMO function b and a Calder´on–Zygmund operator T, one could represent the commutator$[b,T]$ as an average of$[b,S_\omega^{ij}]$ due to Theorem (4.1.3). Then, inorder to prove the upper bound inequality, it suffices to prove that for any $f\in C_0^\infty(\mathbb{R}^d)$,

$$\left\|\sum_{i,j=0}^{\infty}2^{-\max(i,j)\delta/2}\left[b,S_\omega^{ij}\right]\right\|_{L^2}\lesssim\|b\|_{BMO}\|f\|_{L^2}\qquad(1)$$

uniformly in ω. In the following we will write$S^{ij}$for short as the argument doesn't depend on ω explicitly.

As a crucial ingredient in our argument, two kinds of paraproduct-like operators need to be introduced.

The first one is the bilinear operator$B_k$which could be viewed as a generalized dyadic paraproduct:

$$B_k(b,f)\coloneqq\sum_I\beta_I\langle b,h_{I(k)}\rangle\langle b,h_I\rangle h_I\left|I^{(k)}\right|^{-\frac{1}{2}},$$

where$\{\beta_I\}_I$is a sequence satisfying$|\beta_I|\le 1$, $k\ge 0$ is an arbitrary integer, and$I^{(k)}$ denotes the k-th dyadic ancestor of$I$. Note that when $k=0$, this is exactly the classical paraproduct that we have introduced at the end of the previous, whose boundedness is stated in Theorem (4.1.4). Lemma (4.1.6) below shows that such boundedness holds uniformly for any B_k.

The second one is the trilinear operator P defined as

$$P(b,a,f)\sum_I\langle b,h_I\rangle\langle f,h_I\rangle|I|^{-1}\sum_{J:J\subsetneq I}\langle a,h_J\rangle h_J,$$

which will be proved to be bounded on$BMO\times BMO\times L^2\to L^2$ in Lemma$(4.1.7)$.

The main theorem we will prove is the following:

Theorem (4.1.5)[106]: For cancellative dyadic shift S^{ij}, $\left[b,S^{ij}\right]f$ can be represented as a finite linear combination of the following terms:

$$S^{ij}\left(B_k(b,f)\right),B_k\left(b,S^{ij}f\right),\qquad(2)$$

where the integer k is such that $0\le k\le\max(i,j)$and the total number of terms is bounded by $C(1+\max(i,j))$ for some universal dimensional constant C.

For noncancellative dyadic shift S^{00}(dyadic paraproduct) with symbol a, $[b, S^{00}]f$ can be represented as a finite linear combination of the following terms:

$$S^{00}\big(B_0(b,f)\big), B_0(b, S^{00}f), P(b,a,f), P^*(b,a,f), \tag{3}$$

where P^* is understood as the adjoint of P with b and a fixed, and the total number of terms is bounded by a universal dimensional constant.

Lemma (4.1.6)[106]: Given $b \in BMO(\mathbb{R}^d)$ and $k \geq 0$, let

$$B_k(b,f) := \sum_I \beta_I \langle b, h_{I(k)}\rangle\langle b, h_I\rangle h_I \left|I^{(k)}\right|^{-\frac{1}{2}},$$

where all the Haar functions are cancellative. Then $\|B_k(b,f)\|_{L^2} \lesssim \|b\|_{BMO}\|f\|_{L^2}$ with a constant independent of k.

Before we proceed to its proof, note that for the application to our problem, there is no need to include cases when some of the Haar functions in B_k are noncancellative according to the remark above. Hence, $B_k(b,f)$ is in fact a martingale transform whose uniform boundedness follows directly from the observation $\left|\langle b, h_{I(k)}\rangle\right| / \left|I^{(k)}\right|^{-1/2} \leq \|b\|_{BMO}$. However, we will present a different proof via square function in the following, which will provide some insight into the estimates of some other operators and the multi-parameter analogs of the result, where noncancellative Haar functions have to be taken into account.

Proof: For any $g \in L^2(\mathbb{R}^d)$,

$$\langle B_k(b,f), g\rangle = \langle b, \sum_I \beta_I \langle f, h_I\rangle\langle g, h_I\rangle h_{I(k)} \left|I^{(k)}\right|^{-\frac{1}{2}}\rangle.$$

It thus suffices to show that

$$\left\|\sum_I \beta_I \langle f, h_I\rangle\langle g, h_I\rangle h_{I(k)} \left|I^{(k)}\right|^{-\frac{1}{2}}\right\|_{H^1} \lesssim \|f\|_{L^2}\|g\|_{L^2},$$

which is equivalent to

$$\left\|S\left(\sum_I \beta_I \langle f, h_I\rangle\langle g, h_I\rangle h_{I(k)} \left|I^{(k)}\right|^{-\frac{1}{2}}\right)\right\|_{L^1} \lesssim \|f\|_{L^2}\|g\|_{L^2},$$

where in the above S denotes the dyadic square function. To see this, write

$$S\left(\sum_I \beta_I \langle f, h_I\rangle\langle g, h_I\rangle h_{I(k)} \left|I^{(k)}\right|^{-\frac{1}{2}}\right)^2$$

$$= \sum_J \left(\sum_{I:I^{(k)}=J} \beta_I \langle f, h_I\rangle\langle g, h_I\rangle h_{I(k)} |J|^{-\frac{1}{2}}\right)^2 \frac{\chi_J}{|J|}$$

which together with $\|.\|_{\ell^2} \leq \|.\|_{\ell^1}$ and Cauchy–Schwarz inequality implies

$$S\left(\sum_I \beta_I \langle f, h_I\rangle\langle g, h_I\rangle h_{I(k)} \left|I^{(k)}\right|^{-\frac{1}{2}}\right)$$

$$\leq \sum_J \left(\sum_{I:I^{(k)}=J} |\langle f,h_I\rangle||\langle g,h_I\rangle| \frac{\chi_J}{|J|} \right)$$

$$\leq \sum_J \left(\sum_{I:I^{(k)}=J} |\langle f,h_I\rangle|^2 \right)^{\frac{1}{2}} \left(\sum_{I:I^{(k)}=J} |\langle g,h_I\rangle|^2 \right)^{\frac{1}{2}} \frac{\chi_J}{|J|}$$

$$\leq \left(\sum_J \sum_{I:I^{(k)}=J} |\langle f,h_I\rangle|^2 \frac{\chi_J}{|J|} \right)^{\frac{1}{2}} \left(\sum_J \sum_{I:I^{(k)}=J} |\langle g,h_I\rangle|^2 \frac{\chi_J}{|J|} \right)^{\frac{1}{2}}$$

$$=: \left(S^{(k)}f\right)\left(S^{(k)}g\right),$$

where the operator $S^{(k)}f \coloneqq \left(\sum_I \sum_{I:I^{(k)}=J} |\langle f,h_I\rangle|^2 |J|^2 \chi_J\right)^{1/2}$. We claim that $S^{(k)}: L^2 \to L^2$ with norm bounded by a dimensional constant, which does not depend on k. This guarantees that our estimate of B_k becomes independent of k. Combining this with another use of Cauchy– Schwarz will complete the proof.

To show the claim, denote $\alpha_J = \left(\sum_{I:I^{(k)}=J} |\langle f,h_I\rangle|^2\right)^{1/2}$for any J and define $F(x) = \sum_J \alpha_J h_J(x)$. Then

$$\left\|S^{(k)}f\right\|_{L^2}^2 = \left\| \left(\sum_J \alpha_J^2 \frac{\chi_J}{|J|} \right)^{\frac{1}{2}} \right\|_{L^2}^2 = \|SF\|_{L^2}^2 \lesssim \|F\|_{L^2}^2$$

$$= \sum_J \alpha_J^2 = \sum_J \sum_{I:I^{(k)}=J} |\langle f,h_I\rangle|^2 = \sum_I |\langle f,h_I\rangle|^2 = \|f\|_{L^2}^2,$$

where the second to last equality holds because that cube I in the previous summation ranges over all the dyadic cubes exactly once.

Lemma (4.1.7)[106]: For tri-linear operator

$$P(b,a,f) \coloneqq \sum_I \langle b,h_I\rangle\langle f,h_I\rangle |I|^{-1} \sum_{J:J\subsetneq I} \langle a,h_J\rangle h_J,$$

there holds

$$\|P(b,a,f)\|_{L^2} \lesssim \|b\|_{BMO}\|a\|_{BMO}\|f\|_{L^2}.$$

Proof: The idea of the proof is to employ the H^1-BMO duality and the square function characterization of H^1. For any normalized test function $g \in L^2$,

$$\langle P(b,a,f),g\rangle = \langle b, \sum_I \langle f,h_I\rangle |I|^{-1} h_J \sum_{J:J\subsetneq I} \langle a,h_J\rangle\langle g,h_J\rangle\rangle.$$

To see where the BMO norm of a comes into play, observe that for any fixed I and some $1 < p < 2$,

$$\left| \sum_{J:J\subsetneq I} \langle a,h_J\rangle\langle g,h_J\rangle \right| = \left| \sum_{J:J\subsetneq I} \langle a,h_J\rangle h_J, g\chi_I \right|$$

$$\leq \left\| \sum_{J:J\subsetneq I} \langle a,h_J\rangle h_J \right\|_{L^{p'}} \|g\chi_I\|_{L^p}$$

$$\lesssim \left\| \left(\sum_{J:J\subsetneq I} \left|\langle a, h_J\rangle\right|^2 \frac{\chi_J}{|J|} \right)^{\frac{1}{2}} \right\|_{L^{p'}} \|g\chi_I\|_{L^p}$$

$$\lesssim \|a\|_{BMO}|I|^{1/p'}\|g\chi_I\|_{L^p} = \|a\|_{BMO}|I|(\langle |g|^p\rangle_I)^{1/p},$$

where the last inequality follows from John–Nirenberg inequality.
Therefore,

$$S\left(\sum_I \langle f, h_I\rangle |I|^{-1} h_J \sum_{J:J\subsetneq I} \langle a, h_J\rangle\langle g, h_J\rangle\right)$$

$$= \left(\sum_I |\langle f, h_I\rangle|^2 |I|^{-2} \left(\sum_{J:J\subsetneq I} \langle a, h_J\rangle\langle g, h_J\rangle\right)^2 \frac{\chi_I}{|I|}\right)^{\frac{1}{2}}$$

$$\leq \|a\|_{BMO} \left(\sum_I |\langle f, h_I\rangle|^2 \left(\langle |g|^p\rangle_I\right)^{2/p} \frac{\chi_I}{|I|}\right)^{\frac{1}{2}}$$

$$\leq \|a\|_{BMO} \left(\sum_I |\langle f, h_I\rangle|^2 \sup_{I:x\in I} \left(\langle |g|^p\rangle_I\right)^{2/p} \frac{\chi_I}{|I|}\right)^{\frac{1}{2}}$$

$$\leq \|a\|_{BMO} M(|g|^p)^{1/p} S(f),$$

where M is the Hardy–Littlewood maximal function which is bounded on $L^p, 1 < p < \infty$.
Hence,

$$\|P(b, a, f)\|_{L^2} \lesssim \|b\|_{BMO}\|a\|_{BMO} M(|g|^p)^{1/p} \|S(f)\|_{L^2}$$

$$\lesssim \|b\|_{BMO}\|a\|_{BMO}\|f\|_{L^2}.$$

Now we turn to the proof of Theorem(4.1.5) and the strategy is the following. First, we decompose b and f using Haar bases. Second, we split the sum into several parts and represent each of them as a linear combination of terms in Theorem (4.1.5).
To start with, one decomposes$[b, S^{ij}]f$as

$$[b, S^{ij}]f = \sum_{I,J} \langle b, h_J\rangle\langle f, h_J\rangle [h_J, S^{ij}]\, h_J$$

$$= \sum_{I,J} \langle b, h_J\rangle\langle f, h_J\rangle \left(h_I, S^{ij} h_J - S^{ij}(h_I h_J)\right) =: \mathrm{I} + \mathrm{II},$$

where in the following I and II will be referred to as first term and second term, respectively. In order to further organize the sum and extract the correct paraproduct structure, even in the simplest one-parameter case, one needs to divide up the sum into many different parts, depending on the relative sizes ofI, J.

Let's first look at the case when S^{ij} is cancellative, meaning that all the Haar functions appearing are cancellative. Hence,

$$[b, S^{ij}]f = \sum_{I,J} \langle b, h_J\rangle\langle f, h_J\rangle \left(h_J \sum_{J'\subset J^{(i)}}^{(j)} a_{JJ'J^{(i)}} h_{J'} - \sum_K \sum_{I'',J''\subset K}^{(i,j)} a_{I''J''K} \langle h_I h_J, h_{I''}\rangle h_{J''} \right).$$

First, we claim that it suffices to consider the part $I \subset J^{(i)}$. Indeed, it is obvious that when $I \cap J^{(i)} = \emptyset$, both terms in the parentheses are zero. Furthermore, by the cancellation structure of the commutator, when $I \supsetneq J^{(i)}$, the term $[h_I, S^{ij}]h_J$ is also zero. To see this, as h_I is constant on $J^{(i)}$, fixing an arbitrary $x_0 \in J^{(i)}$ implies

$$h_I S^{ij} h_J - S^{ij}(h_I h_J) = h_I(x_0) S^{ij}(h_I(x_0) h_J) = 0.$$

Note that for the case $(i,j) \neq (0,0)$, this is the only part of the proof where one needs the particular cancellation of the commutator structure. Next, we represent the first term and the second term separately.

Based on the discussion above, for any i, j, the first term containing $h_I S^{ij} h_J$ is equal to

$$\sum_J \sum_{I: I \subset J^{(i)}} \langle b, h_I \rangle \langle f, h_J \rangle h_I \sum_{\substack{J': J' \subset J^{(i)} \\ \ell(J') = 2^{i,j} \ell(J)}} a_{JJ'J^{(i)}} h_{J'}.$$

Introducing index $K = J^{(i)}$ allows us to rewrite this as

$$\sum_K \sum_{J: J \subset K}^{(i)} \sum_{I: I \subset K} \langle b, h_I \rangle \langle f, h_J \rangle h_I \sum_{J': J' \subset K}^{(i)} a_{JJ'K} h_{J'}$$

$$\sum_I \langle b, h_I \rangle h_I \left(\sum_{K: K \supset I} \sum_{J: J \subset K}^{(i)} \sum_{J': J' \subset K}^{(i)} a_{JJ'K} \langle f, h_J \rangle h_{J'} \right).$$

Comparing the inner parentheses to the definition of S^{ij} suggests that the expression above is equal to

$$\sum_I \langle b, h_I \rangle h_I \sum_{J': J'^{(j)} \supset I} \langle S^{ij} f, h_{J'} \rangle h_{J'}$$

$$= \sum_I \sum_{J': J' \supsetneq I} \langle b, h_I \rangle \langle S^{ij} f, h_{J'} \rangle h_I h_{J'}$$

$$+ \sum_I \sum_{J': J' \subset I \subset J'^{(j)}} \langle b, h_I \rangle \langle S^{ij} f, h_{J'} \rangle h_I h_{J'} =: \mathrm{I} + \mathrm{II}.$$

Note that there are only parts I and II left because of the supports of Haar functions. For part I, one writes

$$\mathrm{I} = \sum_I \langle b, h_I \rangle h_I \left(\sum_{J': J' \supsetneq I} \langle S^{ij} f, h_{J'} \rangle h_{J'} \right) = \sum_I \langle b, h_I \rangle h_I \langle S^{ij} f, h_I^I \rangle h_I^I$$

$$= \sum_I \langle b, h_I \rangle \langle S^{ij} f, h_I^I \rangle h_I |I|^{-\frac{1}{2}},$$

which is of type $B_0 \langle b, S^{ij} f \rangle$. In order to deal with part II, observe that it can be decomposed into finitely many pieces depending on the relative sizes of I and J', i.e.

$$\mathrm{II} = \sum_{k=0}^{j} \sum_{J'} \langle b, h_{J'^{(k)}} \rangle \langle S^{ij} f, h_{J'} \rangle h_{J'^{(k)}} h_{J'}$$

$$= \sum_{k=0}^{j} \sum_{J'} \beta_{J'} \langle b, h_{J'^{(k)}} \rangle \langle S^{ij} f, h_{J'} \rangle h_{J'} \left| J'^{(k)} \right|^{-\frac{1}{2}} = \sum_{k=0}^{j} B_k \left(b, S^{ij} f \right),$$

where $\beta_{J'} \in \{1, -1\}$ and $0 \leq k \leq j$. Note that the sum at the end contains only $1 + j \leq 1 + \max(i, j)$ terms. Therefore, the representation of the first term is demonstrated.

Now we turn to the second term that contain $S^{ij}\left(h_I h_J\right)$. Due to the supports of Haar functions, this part is nontrivial only when $I \cap J \neq \emptyset$. Hence, one can split this term into three parts: $I \subsetneq J, I = J$, and $J \subsetneq I \subset J^{(i)}$.

For $I \subsetneq J$, note that the second term becomes

$$S^{ij} \left(\sum_{I \subsetneq J} \langle b, h_I \rangle \langle f, h_J \rangle h_I h_J \right) = S^{ij} \left(\sum_{I} \langle b, h_I \rangle h_I \sum_{I \subsetneq J} \langle b, h_I \rangle h_J \right)$$

$$= S^{ij} \left(\sum_{I} \langle b, h_I \rangle h_I \langle f, h_I^I \rangle h_I^I \right)$$

$$= S^{ij} \left(\sum_{I} \langle b, h_I \rangle \langle f, h_I^I \rangle h_I |I|^{-\frac{1}{2}} \right),$$

which is $S^{ij}\left(B_0(b, f)\right)$.

As the diagonal part $I = J$ is obviously of the form $S^{ij}\left(B_0(b, f)\right)$ already, we move on to the last piece $J \subsetneq I \subset J^{(i)}$, which can be written as

$$S^{ij} \left(\sum_{J} \sum_{I : J \subsetneq I \subset J^{(i)}} \langle b, h_I \rangle \langle f, h_I \rangle h_I h_J \right).$$

Observe that what's inside the parentheses is of an almost identical form as part II that appeared at the end of the discussion of the first term except that j is changed to i and that f takes the place of $S^{ij} f$. Hence, the same reasoning implies that it is a sum of at most $i \leq \max(i, j)$ terms of $S^{ij}\left(B_k(b, f)\right)$, $1 \leq k \leq i$. This proves the representation of the second term as well as completes the discussion of the case when S^{ij} is cancellative.

It suffices to assume that

$$S^{00} f = \sum_{I} a_I \langle f, h_I^I \rangle h_I,$$

where $a_I \coloneqq \langle a_I, h_I \rangle |I|^{-1/2}$ with $\|a\|_{BMO} \leq 1$. Because if we switch the positions of cancellative and noncancellative Haar functions, what we obtain is none other than its adjoint. Moreover, for the Haar expansion

$$[b, S^{00}] f = \sum_{I,J} \langle b, h_I \rangle \langle f, h_J \rangle [h_I, S^{00}] h_J,$$

it is not hard to see, according to a discussion similar to the one at the beginning of the case $(i, j) \neq (0,0)$, that one needs only to consider the part $I \subset J$ thanks to the commutator structure. We then split the sum into two parts: $I \subsetneq J$ and $I = J$.

We consider the first term containing $h_I S^{00} h_J$ and the second term containing $S^{00}\left(h_I h_J\right)$ separately, without need to exploit more of the cancellation of the commutator. The second term can be dealt with exactly the same as how we treated the $I \subsetneq J$ part of the second term in the case$(i,j) \neq (0,0)$, which we omit. To study the first term, one observes that for any h_J,

$$S^{00}h_J = \sum_{I \subsetneq J} a_I \langle h_J, h_I^I \rangle h_I = \sum_{I \subsetneq J} a_I |I|^{\frac{1}{2}} h_I h_J.$$

Hence, the first term becomes

$$\sum_J \sum_{I:I' \subsetneq J} \langle b, h_I \rangle h_I \langle f, h_J \rangle a_{I'} |I'|^{\frac{1}{2}} h_{I'} h_J = \sum_J \sum_{I \subset I' \subsetneq J} + \sum_J \sum_{I' \subsetneq I \subsetneq J} =: I + II.$$

One writes

$$I = \sum_I \langle b, h_I \rangle h_I \left(\sum_{I':I \subset I'} \sum_{J:I' \subsetneq J} a_{I'} \langle f, h_J \rangle h_J |I'|^{\frac{1}{2}} h_{I'} \right)$$

$$= \sum_I \langle b, h_I \rangle h_I \left(\sum_{I':I \subset I'} a_{I'} |I'|^{\frac{1}{2}} h_{I'} \langle f, h_J \rangle h_{I'}^1 \right)$$

$$= \sum_I \langle b, h_I \rangle h_I \left(\sum_{I':I \subset I'} a_{I'} \langle f, h_I^1 \rangle h_{I'} \right)$$

$$= \sum_I \langle b, h_I \rangle h_I \left(\sum_{I':I \subset I'} a_{I'} \langle S^{00} f, h_{I'} \rangle h_{I'} \right)$$

$$= \sum_I \langle b, h_I \rangle h_I \langle S^{00} f, h_I \rangle \, h_I + \sum_I \langle b, h_I \rangle h_I \langle S^{00} f, h_I^1 \rangle h_I^1$$

$$= \sum_I \beta_I \langle b, h_I \rangle \langle S^{00} f, h_I \rangle h_I^{\epsilon} |I|^{\frac{1}{2}} + \sum_I \langle b, h_I \rangle \langle S^{00} f, h_I^1 \rangle |I|^{-\frac{1}{2}},$$

which is the sum of two $B_0(b, S^{00} f)$ with $\beta_I \in \{1, -1\}$.
To deal with partII, observe that

$$\mathrm{II} = \sum_{I' \subsetneq I} \langle b, h_I \rangle h_I a_{I'} |I'|^{\frac{1}{2}} h_{I'} \langle f, h_I^1 \rangle h_I^1,$$

by first summing over index$J$. Thus,

$$\mathrm{II} = \sum_{I' \subsetneq I} a_{I'} |I'|^{\frac{1}{2}} h_{I'} \left(\sum_{I:I \supsetneq I'} \langle b, h_I \rangle |I|^{-\frac{1}{2}} \langle f, h_I^1 \rangle h_I \right)$$

$$=: \sum_{I'} a_{I'} |I'|^{\frac{1}{2}} h_{I'} \sum_{I:I \supsetneq I'} \langle S_b f, h_I \rangle h_I$$

$$= \sum_{I'} a_{I'} \langle S_b f, h_I^1 \rangle h_{I'} = S^{00} \langle S_b f \rangle,$$

where the operator$S_b f \coloneqq \sum_I \langle b, h_I \rangle |I|^{-1/2} \langle f, h_I^1 \rangle h_I$is a classical para-product $B_0(b,f)$, and this completes the discussion of part$I \subsetneq J$.

In this special case, what we try to decompose becomes

$$\sum_I \sum_{\epsilon, \epsilon' \in \{0,1\}^d \{\vec{1}\}} \langle b, h_I^{\epsilon} \rangle \langle f, h_I^{\epsilon'} \rangle \left(h_I^{\epsilon} S^{00} h_I^{\epsilon'} - S^{00} \left(h_I^{\epsilon} h_I^{\epsilon'} \right) \right). \tag{4}$$

Here, in order to avoid possible confusion, we wrote out the sum over indexϵ, ϵ', explicitly. Recall that for each cube I, there are2^ddifferent Haar functions associated:$\{h_I^\epsilon\}, \epsilon \in \{0,1\}^d$, and the Haar function is noncancellative if and only if $\epsilon = \vec{1}$. First, it is useful to observe that if$\epsilon \neq \epsilon'$, $[h_I^\epsilon, S^{00}]h_I^{\epsilon'} = 0$. Indeed, for any fixed I andϵ, ϵ',

$$h_I^\epsilon S^{00} h_I^{\epsilon'} = \sum_{J:J\subsetneq I} a_J |J|^{\frac{1}{2}} h_J\left(h_I^\epsilon h_I^{\epsilon'}\right),$$

and

$$S^{00}\left(h_I^\epsilon h_I^{\epsilon'}\right) = \sum_{J:J\supset I} a_J |J|^{-\frac{1}{2}} h_J \left(\int_1 h_I^\epsilon h_I^{\epsilon'}\right) + \sum_{J:J\subsetneq I} a_J |J|^{\frac{1}{2}} h_J\left(h_I^\epsilon h_I^{\epsilon'}\right).$$

As a result of cancellation and the fact that$\int_1 h_I^\epsilon h_I^{\epsilon'}$is nonzero if and only if$\epsilon = \epsilon'$, i.e. $h_I^\epsilon h_I^{\epsilon'} = |I|^{-1}\chi_I$, $[h_I^\epsilon, S^{00}]h_I^{\epsilon'} \neq 0$only when $\epsilon = \epsilon'$.Therefore, one can safely suppress the dependence onϵ when studying this part of the sum.

Furthermore, it is easily seen that the second term containing$S^{00}(h_I h_I)$here can be estimated exactly the same as before, it thus suffices to deal with the first term containing$h_I S^{00} h_I$, which is equal to

$$\sum_I \langle b, h_I\rangle\langle f, h_I\rangle h_I S^{00} h_I = \sum_I \langle b, h_I\rangle\langle f, h_I\rangle |I|^{-1} \sum_{J:J\subsetneq I} \left(a, h_J\right) h_J$$
$$= P(b, a, f),$$

hence the proof is complete.

We present the proof of the main theorem in the general setting by iterating the one-parameter result, i.e. Theorem(4.1.5). For the sake of brevity, we consider the bi-parameter case as an example, while the strategy can be easily generalized to work for arbitrarily many parameters. We show that the commutator can be represented as a finite linear combination of the bi-parameter analogs of terms in Theorem (4.1.5),for which one needs to define and estimate the following new bi-parameter operators, including all the possible "tensor products" of the one-parameter operators B_k and P.

Lemma (4.1.8)[106]: Given$b \in BMO_{prod}(\mathbb{R}^n \times \mathbb{R}^n)$and integers $k, l \geq 0$, define the following operators

$$B_{k,l}(b,f) = \sum_{I_1,I_2} \beta_{I_1 I_2} \langle b, h_{I_1^{(k)}} \otimes u_{I_2^{(l)}}\rangle \langle f, h_{I_1}^{\epsilon_1} \otimes u_{I_2^{\epsilon_2}}\rangle\, h_{I_1}^{\epsilon_1'} \otimes u_{I_2}^{\epsilon_2'} \left|I_1^{(k)}\right|^{-\frac{1}{2}} \left|I_2^{(l)}\right|^{-\frac{1}{2}},$$

where$\beta_{I_1 I_2}$ is a sequence satisfying$\left|\beta_{I_1 I_2}\right| \leq 1$. When $k > 0$, all the Haar functions in the first variable are cancellative, while when $k = 0$, there is at most one of $h_I^\epsilon, h_I^{\epsilon'}$being noncancellative. The same assumption goes for the second variable. Then,$\left\|B_{k,l}(b,f)\right\|_{L^2} \lesssim \|b\|_{BMO_{prod}} \|f\|_{L^2}$with a constant independent of k, l.

In the above, we use$u_{I_2}$to denote Haar functions in the second variable, for any dyadic cube $I_2 \subset \mathbb{R}^m$.

Note that when $k = l = 0$, $B_{k,l}$becomes the classical bi-parameter B_0. When all the Haar functions are cancellative, the proof of the lemma proceeds exactly the same as its one-parameter counterpart, except that one needs bi-parameter dyadic square function as majorization instead. Therefore in the following, we will only prove the lemma assuming that $k = 0, l > 0$, and $h_{I_1}^{\epsilon_1} = h_{I_1}^1$ is the only noncancellative Haar. Note that in the setting of arbitrarily many parameters, parallel results still hold.

Proof: We are going to follow the strategy in the proof of Lemma (4.1.6) and use hybrid maximal-square functions as majorization.

Pairing $B_{0,l}(b,f)$with a normalized L^2 function g and applying the product H^1-BMO duality, it suffices to show that

$$\left\| SS\left(\sum_{I_1,I_2} \beta_{I_1I_2}\langle b, h^1_{I_1}\otimes u_{I_2}\rangle\langle g, h_{I_1}\otimes u_{I_2}\rangle h_{I_1}\otimes u_{I_2^{(l)}}\,|I_1|^{-\frac{1}{2}}\left|I_2^{(l)}\right|^{-\frac{1}{2}}\right)\right\|_{L^1} \lesssim \|f\|_{L^2},$$

where SS is the dyadic double square function whoseL^1norm characterizes product H^1. To see this, one calculates

$$SS\left(\sum_{I_1,I_2} \beta_{I_1I_2}\langle b, h^1_{I_1}\otimes u_{I_2}\rangle\langle g, h_{I_1}\otimes u_{I_2}\rangle h_{I_1}\otimes u_{I_2^{(l)}}\,|I_1|^{-\frac{1}{2}}\left|I_2^{(l)}\right|^{-\frac{1}{2}}\right)^2$$

$$=\sum_{I_1,I_2}\left(\sum_{J_1:J_2^{(l)}=I_2}\langle f, h^1_{I_1}\otimes u_{I_2}\rangle\langle g, h_{I_1}\otimes u_{I_2}\rangle\,|I_1|^{-\frac{1}{2}}|I_2|^{-\frac{1}{2}}\right)^2\frac{\chi_{I_1}\otimes\chi_{I_2}}{|I_1||I_2|}$$

$$\leq\sum_{I_1}\left(\sum_{I_2}\sum_{J_2:J_2^{(l)}=I_2}\sup_{I_1}\left(\langle\langle f, u_{J_2}\rangle_2\rangle_{I_1}\right)\langle g, h_{I_1}\otimes u_{I_2}\rangle\frac{\chi_{I_2}}{|I_2|}\right)^2\frac{\chi_{I_1}}{|I_1|},$$

where the last inequality follows from$\|\cdot\|_{\ell^2}\leq\|\cdot\|_{\ell^1}$, and$\langle\cdot\rangle_{I_1}$denotes the average value over I_1. Then the above is controlled by

$$\sum_{I_1}\left(\sum_{I_2}\sum_{J_2:J_2^{(l)}=I_2}M_1\left(\langle f, u_{J_2}\rangle_2\right)\langle g, h_{I_1}\otimes u_{I_2}\rangle\frac{\chi_{I_2}}{|I_2|}\right)^2\frac{\chi_{I_1}}{|I_1|},$$

where M_1 is the Hardy–Littlewood maximal function in the first variable. Next, Cauchy–Schwarz inequality implies that

$$\leq\sum_{I_1}\left(\sum_{I_2}\sum_{J_2:J_2^{(l)}=I_2}M_1\left(\langle f, u_{J_2}\rangle_2\right)^2\frac{\chi_{I_2}}{|I_2|}\right)$$

$$\times\left(\sum_{I_2}\sum_{J_2:J_2^{(l)}=I_2}\left|\langle g, h_{I_1}\otimes u_{I_2}\rangle\right|^2\frac{\chi_{I_2}}{|I_2|}\right)\frac{\chi_{I_1}}{|I_1|}$$

$$=\left(\sum_{I_2}\sum_{J_2:J_2^{(l)}=I_2}M_1\left(\langle f, u_{J_2}\rangle_2\right)^2\frac{\chi_{I_2}}{|I_2|}\right)$$

$$\times\left(\sum_{I_1}\sum_{I_2}\sum_{J_2:J_2^{(l)}=I_2}\left|\langle g, h_{I_1}\otimes u_{I_2}\rangle\right|^2\frac{\chi_{I_1}\otimes\chi_{I_2}}{|I_1||I_2|}\right)=:\mathrm{I}.\mathrm{II}.$$

II could be written as the square of SS acting on a normalized L^2 function, similarly as the last part of the proof of Lemma (4.1.6).For I, Fefferman–Stein inequality implies that

$$\left\|I^{\frac{1}{2}}\right\|_{L^2(\mathbb{R}^n\times\mathbb{R}^m)} = \left(\int_{\mathbb{R}^m}\left\|\left(\sum_{I_2}\sum_{J_2:J_2^{(l)}=I_2} M_1\left(\langle f,u_{J_2}\rangle_2\right)^2\frac{\chi_{I_2}}{|I_2|}\right)^{\frac{1}{2}}\right\|_{L^2(\mathbb{R}^n)}^2 dx_2\right)^{\frac{1}{2}}$$

$$\lesssim \left(\int_{\mathbb{R}^m}\left\|\left(\sum_{I_2}\sum_{J_2:J_2^{(l)}=I_2} \left|\langle f,u_{J_2}\rangle_2\right|^2\frac{\chi_{I_2}}{|I_2|}\right)^{\frac{1}{2}}\right\|_{L^2(\mathbb{R}^n)}^2 dx_2\right)^{\frac{1}{2}}$$

$$\lesssim \left(\int_{\mathbb{R}^m}\|f(.,x_2)\|_{L^2(\mathbb{R}^n)}^2\,dx_2\right)^{\frac{1}{2}} = \|f\|_{L^2(\mathbb{R}^n\times\mathbb{R}^m)},$$

where once again the last inequality is due to the same argument in the last part of the proof of Lemma(4.1.6),thus the proof is complete.

Lemma (4.1.9)[106]: Given$b,a\in BMO_{prod}(\mathbb{R}^n\times\mathbb{R}^m)$, define

$$PP(b,a,f):=\sum_{I_1,I_2}\langle b,h_{I_1}\otimes u_{I_2}\rangle\langle f,h_{I_1}\otimes u_{I_2}\rangle\,|I_1|^{-1}|I_2|^{-1}$$
$$\times\sum_{J_1:J_1\subsetneq I_1}\sum_{J_2:J_2\subsetneq I_2}\langle a,h_{I_1}\otimes u_{I_2}\rangle h_{I_1}\otimes u_{I_2},$$

and let PP_1be its partial adjoint in the first variable with b,a fixed. Then,

$$\|PP(b,a,f)\|_{L^2}\lesssim\|b\|_{BMO_{prod}}\|b\|_{BMO_{prod}}\|f\|_{L^2}, \quad (5)$$
$$\|PP_1(b,a,f)\|_{L^2}\lesssim\|b\|_{BMO_{prod}}\|b\|_{BMO_{prod}}\|f\|_{L^2}. \quad (6)$$

Recall that for a bi-parameter singular integral T, its partial adjoint T_1 is defined via

$$\langle T(f_1\otimes f_2),g_1\otimes g_2\rangle=\langle T_1(g_1\otimes f_2),f_1\otimes g_2\rangle.$$

It is known that theL^2boundedness of T does not imply the L^2 bound- edness ofT_1 (see [16] or [114] for a detailed discussion and counterexamples). Hence, in the following, we need to prove the boundedness of PP and PP_1 separately.

Proof: We first note that the proof of PP is essentially the same as Lemma (4.1.7).Inthebi-parametersetting, one needs to use the double square function SS to characterize product $H^1$and the strong maximal function M_S as majorization. The key observation is that there holds the following bi-parameter John–Nirenberg inequality (see[107]):

$$\left\|\left(\sum_{R\subset\Omega}|\langle a,h_R\rangle|^2\frac{\chi_R}{|R|}\right)^{\frac{1}{2}}\right\|_{L^p}\le\|a\|_{BMO_{prod}}|\Omega|^{1/p},1<p<\infty\,,$$

whereΩis any open set in $\mathbb{R}^n\times\mathbb{R}^m$ of finite measure, and R denotes dyadic rectangles. It thus easy to verify that a same argument as in Lemma(4.1.7) implies (5). The estimate of(6) involves the hybrid maximal-square functions, which we have seen in the proof of Lemma (4.1.9).Tobespecific, let $g\in L^2$be a normalized test function,

$$\langle PP_1(b,a,f),g\rangle$$
$$= \langle b, \sum_{I_1,I_2} |I_1|^{-1}|I_2|^{-1}\, h_{I_1}\otimes u_{I_2}$$
$$\times \sum_{J_1:J_1\subsetneq I_1} \sum_{J_2:J_2\subsetneq I_2} \langle a, h_{J_1}\otimes u_{J_2}\rangle\, \langle f, h_{J_1}\otimes u_{I_2}\rangle \langle g, h_{I_1}\otimes u_{J_2}\rangle\rangle.$$

Note that by bi-parameter John–Nirenberg inequality,

$$\left|\sum_{J_1:J_1\subsetneq I_1} \sum_{J_2:J_2\subsetneq I_2} \langle a, h_{J_1}\otimes u_{J_2}\rangle\, \langle f, h_{J_1}\otimes u_{I_2}\rangle \langle g, h_{I_1}\otimes u_{J_2}\rangle\right|$$
$$= \left|\sum_{J_1:J_1\subsetneq I_1} \sum_{J_2:J_2\subsetneq I_2} \langle a, h_{J_1}\otimes u_{J_2}\rangle\, \langle \langle f, u_{I_2}\rangle_2 \otimes \langle g, u_{I_2}\rangle_1, h_{I_1}\otimes u_{J_2}\rangle\right|$$
$$\leq \|a\|_{BMO_{prod}} |I_1||I_2| \left(\langle |\langle f, u_{I_2}\rangle_2|^p\rangle_{I_1}\right)^{1/p} \left(\langle |\langle g, h_{I_1}\rangle_1|^p\rangle_{I_2}\right)^{1/p},$$

for some$1 < p < 2$. Hence,

$$SS\left(\sum_{I_1,I_2} |I_1|^{-1}|I_2|^{-1}\, h_{I_1}\otimes u_{I_2} \times \sum_{J_1:J_1\subsetneq I_1} \sum_{J_2:J_2\subsetneq I_2} \langle a, h_{J_1}\otimes u_{J_2}\rangle\, \langle f, h_{J_1}\otimes u_{I_2}\rangle \langle g, h_{I_1}\otimes u_{J_2}\rangle\right)$$
$$\leq \|a\|_{BMO_{prod}} \left(\sum_{I_1,I_2} \left(\langle |\langle f, u_{I_2}\rangle_2|^p\rangle_{I_1}\right)^{2/p} \left(\langle |\langle g, h_{I_1}\rangle_1|^p\rangle_{I_2}\right)^{2/p} \frac{\chi_{I_1}\otimes\chi_{I_2}}{|I_1||I_2|}\right)^{\frac{1}{2}}$$
$$\leq \|a\|_{BMO_{prod}} \left(\sum_{I_2} M_1\left(|\langle f, u_{I_2}\rangle_2|^p\right)^{2/p} \frac{\chi_{I_2}}{|I_2|}\right)^{1/2} \times \left(\sum_{I_1} M_2\left(|\langle g, h_{I_2}\rangle_1|^p\right)^{2/p} \frac{\chi_{I_1}}{|I_1|}\right)^{1/2}.$$

The two terms on the last line above can be viewed as generalized hybrid maximal-square functions, whose boundedness is easy to obtain. For example,

$$\left\|\left(\sum_{I_2} M_1\left(|\langle f, u_{I_2}\rangle_2|^p\right)^{2/p} \frac{\chi_{I_2}}{|I_2|}\right)^{1/2}\right\|_{L^2}$$
$$= \left(\int_{\mathbb{R}^n} \left\|\left(\sum_{I_2} M_1\left(|\langle f, u_{I_2}\rangle_2|^p\right)^{2/p} \frac{\chi_{I_2}}{|I_2|}\right)^{\frac{1}{2}}\right\|_{L^2(\mathbb{R}^m)}^2 dx_1\right)^{\frac{1}{2}}$$
$$= \left(\int_{\mathbb{R}^n} \sum_{I_2} M_1\left(|\langle f, u_{I_2}\rangle_2|^p\right)^{2/p} dx_1\right)^{\frac{1}{2}}$$

$$\lesssim \left(\sum_{I_2} \int_{\mathbb{R}^n} \left| \langle f, u_{I_2} \rangle_2 \right|^2 dx_1 \right)^{\frac{1}{2}} = \|f\|_{L^2}.$$

Therefore, $\|PP_1(b, a, f)\|_{L^2} \lesssim \|b\|_{BMO_{prod}} \|a\|_{BMO_{prod}} \|f\|_{L^2}$.

In addition to the above two types of operators, in the bi-parameter setting, a new type of operator that mixes the paraproduct and P arise naturally in our argument. We show that they have the following uniform BMO estimates.

Lemma (4.1.10)[106]: Given $b \in BMO_{prod}(\mathbb{R}^n \times \mathbb{R}^m)$, $a^1 \in BMO(\mathbb{R}^n)$, and $a^2 \in BMO(\mathbb{R}^m)$. For integers $k, l \geq 0$, define

$$BP_k(b, a^2, f) := \sum_{I_1, I_2} \beta_{I_1} \langle b, h_{I_1^{(k)}} \otimes u_{I_2} \rangle \langle f, h_{I_1}^{\epsilon_1} \otimes u_{I_2} \rangle \left| I_1^{(k)} \right|^{-\frac{1}{2}} |I_2|^{-1} h_{I_1}^{\epsilon_1'}$$

$$\times \sum_{J_2 : J_2 \subsetneq I_2} \langle a^2, u_{J_2} \rangle_2 h_{J_2}$$

$$PB_l(b, a^1, f) := \sum_{I_1, I_2} \beta_{I_2} \langle b, h_{I_1} \otimes u_{I_2^{(l)}} \rangle \langle f, h_{I_1} \otimes u_{I_2}^{\epsilon_2} \rangle |I_1|^{-1} \left| I_2^{(l)} \right|^{-\frac{1}{2}} h_{I_2}^{\epsilon_2'}$$

$$\times \sum_{J_1 : J_1 \subsetneq I_1} \langle a^1, u_{J_1} \rangle_1 h_{J_1},$$

where β_{I_1}, β_{I_2} are sequences satisfying $\left|\beta_{I_1}\right|, \left|\beta_{I_2}\right| \leq 1$. When $k > 0$, all the Haar functions in the first variable are cancellative, while when $k = 0$, there is at most one of $h_{I_1}^{\epsilon_1}, h_{I_1}^{\epsilon_1'}$ being noncancellative. The same assumption goes for the second variable. Then, there holds

$$\|BP_k(b, a^2, f)\|_{L^2} \lesssim \|b\|_{BMO_{prod}} \|a^2\|_{BMO_{prod}} \|f\|_{L^2},$$

$$\|PB_l(b, a^1, f)\|_{L^2} \lesssim \|b\|_{BMO_{prod}} \|a^1\|_{BMO_{prod}} \|f\|_{L^2}.$$

Proof: By symmetry, it suﬃces to estimate PB_l. The strategy is similar as before: a square function argument encoding the product BMO estimate of b, combined with a John–Nirenberg inequality taking advantage of the BMO estimate of a^1. Note that the arguments slightly vary depending on whether noncancellative Haar functions appear. Taking g such that $\|g\|_{L^2} \leq 1$,

$$\langle PB_l(b, a^1, f), g \rangle$$

$$= \langle b, \sum_{I_1, I_2} \langle f, h_{I_1} \otimes u_{I_2}^{\epsilon_2} \rangle |I_1|^{-1} \left| I_2^{(l)} \right|^{-\frac{1}{2}} h_{I_1} \otimes u_{I_2}^{(l)}$$

$$\times \sum_{J_1 : J_1 \subsetneq I_1} \langle a^1, h_{J_1} \rangle_1 \langle g, h_{I_1} \otimes u_{I_2}^{\epsilon_2'} \rangle \rangle.$$

A similar application of John–Nirenberg inequality as before implies that

$$SS\left(\sum_{I_1, I_2} \langle f, h_{I_1} \otimes u_{I_2}^{\epsilon_2} \rangle |I_1|^{-1} \left| I_2^{(l)} \right|^{-\frac{1}{2}} h_{I_1} \otimes u_{I_2^{(l)}} \times \sum_{J_1 : J_1 \subsetneq I_1} \langle a^1, h_{J_1} \rangle_1 \langle g, h_{I_1} \otimes u_{I_2}^{\epsilon_2'} \rangle \right)$$

$$\leq \|a^1\|_{BMO}\left(\sum_{I_1,J_2}\left(\sum_{I_2\subset J_2}^{(l)}\langle f,h_{I_1}\otimes u_{I_2}^{\epsilon_2}\rangle\left(\langle\left|\langle g,u_{I_2}^{\epsilon_2'}\rangle_2\right|^p\rangle_{I_1}\right)^{\frac{1}{p}}\right)^2\frac{\chi_{I_1}\otimes\chi_{J_2}}{|I_1||J_2|}\right)^{\frac{1}{2}}. \quad (7)$$

(a) Case $l>0$.

In this case, all the Haar functions that appear are cancellative, hence by omitting the dependence on ϵ_2,ϵ_2' and applying Cauchy–Schwarz inequality, there holds

$$(7)\leq \|a^1\|_{BMO}\left(\sum_{I_1,J_2}\left(\sum_{I_2\subset J_2}^{(l)}\left|\langle f,h_{I_1}\otimes u_{I_2}\rangle\right|^2\right)\times\left(\sum_{I_2\subset J_2}^{(l)}\left(\langle\left|\langle g,u_{I_2}\rangle_2\right|^p\rangle_{I_1}\right)^{2/p}\right)\frac{\chi_{I_1}\otimes\chi_{J_2}}{|I_1||J_2|}\right)^{\frac{1}{2}}$$

$$\leq \|a^1\|_{BMO}\left(\sum_{J_2}\left(\sum_{I_1}\sum_{I_2\subset J_2}^{(l)}\left|\langle f,h_{I_1}\otimes u_{I_2}\rangle\right|^2\frac{\chi_{I_1}}{|I_1|}\right)\times\left(\sum_{I_2\subset J_2}^{(l)}M_1\left(\left|\langle g,u_{I_2}\rangle_2\right|^p\right)^{2/p}\right)\frac{\chi_{J_2}}{|J_2|}\right)^{\frac{1}{2}}.$$

which by $\|\cdot\|_{\ell^2}\leq\|\cdot\|_{\ell^1}$ and another use of Cauchy–Schwarz is bounded by

$$\|a^1\|_{BMO}\left(\sum_{I_1}\sum_{J_2}\sum_{I_2\subset J_2}^{(l)}\left|\langle f,h_{I_1}\otimes u_{I_2}\rangle\right|^2\frac{\chi_{I_1}\otimes\chi_{J_2}}{|I_1||J_2|}\right)^{\frac{1}{2}}$$

$$\times\left(\sum_{J_2}\sum_{I_2\subset J_2}^{(l)}M_1\left(\left|\langle g,u_{I_2}\rangle_2\right|^p\right)^{2/p}\frac{\chi_{J_2}}{|J_2|}\right)^{\frac{1}{2}}.$$

Therefore, a similar double square function and hybrid maximal-square function argument as in Lemma(4.1.8) and Lemma(4.1.9)implies that

$$(7)\lesssim\|a^1\|_{BMO}\|f\|_{L^2}\|g\|_{L^2}.$$

(b) Case $l=0$ and $\epsilon_2=\vec{1}$.

In this case,

$$(7)=\|a^1\|_{BMO}\left(\sum_{I_1,I_2}\left(\langle\langle f,h_{I_2}\rangle_1\rangle_{I_1}\right)\left(\langle\left|\langle g,u_{I_2}\rangle_2\right|^p\rangle_{I_1}\right)^{2/p}\frac{\chi_{I_1}\otimes\chi_{I_2}}{|I_1||I_2|}\right)^{\frac{1}{2}}.$$

$$\leq\left(\sum_{I_1}M_2\left(\langle f,h_{I_2}\rangle_1\right)^2\frac{\chi_{I_1}}{|I_1|}\right)^{\frac{1}{2}}\left(\sum_{I_2}M_1\left(\left|\langle g,u_{I_2}\rangle_2\right|^p\right)^{\frac{2}{p}}\frac{\chi_{I_2}}{|I_2|}\right)^{\frac{1}{2}}.$$

Which shows that

$$\|(7)\|_{L^1}\lesssim\|a^1\|_{BMO}\|f\|_{L^2}\|g\|_{L^2}.$$

(c) Case $l=0$ and $\epsilon_2'=\vec{1}$. This last case can be dealt with similarly by noticing that

$$(7)=\|a^1\|_{BMO}\left(\sum_{I_1,I_2}\left|\langle f,h_{I_1}\otimes u_{I_2}\rangle\right|^2\left(\langle\left|\langle g\rangle_{I_2}\right|^p\rangle_{I_1}\right)^{2/p}\frac{\chi_{I_1}\otimes\chi_{I_2}}{|I_1||I_2|}\right)^{\frac{1}{2}}.$$

$$\leq\|a^1\|_{BMO}\left(M_1(|M_2(g)|^p)\right)^{1/p}SS(f).$$

The boundedness of M_1 and M_2 in each variable implies that

$$\left\|\left(M_1(|M_2(g)|^p)\right)^{1/p}\right\|_{L^2} \lesssim \|g\|_{L^2}.$$

To conclude, we've demonstrated in each case that

$$\|PB_l(b,a^1,f)\|_{L^2} \lesssim \|b\|_{BMO_{prod}}\|a^1\|_{BMO_{prod}}\|f\|_{L^2}$$

which completes the proof.

Now let's proceed with the proof of Theorem(4.1.1).Using Theorem(4.1.3) twice for both variables we have

$$[[b,T_1],T_2]f = c\|T_1\|CZc\|T_2\|CZ\mathbb{E}_{\omega_1}\mathbb{E}_{\omega_2}$$

$$\times \sum_{i_1 j_1=0}^{\infty}\sum_{i_2 j_2=0}^{\infty} 2^{-\max(i_1 j_1)\frac{\delta}{2}}2^{-\max(i_2 j_2)\frac{\delta}{2}}\left[\left[b,S_{\omega_1}^{i_1 j_1}\right],S_{\omega_2}^{i_2 j_2}\right]f. \tag{8}$$

Since our estimate in the following doesn't depend on the parameters ω_1,ω_2 explicitly, we will omit them in the notation. Our goal is to prove that

$$\left\|\left[\left[b,S_1^{i_1 j_1}\right],S_2^{i_2 j_2}\right]f\right\|_{L^2(\mathbb{R}^n\times\mathbb{R}^m)}$$

$$\lesssim (1+\max(i_1,j_1))(1+\max(i_2,j_2))\|b\|_{BMO_{\mathrm{prod}}(\mathbb{R}^n\times\mathbb{R}^m)}\|f\|_{L^2(\mathbb{R}^n\times\mathbb{R}^m)},$$

which can be achieved by showing that any$\left[\left[b,S_1^{i_1 j_1}\right],S_2^{i_2 j_2}\right]f$can be represented as a finite linear combination of the following terms and their adjoints (which is understood as the adjoint operator withb,a^i fixed):

$$B_{k,l}\left(b,S_1^{i_1 j_1}S_2^{i_2 j_2},f\right), \qquad S_1^{i_1 j_1}\left(B_{k,l}\left(b,S_2^{i_2 j_2},f\right)\right), \tag{9}$$

$$BP_k\left(b,a^2,S_1^{i_1 j_1}f\right) \quad PB_l\left(b,a^1,S_2^{i_2 j_2}f\right). \tag{10}$$

$$PP(b,a^1\otimes a^2,f) \quad PP_1(b,a^1\otimes a^2,f), \tag{11}$$

where $k,l\geq 0$, and a^i is the BMO symbol of the dyadic shift S^{00}if it appears in the i-th variable. The total number of terms in the representation is no greater than $C(1+\max(i_1,j_1))(1+\max(i_2,j_2))$for some universal constant C. Note that for $a^1\in BMO(\mathbb{R}^m)$ and $a^2\in BMO(\mathbb{R}^m)$, there holds $a^1\otimes a^2\in BMO_{\mathrm{prod}}(\mathbb{R}^n\times\mathbb{R}^m)$. Hence, implied by Theorem(4.1.4), Lemma(4.1.8), Lemma(4.1.9), and Lemma(4.1.10), the L^2 norm of all of the terms above are uniformly bounded, independent of k,l in particular.

To derive the desired representation, we argue by an iteration of Theorem(4.1.5).

In the case when both $S_1^{i_1 j_1}$ and $S_2^{i_2 j_2}$ are cancellative, only operators $B_{k,l}$ need to be involved. In order to make the notations clear, in the following, we will use B_l^{τ} to denote the one-parameter paraproducts that appeared for the τ-th variable, where $k\geq 0$ and$\tau=1,2$. Calculation shows that

$$\left[\left[b,S_1^{i_1 j_1}\right],S_2^{i_2 j_2}\right]f = \sum_{I_1:J_1}\sum_{I_2:J_2}\langle b,h_{I_1}\otimes u_{I_2}\rangle\langle f,h_{J_1}\otimes u_{J_2}\rangle\left[h_{I_1},S_1^{i_1 j_1}\right]h_{J_1}\otimes\left[u_{I_2},S_2^{i_2 j_2}\right]u_{J_2},$$

which by iteration equals

$$\sum_{I_1:J_1}\left(\sum_{t_1\in\Lambda_1}B_{k,t_1}^1\left(\langle b,u_{I_2}\rangle_2,S_1^{i_1 j_1}(\langle f,u_{J_2}\rangle_2)\right)\right.$$

$$\left.+\sum_{t_2\in\Lambda_2}S_1^{i_1 j_1}\left(B_{k,t_2}^1(\langle b,u_{I_2}\rangle_2,\langle f,u_{J_2}\rangle_2)\right)\right)\otimes\left(\left[u_{I_2},S_2^{i_2 j_2}\right]u_{J_2}\right),$$

where B^1_{k,t_i} are paraproducts of type B^1_k in the first variable, and for each t_i, k is an arbitrary nonnegative integer.

Note that in the first parentheses we have a finite linear combination of terms that have already been studied, and all of the index set Λ_i satisfy $|\Lambda_i| \le C(1+\max(i_1,j_1)), i = 1,2$. Since the terms inside the first parentheses can be treated similarly, let's study one of the terms B^1_{k,t_1} as an example. We will also omit the subscript t_1 as the choice is arbitrary. Then, the sum corresponding to B^1_k is equal to

$$\sum_{I_1,J_1} B^1_k\left(\langle b,u_{I_2}\rangle_2, S_1^{i_1j_1}(\langle f,u_{J_2}\rangle_2)\right)\otimes\left([u_{I_2},S_2^{i_2j_2}]u_{J_2}\right)$$

$$=\sum_{I_2:J_2}\sum_{I_1}\beta_{I_1}\langle b,h_{I_1^{(k)}}\otimes u_{I_2}\rangle\langle S_1^{i_1j_1}(f),h_{I_1}^{\epsilon_1}\otimes u_{J_2}\rangle h_{I_1}^{\epsilon_1'}\left|I_1^k\right|^{-\frac{1}{2}}\otimes\left([u_{I_2},S_2^{i_2j_2}]u_{J_2}\right)$$

$$=\sum_{I_1}\beta_{I_1}h_{I_1}^{\epsilon_1'}\left|I_1^k\right|^{-\frac{1}{2}}\otimes\left(\left[\langle b,h_{I_1^{(k)}}\rangle_1, S_1^{i_1j_1}\right]\langle S_1^{i_1j_1}f,h_{I_1}^{\epsilon_1}\rangle_1\right)$$

$$=\sum_{I_1}\beta_{I_1}h_{I_1}^{\epsilon_1'}\left|I_1^k\right|^{-\frac{1}{2}}\otimes\left(\sum_{s_1\in\Gamma_1}B^2_{l,s_1}\left(\langle b,h_{I_1^{(k)}}\rangle_1, S_2^{i_2j_2}\left(\langle S_1^{i_1j_1}(f),h_{I_1}^{\epsilon_1}\rangle_1\right)\right)\right.$$

$$\left.+\sum_{s_2\in\Gamma_2}S_2^{i_2j_2}\left(B^2_{l,s_2}\left(\langle b,h_{I_1^{(k)}}\rangle_1,\langle S_1^{i_1j_1}(f),h_{I_1}^{\epsilon_1}\rangle_1\right)\right)\right)$$

where B^2_{l,s_i} are paraproducts of type B^2_l in the second variable, and all the index sets Γ_i satisfy $|\Gamma_i| \le C(1+\max(i_2,j_2)), i = 1,2$. Again, since all the terms in the parentheses are similar, we only consider one of B^2_{l,s_2} and omit the subscript s_2. This is a mixed case, and all the other combinations follow similarly. Thus, noticing that

$$\sum_{I_1}\beta_{I_1}h_{I_1}^{\epsilon_1'}\left|I_1^k\right|^{-\frac{1}{2}}\otimes S_2^{i_2j_2}\left(B^2_l\left(\langle b,h_{I_1^{(k)}}\rangle_1,\langle S_1^{i_1j_1}(f),h_{I_1}^{\epsilon_1}\rangle_1\right)\right)$$

$$= S_2^{i_2j_2}\left(\sum_{I_1,I_2}\beta_{I_1}\beta_{I_2}\langle b,h_{I_1^{(k)}}\otimes u_{I_2^{(l)}}\rangle\langle S_1^{i_1j_1}f,h_{I_1}^{\epsilon_1}\otimes u_{I_2}^{\epsilon_2}\rangle h_{I_1}^{\epsilon_1'}\otimes u_{I_2}^{\epsilon_2'}\left|I_1^{(k)}\right|^{-\frac{1}{2}}\left|I_2^{(l)}\right|^{-\frac{1}{2}}\right)\quad(12)$$

is exactly $S_2^{i_2j_2}\left(B_{k,l}\left(bS_1^{i_1j_1}f\right)\right)$, where $B_{k,l}$ is the bi-parameter para- product we've studied in Lemma(4.1.8),and the only case involving non-cancellative Haar functions is when the corresponding k or l is 0. We therefore obtain the desired representation of this term. All the other terms can be treated similarly, by noticing that paraproducts $B_{k,l}$ can be obtained by combining B^1_k and B^2_l through the same process described above. And it is easily seen that the total number of terms is bounded by $(1+\max(i_1,j_1))(1+\max(i_2,j_2))$ up to a dimensional constant.

We assume that $S_2^{00}f = \sum_{I_2}\langle a^2,u_{I_2}\rangle_2|I_2|^{-1/2}\langle f,u^1_{I_2}\rangle_2 u_{I_2}$. Following from Theorem(4.1.5), in the first variable, the commutator can be represented as a linear combination of paraproducts, i.e.

$$\left[[b,S_1^{i_1j_1}],S_2^{00}\right]f$$

$$= \sum_{I_1 \subset J_1^{(i_1)}} \sum_{I_2 \subset J_2} \langle b, h_{I_1} \otimes u_{I_2} \rangle \langle f, h_{J_1} \otimes u_{J_2} \rangle \left[h_{I_1}, S_1^{i_1 j_1} \right] h_{J_1} \otimes \left[u_{I_2}, S_2^{00} \right] u_{J_2}$$

$$= \sum_{I_1 \subset J_1} \left(\sum_{t_1 \in \Lambda_1} B_{k,t_1}^1 \left(\langle b, u_{I_2} \rangle_2, S_1^{i_1 j_1} (\langle f, u_{J_2} \rangle_2) \right) \right.$$

$$\left. + \sum_{t_2 \in \Lambda_2} S_1^{i_1 j_1} \left(B_{k,t_2}^1 (\langle b, u_{I_2} \rangle_2, \langle f, u_{J_2} \rangle_2) \right) \right) \otimes \left(\left[u_{I_2}, S_2^{00} \right] u_{J_2} \right).$$

Recall that by Theorem(4.1.5), in the one-parameter setting, the noncan- cellative dyadic shiftS^{00}can be represented as a finite linear combination of paraproducts (corresponding to the sum over$I \subsetneq J$and the second term in the sum over$I = J$) and operator P (corresponding to the first term in the sum over$I = J$). Hence,

$$\sum_{I_2 \subset J_2} B_{k,t_1}^1 \left(\langle b, u_{I_2} \rangle_2, S_1^{i_1 j_1} (\langle f, u_{J_2} \rangle_2) \right) \otimes \left(\left[u_{I_2}, S_2^{00} \right] u_{J_2} \right)$$

$$= \sum_{I_1} \beta_{I_1} h_{I_1}^{\epsilon_1'} \left| I_1^{(k)} \right|^{-\frac{1}{2}} \otimes \left(\left[\langle b, h_{I_1^{(k)}} \rangle_1, S_2^{00} \right], \langle S_1^{i_1 j_1} f, h_{I_1}^{\epsilon_1} \rangle_1 \right)$$

$$= \sum_{I_1} \beta_{I_1} h_{I_1}^{\epsilon_1'} \left| I_1^{(k)} \right|^{-\frac{1}{2}} \otimes \left(\sum_{s_1 \in \Gamma_1} B_{l,s_1}^2 \left(\langle b, h_{I_1^{(k)}} \rangle_1, S_2^{i_2 j_2} (\langle S_1^{i_1 j_1}(f), h_{I_1}^{\epsilon_1} \rangle_1) \right) \right.$$

$$+ \sum_{s_2 \in \Gamma_2} S_2^{i_2 j_2} \left(B_{l,s_2}^2 \left(\langle b, h_{I_1^{(k)}} \rangle_1, \langle S_1^{i_1 j_1}(f), h_{I_1}^{\epsilon_1} \rangle_1 \right) \right)$$

$$\left. + P \left(\langle b, h_{I_1^{(k)}} \rangle_1, a^2, \langle S_1^{i_1 j_1} f, h_{I_1}^{\epsilon_1} \rangle_1 \right) \right)$$

$$= \left(\sum_{s_1 \in \Gamma_1} B_{k,0,,s_1} (b, S_1^{i_1 j_1} S_2^{00} f) \right) + \left(\sum_{s_2 \in \Gamma_2} S_2^{00} \left(B_{k,0,,s_2} (b, S_1^{i_1 j_1} f) \right) \right)$$

$$+ BP_k (b, a^2, S_1^{i_1 j_1} f).$$

Similarly, the other term can be treated exactly the same:

$$\sum_{I_2 \subset J_2} S_1^{i_1 j_1} \left(B_{k,t_2}^1 (\langle b, u_{I_2} \rangle_2, \langle f, u_{J_2} \rangle_2) \right) \otimes \left[u_{I_2}, S_2^{00} \right] u_{J_2}$$

$$= \left(\sum_{s_1 \in \Gamma_1} S_1^{i_1 j_1} \left(B_{k,0,,s_1} (b, S_2^{00} f) \right) \right) + \left(\sum_{s_2 \in \Gamma_2} S_1^{i_1 j_1} S_2^{00} \left(B_{k,0,,s_2} (b, f) \right) \right)$$

$$+ S_1^{i_1 j_1} \left(BP_k (b, a^2, f) \right).$$

The desired representation is hence obtained. Note that by symmetry and duality, this implies the boundedness of other types of the mixed cases as well.

$$\left[[b, S_1^{00}], S_2^{00} \right] f$$

$$= \sum_{I_1 \subset J_1} \sum_{I_2 \subset J_2} \langle b, h_{I_1} \rangle_1 [h_{I_1}, S_1^{00}] u_{J_1} \otimes [h_{I_2}, S_2^{00}] u_{J_2}.$$

First, we deal with the case when both S_1^{00} and S_2^{00} are of the same type, for instance,

$$S_1^{00} f := \sum_{I_1} \langle a^1, h_{I_1} \rangle_1 \, |I_1|^{-\frac{1}{2}} \langle f, h_{I_1}^1 \rangle h_{I_1},$$

$$S_2^{00} f := \sum_{I_2} \langle a^2, u_{I_2} \rangle_2 \, |I_2|^{-\frac{1}{2}} \langle f, u_{I_2}^1 \rangle_2 u_{I_2}.$$

Observe that compared, after decomposing the commutator in each variable into paraproducts and operator P, the only new case that arises here is the "tensor product" of operator P in both variables, which is equalto

$$\sum_{I_1, I_2} \langle b, h_{I_1} \otimes u_{I_2} \rangle \langle f, h_{I_1} \otimes u_{I_2} \rangle \, |I_1|^{-1} |I_2|^{-1}$$

$$\times \sum_{J_1 : J_1 \subsetneq I_1} \sum_{J_2 : J_2 \subsetneq I_2} \langle a^1 \otimes a^2, h_{J_1} \otimes u_{J_2} \rangle h_{J_1} \otimes u_{J_2}$$

$$= PP(b, a^1 \otimes a^2, f).$$

Second, we discuss the case when S_1^{00} and S_2^{00} are of different types, for instance,

$$S_1^{00} f := \sum_{I_1} \langle a^1, h_{I_1} \rangle_1 \, |I_1|^{-\frac{1}{2}} \langle f, h_{I_1} \rangle h_{I_1}^1,$$

$$S_2^{00} f := \sum_{I_2} \langle a^2, u_{I_2} \rangle_2 \, |I_2|^{-\frac{1}{2}} \langle f, u_{I_2}^1 \rangle_2 u_{I_2}.$$

It is implied by Theorem(4.1.5) that in the first variable, the commutator is a linear combination of paraproducts and operator P^*. Therefore, the only new case that arises here in the representation is P^* in the first variable mixed with P in the second variable, which is

$$\sum_{I_1, I_2} \langle b, h_{I_1} \otimes u_{I_2} \rangle \, |I_1|^{-1} |I_2|^{-1} \times \sum_{J_1 : J_1 \subsetneq I_1} \sum_{J_2 : J_2 \subsetneq I_2} \langle a^1 \otimes a^2, h_{J_1} \otimes u_{J_2} \rangle \langle f, h_{I_1} \otimes u_{I_2} \rangle h_{J_1} \otimes u_{J_2}$$

$$= PP_1(b, a^1 \otimes a^2, f).$$

Hence the main theorem in the bi-parameter setting is proved. As a final remark, the proof in the multi-parameter setting proceeds exactly the same as this one. Clearly, in the desired representation of commutators with dyadic shifts, one needs to involve a larger number of basic operators which mix together B_k and P in each variable, but the uniform boundedness of such operators can all be obtained similarly as in Lemmas (4.1.8), (4.1.9),and(4.1.10).

Section (4.2): Characterizations of Multi-Parameter BMO

As dual of the Hardy space H^1, the classical space of functions of bounded mean oscillation, BMO, arises naturally in many endpoint results in analysis, partial differential equations and probability. When entering a setting with several free parameters, a large variety of spaces are encountered, some of which lose the feature of mean oscillation itself. We are interested in characterizations of multi-parameter BMO spaces through boundedness of commutators.

A classical result of Nehari [9] shows that a Hankel operator with anti-analytic symbol b mapping analytic functions into the space of anti-analytic functions by $f \longmapsto P\_bf$ is bounded with respect to the L^2 norm if and only if the symbol belongs to BMO. This

theorem has an equivalent formulation in terms of the boundedness of the commutator of the multiplication operator with symbol function b and the Hilbert transform $[H,b] = Hb - Hb$.
Ferguson–Sadosky in [5] and later Ferguson–Lacey in their ground breaking [1] study the symbols of bounded 'big' and 'little' Hankel operators on the bidisk through commutators of the tensor product or of the iterated form

$$[H_1, H_2, b], \text{and } [H_1[H_2, b]].$$

Here $b = b(x_1, x_2)$ and the H_k are the Hilbert transforms acting in the kth variable. A full characterization of different two-parameter BMO spaces, Cotlar–Sadosky's little BMO and Chang–Fefferman's product BMO space, is given through these commutators.

Through the use of completely different real variable methods, in [3] Coifman–Rochberg–Weiss extended Nehari's one-parameter theory to real analysis in the sense that the Hilbert transform was replaced by Riesz transforms. These one-parameter resultsi [3] were treated in the multi-parameter setting in Lacey–Petermichl–Pipher–Wick [111]. Both the upper and lower estimate have proofs very different from those in one parameter. In addition, in both cases it is observed that the Riesz transforms are a representative testing class in the sense that BMO also ensures boundedness for (iterated) commutators with more general Calderón–Zygmund operators, a result now known in full generality due to Dalenc–Ou [106]. Notably the Riesz commutator has found striking applications to compensated compactness and div-curl lemmas,[118], [122].
Our extension to the multi-parameter setting is two-fold. On the one hand we replace the Calderón–Zygmund operators by Journé operators J_i and on the other hand we also iterate the commutator:

$$[J_1, \dots, [J_t, b] \dots].$$

We prove the remarkable fact that a multi-parameter BMO class still ensures boundedness in this situation and that the collection of tensor products of Riesz transforms remains the representative testing class. The BMO class encountered is a mix of little BMO and product BMO that we call a little product BMO. Its precise form depends upon the distribution of variables in the commutator. In this case, lower estimates were only known in the case of the double Hilbert transform[5]. The sufficiency of the little BMO class for boundedness of Journé commutators had never been observed.
It is a general fact that two-sided commutator estimates have an equivalent formulation in terms of weak factorization. We find the preduals of our little product BMO spaces and prove a corresponding weak factorization result. Necessity of the little product BMO condition is shown through a lower estimate on the commutator. There is a sharp contrast when tensor products of Riesz transforms are considered instead of multiple Hilbert transforms and when iterations are present.
In the Hilbert transform case, Toeplitz operators with operator symbol arise naturally. Using Riesz transforms in $\mathbb{R}^d$ as a replacement, there is an absence of analytic structure and tools relying on analytic projection or orthogonal spaces are not readily available. We overcome part of this difficulty through the use of Calderón–Zygmund operators whose Fourier multiplier symbols are adapted to cones. This idea is inspired by[111]. Such operators are also mentioned in [128]. A class of operators of this type classifies little product BMO through two-sided commutator estimates, but it does not allow the passage to a classification through iterated commutators with tensor products of Riesz transforms.

In a second step, we find it necessary to consider upper and lower commutator estimates using a well-chosen family of Journé operators that are not of tensor product type. Through geometric considerations and an averaging procedure of zonal harmonics on products of spheres, we construct the multiplier of a special Journé operator that preserves lower commutator estimates and resembles the multiple Hilbert transform: it has large plateaus of constant values and is a polynomial in multiple Riesz transforms. We expect that this construction allows other applications.
There is an increase in difficulty when the dimension is greater than two, due to the simpler structure of the rotation group on $\mathbb{S}^1$. In higher dimension, there is a rise in difficulty when tensor products involve more than two Riesz transforms. The actual passage to the Riesz transforms requires a stability estimate in commutator norms for certain multi-parameter singular integrals in terms of the mixed BMO class. We prove a qualitative upper estimate for iterated commutators using paraproduct free Journé operators. We make use of recent versions of $T(1)$ theorems in this setting. These recent advances are different from the corresponding theorem of Journé[16]. The results we allude to have the additional feature of providing a convenient representation formula for bi-parameter in[123] and even multi-parameter in[125] Calderón–Zygmund operators by dyadic shifts.

This contains some review on Hardy spaces in several parameters as well as some new definitions and lemmas relevant to us.

We describe the elements of product Hardy space theory, as developed by Chang and Fefferman as well as Journé. By this we mean the Hardy spaces associated with domains like the poly-disk or $\mathbb{R}^{d} \coloneqq \otimes_{s=1}^{t} \mathbb{R}^{d_s}$ for $d = (d_1, \ldots, d_t)$. While doing so, we typically do not distinguish whether we are working on $\mathbb{R}^d$ or $\mathbb{T}^d$. In higher dimensions, the Hilbert transform is usually replaced by the collection of Riesz transforms.
The (real) one-parameter Hardy space $H^1_{Re}(\mathbb{R}^d)$ denotes the class of functions with the norm

$$\sum_{j=0}^{d} \|R_j f\|_1$$

where R_j denotes the jth Riesz transform or the Hilbert transform if the dimension is one. Here and below we adopt the convention that R_0, the 0th Riesz transform, is the identity. This space is invariant under the one-parameter family of isotropic dilations, while the product Hardy space $H^1_{Re}(\mathbb{R}^d)$ is invariant under dilations of each coordinate separately. That is, it is invariant under a t parameter family of dilations, hence the terminology 'multi-parameter' theory. One way to define a norm on $H^1_{Re}(\mathbb{R}^d)$ is

$$\|f\|_{H^1} \sim \sum_{0 \le j_l \le d_l} \left\| \bigotimes_{l=1}^{t} R_{l,j_l} f \right\|_1 .$$

R_{l,j_l} is the Riesz transform in the j_lth direction of the lth variable, and the 0th Riesz transform is the identity operator.
The dual of the real Hardy space $H^1_{Re}(\mathbb{R}^d)^*$ is $BMO(\mathbb{R}^d)$, the t-fold product BMO space. It is a theorem of S.-Y. Chang and R. Fefferman [2], [107] that this space has a characterization in terms of a product Carleson measure.
Define

$$\|b\|_{BMO(\mathbb{R}^d)} := \sup_{U\subset\mathbb{R}^d}\left(|U|^{-1}\sum_{R\subset U}\sum_{\varepsilon\in\mathrm{sig}_d}|\langle b, w_R^{\varepsilon}\rangle|^2\right)^{\frac{1}{2}}. \qquad (13)$$

Here the supremum is taken over all open subsets $U \subset \mathbb{R}^d$with finite measure, and we use a wavelet basis w_R^{ϵ} adapted to rectangles $R = Q_1 \times \ldots \times Q_t$, where each Q_l is a cube. The superscript ϵ reflects the fact that multiple wavelets are associated to any dyadic cube, see[111] for details. The fact that the supremum admits all open sets of finite measure cannot be omitted, as Carleson's example shows[117]. This fact is responsible for some of the difficulties encountered when working with this space.

Theorem(4.2.1)[115]: (Chang, Fefferman). We have the equivalence of norms

$$\|b\|_{\left(H^1_{Re}(\mathbb{R}^d)\right)^*}\|b\|_{BMO(\mathbb{R}^d)}.$$

That is, $BMO(\mathbb{R}^d)$is the dual to$H^1_{Re}(\mathbb{R}^d)$.

This BMOnorm is invariant under a t-parameter family of dilations. Here the dilations are isotropic in each parameter separately. See also[4] and[121].

Following[119] and[5], we recall some facts about the space littleBMO, often written as 'bmo', and its predual. A locally integrable function$b: \mathbb{R}^d = \mathbb{R}^{d_1} \times \ldots \times \mathbb{R}^{d_s} \to \mathbb{C}$is in bmo if and only if

$$\|b\|_{bmo} = \sup_{Q=Q_1\times\ldots\times Q_s}|Q|^{-1}\int_Q \left|b(x) - b_Q\right| < \infty$$

Here the Q_k are d_k-dimensional cubes and$b_Q$denotes the average of b over Q.

It is easy to see that this space consists of all functions that are uniformly in BMO in each variable separately. Let$x_{\hat{v}} = (x_1, \ldots, x_{v-1},; x_{v+1}, \ldots, x_s)$. Then$b(x_{\hat{v}})$is a function in$x_v$only with the other variables fixed. Its BMO norm in$x_v$is

$$\|b(x_{\hat{v}})\|_{BMO} = \sup_{Q_v}|Q_v|^{-1}\int_{Q_v}|b(x) - b(x_{\hat{v}})Q_v|\,dx_v$$

and the little BMO norm becomes

$$\|b\|_{bmo} = \max_v\left\{\sup_{x_{\hat{v}}}\|b(x_{\hat{v}})\|_{BMO}\right\}.$$

On the bi-disk, this becomes

$$\|b\|_{bmo} = \max_v\left\{\sup_{x_1}\|b(x_1,.)\|_{BMO}, \sup_{x_2}\|b(x_2,.)\|_{BMO}\right\},$$

the space discussed in[5]. Here, the pre-dual is the space$H^1(\mathbb{T})\otimes L^1(\mathbb{T}) + L^1(\mathbb{T})\otimes H^1(\mathbb{T})$. All other cases are an obvious generalization, at the cost of notational inconvenience.

We define a BMO space which is in between little BMO and productBMO. As mentioned, we aim at characterizing BMO spaces consisting for example of those functions $b(x_1, x_2, x_3)$qsuch that $b(x_1,.,.)$ and $b(.,., x_3)$qare uniformly in product BMO in the remaining two variables.

Definition(4.2.2)[115]: Let $b: \mathbb{R}^d \to \mathbb{C}$ with $d = (d_1, \ldots, d_t)$. Take a partition $\mathcal{I} = \{I_s: I \le s \le l\}$ of $\{1,2, \ldots, t\}$so that $\dot\cup_{1\le s\le l} I_s = \{1,2, \ldots, t\}$. We say that $b \in BMO_{\mathcal{I}}(\mathbb{R}^d)$ if for any choices $v = (v_s), v_s \in I_s, b$is uniformly in product BMO in the variables indexed byv_s. We call a BMOspace of this type a 'little product BMO'. If for any$x = (x_1, \ldots, x_t) \in$

$\mathbb{R}^d$, we define$x_{\hat{v}}$by removing those variables indexed by v_s, the little product BMO norm becomes

$$\|b\|_{BMO_{\mathcal{J}}} = \max_{v}\left\{\sup_{x_{\hat{v}}}\|b(x_{\hat{v}})\|_{BMO}\right\}$$

where the BMO norm is product BMO in the variables indexed byv_s.

For example, when$d = (1,1,1) = 1$, when$t = 3$ and $l = 2$ with$I_1 = (13)$and$I_2 = (2)$, writing $\mathcal{J} = (13)(2)$ the space$BMO_{(13)(2)}(\mathbb{T}^1)$arises, which consists of those functions that are uniformly in product BMO in the variables$(1,2)$and$(3,2)$ respectively, as described above. Moreover, as degenerate cases, it is easy to see that$BMO_{(12\ldots t)}$and$BMO_{(1)(2)\ldots(t)}$are exactly little BMO and product BMO respectively, the spaces we are familiar with.

Little product BMO spaces on$\mathbb{T}^d$can be defined in the same way. Now we find the predual of$BMO_{(13)(2)}$, which is a good model for other cases. We choose the order of variables most convenient for us.

Theorem (4.2.3)[115]: The pre-dual of the space$BMO_{(13)(2)}(\mathbb{T}^1)$is equal to the space

$$H^1_{Re}\left(\mathbb{T}^{(1.1)}\right)\otimes L^1(\mathbb{T}) + L^1(\mathbb{T})\otimes H^1_{Re}\left(\mathbb{T}^{(1.1)}\right)$$
$$:= \left\{f + g : f \in H^1_{Re}\left(\mathbb{T}^{(1.1)}\right)\otimes L^1(\mathbb{T}) and g \in L^1(\mathbb{T})\otimes H^1_{Re}\left(\mathbb{T}^{(1.1)}\right)\right\}.$$

Proof: The space

$$H^1_{Re}\left(\mathbb{T}^{(1.1)}\right)\otimes L^1(\mathbb{T}) = \{f \in L^1(\mathbb{T}^3):, H_2 f H_1 H_2 f, L^1(\mathbb{T}^3)\}$$

equipped with the norm$\|f\| = \|f\|_1 + \|H_1 f\|_1 + \|H_2 f\|_1 + \|H_1 H_2 f\|_1$is a Banach space. Let$W^1 = L^1(\mathbb{T}^3) \times L^1(\mathbb{T}^3) \times L^1(\mathbb{T}^3) \times L^1(\mathbb{T}^3)$equipped with the norm

$$\|(f_1, f_2, f_3, f_4)\|_{W_1} = \|f_1\|_1 + \|f_2\|_1 + \|f_3\|_1 + \|f_4\|_1.$$

Then we see that$H^1_{Re}\left(\mathbb{T}^{(1.1)}\right)\otimes L^1(\mathbb{T})$is isomorphically isometric to the closed subspace

$$V = \left\{\left(f, H_1(f), H_2(f), H_1 H_2(f)\right) : f \in H^1\left(\mathbb{T}^{(1.1)}\right)\otimes L^1(\mathbb{T}^3)\right\}$$

ofW^1. Now, the dual of$W^1$is equal to$W^\infty = L^\infty(\mathbb{T}^3) \times L^\infty(\mathbb{T}^3) \times L^\infty(\mathbb{T}^3) \times L^\infty(\mathbb{T}^3)$equipped with the norm$\|(g_1, g_2, g_3, g_4)\|_\infty = \max\{\|g_i\|_\infty : 1 \le i \le 4\}$so the dual space of V is equal to the quotient of$W^\infty$by the annihilator U of the subspace V inW^∞. But, using the fact that the Hilbert transforms are self-adjoint up to a sign change, we see that

$$U = \{(g_1, g_2, g_3, g_4) : g_1 + H_1 g_2 + H_2 g_3 + H_1 H_2 g_4 = 0\}$$

and so:

$$V^* \cong W^\infty / U \cong \operatorname{lm} \theta$$

where

$$\theta(g_1, g_2, g_3, g_4) : g_1 + H_1 g_2 + H_2 g_3 + H_1 H_2 g_4 = 0$$

since$U = ker(\theta)$. But

$$\operatorname{lm}(\theta) = L^\infty(\mathbb{T}^3) + H_1\left(L^\infty(\mathbb{T}^3)\right) + H_2\left(L^\infty(\mathbb{T}^3)\right) + H_1\left(H_2\left(L^\infty(\mathbb{T}^3)\right)\right)$$

is equal to the functions that are uniformly in product BMO in variables1 and2. Using the same reasoning we see that the dual of $L^1(\mathbb{T})\otimes H^1_{Re}\left(\mathbb{T}^{(1.1)}\right)$is equal to $L^\infty(\mathbb{T}^3) + H_2\left(L^\infty(\mathbb{T}^3)\right) + H_3\left(L^\infty(\mathbb{T}^3)\right) + H_2 H_3\left(L^\infty(\mathbb{T}^3)\right)$, which is equal to the space of functions that are uniformly in product BMO in variables 2 and 3. Now, we consider the 'L^1 sum' of the spaces$H^1_{Re}\left(\mathbb{T}^{(1.1)}\right)\otimes L^1(\mathbb{T})andL^1(\mathbb{T})\otimes H^1_{Re}\left(\mathbb{T}^{(1.1)}\right)$; that is

$$M_{(13)(2)} = \left\{(f, g) : f \in H^1_{Re}\left(\mathbb{T}^{(1.1)}\right)\otimes L^1(\mathbb{T}); g \in L^1(\mathbb{T})\otimes H^1_{Re}\left(\mathbb{T}^{(1.1)}\right)\right\}$$

equipped with the norm

$$\|(f,g)\| = \|f\|_{H^1_{Re}(\mathbb{T}^{(1.1)})\otimes L^1(\mathbb{T})} + \|g\|_{L^1(\mathbb{T})\otimes H^1_{Re}(\mathbb{T}^{(1.1)})}.$$

We see that, if$\phi: M_{(13)(2)} \to L^1(\mathbb{T}^3)$is defined by$\phi(f,g) = f + g$, then the image of$\phi$is isometrically isomorphic to the quotient of$M_{(13)(2)}$by the space

$$N = \{(f,g) \in M_{(13)(2)}: f + g = 0\}$$
$$= \{(f,-f): f \in H^1_{Re}(\mathbb{T}^{(1.1)})\otimes L^1(\mathbb{T}) \cap L^1(\mathbb{T})\otimes H^1_{Re}(\mathbb{T}^{(1.1)})\}.$$

Now, recall that the dual of the quotient M/N is equal to the annihilator of$N$. It is easy to see that the annihilator of N is equal to the set of ordered pairs(ϕ,ϕ) qwithϕin the intersection of the duals of the two spaces. Thus the dual of the image of$\theta$is equal to$BMO_{(13)(2)}(\mathbb{R}^d)$. The norm of an element in the predual is equal to its norm as an element of the double dual which is easily computed. Following this example, the reader may easily find the correct formulation for the predual of other little product BMO spaces as well those in several variables, replacing the Hilbert transform by all choices of Riesz transforms. For instance, one can prove that the predual of the space$BMO_{(13)(2)}(\mathbb{R}^d)$is equal to $H^1_{Re}(\mathbb{R}^{(d_1,d_2)})\otimes L^1(\mathbb{R}^{d_3}) + L^1(\mathbb{R}^{d_1})\otimes H^1_{Re}(\mathbb{R}^{(d_2,d_3)})$.

We characterize the boundedness of commutators of the form$[H_2[H_3H_1,b]]$, bss as operators on$L^2(\mathbb{T}^3)$. In the case of the Hilbert transform, this case is representative of the general case and provides a starting point that is easier to read because of the simplicity of the expression of products and sums of projection onto orthogonal subspaces. Its general form can be found at the beginning.

Now let $b \in L^1(\mathbb{T}^n)$and let P and Q denote orthogonal projections onto subspaces of $L^2(\mathbb{T}^n)$. We shall describe relationships between functions in the little product BMOs and several types of projection-multiplication operators. These will be Hilbert transform-type operators of the form $P - P^\perp$; and iterated Hankel or Toeplitz type operators of the form $Q^\perp bQ$(Hankel), PbP(Toeplitz),$PQ^\perp bQP$ (mixed), where b means the (not a priori bounded) multiplication operator $M_b$on $L^2(\mathbb{T}^n)$.

We shall use the following simple observation concerning Hilbert transform type operators again and again:

Corollary(4.2.4)[115]: We have the following two-sided estimate

$$\|b\|_{BMO_{(13)(2)}} \lesssim \|[H_2[H_3H_1,b]]\|_{L^2(\mathbb{T}^3)\to L^2(\mathbb{T}^3)} \lesssim \|b\|_{BMO_{(13)(2)}}.$$

It will be useful to denote by Q_{13} orthogonal projection on the subspace of functions which are either analytic or anti-analytic in the first and third variables; $Q_{13} = P_1P_3 + P_1^\perp P_3^\perp$. Then the projection $Q_{13}^\perp$ onto the orthogonal of this subspace is defined by $Q_{13}^\perp = P_1^\perp P_3 + P_1P_3^\perp$. We reformulate properties (ii) and (iii) in the statement of Theorem (4.2.3) in terms of Hankel Toeplitz type operators.

Lemma (4.2.5)[115]: We have the following algebraic facts on commutators and projection operators.

(i) The commutators$[H_2[H_1,b]]$and$[H_2[H_3,b]]$are bounded on$L^2(\mathbb{T}^3)$if and only if the operators$P_iP_2bP_i^\perp P_2^\perp, P_i^\perp P_2bP_iP_2^\perp bP_i^\perp P_2, P_i^\perp P_2^\perp bP_iP_2$with$i \in \{1,3\}$are bounded on $L^2(\mathbb{T}^3)$.

(ii) The commutator$[H_2[H_3H_1,b]]$is bounded on$L^2(\mathbb{T}^3)$if and only if all four operators $P_2Q_{13}bQ_{13}^\perp P_2^\perp, P_2^\perp Q_{13}^\perp bQ_{13}P_2, P_2Q_{13}^\perp bQ_{13}P_2^\perp, P_2^\perp Q_{13}bQ_{13}^\perp P_2$are bounded on $L^2(\mathbb{T}^3)$.

Proof: Using Remark1 it is easy to see that

$$[H_2[H_1,b]] = 4\left((P_2P_1bP_1^\perp P_2^\perp - P_2P_1^\perp bP_1P_2^\perp) - (P_2^\perp P_1bP_1^\perp P_2 - P_2^\perp P_1^\perp bP_1P_2)\right)$$

and that the corresponding equation for$[H_2[H_3,b]]$, bss is also true. This, along with the observation that the ranges of all arising summands are mutually orthogonal, gives

assertion (i). To prove (ii) we just notice that $H_1H_3 = Q_{13} - Q_{13}^{\perp}$ is a Hilbert transform type operator which permits us to repeat the above argument replacing P_1 by Q_{13}.
The following lemma will allow us to insert an additional Hilbert transform into the commutator without reducing the norm.

Lemma (4.2.6)[115]: $\|P_3P_1^{\perp}P_2^{\perp}bP_1P_2P_3\|_{L^2\to L^2} = \|P_1^{\perp}P_2^{\perp}bP_1P_2\|_{L^2\to L^2}$.

Proof: The inequality $\le$ is trivial, since P_3 is a projection which commutes with $P_1^{\perp}$ and $P_2^{\perp}$. To see $\ge$, notice that $P_3P_1^{\perp}P_2^{\perp}bP_1P_2P_3$ is a Toeplitz operator with symbol $P_1^{\perp}P_2^{\perp}bP_1P_2$. So $\|P_3P_1^{\perp}P_2^{\perp}bP_1P_2P_3\| = \sup_{x_3}\|P_1^{\perp}P_2^{\perp}b(.,.,x_3)P_1P_2\|$. The latter is just $\|P_1^{\perp}P_2^{\perp}bP_1P_2\|$. For convenience we include a sketch of the facts about Toeplitz operators we use. Compare [116]. Let W_3 be the operator of multiplication by $\mathcal{Z}_3, W_3(f) = \mathcal{Z}_3 f$, acting on $L^2(\mathbb{T}^3)$. If we define $B = P_1^{\perp}P_2^{\perp}bP_1P_2$ as well as

$$A_n = W_3^{*n}(P_3P_1^{\perp}P_2^{\perp}bP_1P_2P_3)W_3^n \text{ and } C_n = W_3^n(P_3^{\perp}P_1^{\perp}P_2^{\perp}bP_1P_2P_3^{\perp})W_3^{*n}$$

as operators acting on $L^2(\mathbb{T}^3)$ then the sequences A_n and C_n converge to B in the strong operator topology: it is easy to see that W_3, W_3^*; and P_3 commute with $P_1, P_2, P_1^{\perp}$ and $P_2^{\perp}$. The multiplier b satisfies the equation $W_3^{*n}bW_3^{*n} = b$ and $W_3^*W_3^{*n} = Id$. So we see that

$$A_n = P_1^{\perp}P_2^{\perp}(W_3^{*n}P_3W_3^n)bP_1P_2(W_3^{*n}P_3W_3^n).$$

But if $f \in L^2(\mathbb{T}^3)$, then, since W_3^n is a unitary operator:

$$\|W_3^{*n}P_3W_3^n(f) - f\| = \|P_3W_3^n(f) - W_3^n(f)\| = \|(P_3 - I)(W_3^n)(f)\| \to 0 \ (n \to \infty),$$

as tail of a convergent Fourier series. This means that $W_3^{*n}P_3W_3^n$ converges to the identity in the strong operator topology. Thus, for each $f \in L^2(\mathbb{T}^3)$ we have $\|(A_n - B)(f)\| \to 0$. So

$$\|P_1^{\perp}P_2^{\perp}bP_1P_2\| \le \sup_{n\in\mathbb{N}}\|W_3^{*n}(P_3P_1^{\perp}P_2^{\perp}bP_1P_2P_3)W_3^n\|$$
$$\le \|P_3P_1^{\perp}P_2^{\perp}bP_1P_2P_3\|.$$

We show the proof of the main theorem.

Theorem (4.2.7)[115]: Let $b \in L^1(\mathbb{T}^3)$. Then the following are equivalent with linear dependence on the respective norms.

(i) $b \in BMO_{(13)(2)}$.

(ii) The commutators $[H_2[H_1, b]]$ and $[H_2[H_3, b]]$ are bounded on $L^2(\mathbb{T}^3)$.

(iii) The commutator $[H_2[H_3H_1, b]]$ is bounded on $L^2(\mathbb{T}^3)$.

Proof: We show (i)⟺(ii) and (ii)⟺(iii).

(i)⟺(ii). Consider $f = f(x_1, x_2)$ and $g = g(x_3)$. Then

$$[H_2, [H_1, b]](fg) = g.[H_2, [H_1, b]](f).$$

So $\|[H_2, [H_1, b]](fg)\|_{L^2(\mathbb{T}^3)}^2 = \|Fg\|_{L^2(\mathbb{T})}^2$ where $F(x_2) = \|[H_2[H_1, b]](f)\|_{L^2(\mathbb{T}^2)}$. The map $g \mapsto Fg$ has $L^2(\mathbb{T})$ operator norm $\|F\|_{\infty}$. Now change the roles of x_1 and x_3. The Ferguson–Lacey equivalences $\|[H_2[H_i, b]]\| \sim \|b\|_{BMO}$ give the desired result.

(ii)⟹(iii). Boundedness of the commutators $[H_2, [H_1, b]]$ and $[H_2, [H_3, b]]$ implies the boundedness of the mixed commutator $[H_2, [H_1H_3, b]]$ by the identity $[H_2, [H_1H_3, b]] = H_1[H_2, [H_3, b]] + [H_2, [H_1, b]]H_3$.

(iii)⟹(ii). This part relies on Lemma (4.2.6). We wish to conclude from the boundedness of $[H_2, [H_3H_1, b]]$ the boundedness of $[H_2, [H_1, b]]$ and $[H_2, [H_3, b]]$. To see boundedness of $[H_2, [H_1, b]]$ let us look at one of the Hankels from Lemma (4.2.5). Lemma (4.2.6) shows that $P_2^{\perp}P_1^{\perp}bP_2P_1$ is bounded if and only if the operator $P_3P_1^{\perp}P_2^{\perp}bP_1P_2P_3$ is. And the latter is an operator found in the list from part (ii) of Lemma (4.2.5). The analogous reasoning shows that all eight Hankels in 1 are bounded and so (ii) is proved.

We are again in $\mathbb{R}^d$ with $d = (d_1, \dots, d_t)$ and a partition $\mathcal{I} = (I_s)_{1 \le s \le l}$ of$\{1, \dots, t\}$. We show the following characterization theorem of the space $BMO_{\mathcal{I}}(\mathbb{R}^{\boldsymbol{d}})$.

Such two-sided estimates also hold in L^p for $1 < p < \infty$. From the inductive nature of our arguments, it will also be apparent that the characterization holds when we consider intermediate cases, meaning commutators with any fixed number of Riesz transforms in each iterate. Below we state our most general two-sided estimate through Riesz transforms.

Theorem (4.2.8)[115]: Let$1 < p < \infty$. Under the same assumptions as Corollary$(4.2.9)$ and for any fixed $\boldsymbol{n} = (n_s)$ where$1 \le n_s \le |I_s|$, we have the two-sided estimate

$$\|b\|_{BMO_{\mathcal{I}}(\mathbb{R}^{\boldsymbol{d}})} \lesssim \sup_{j} \left\| \left[\boldsymbol{R}_{1,j^{(1)}}, \dots, \left[\boldsymbol{R}_{l,j^{(l)}}, b \right] \dots \right] \right\|_{L^p(\mathbb{R}^{\boldsymbol{d}})\circlearrowleft} \lesssim \|b\|_{BMO_{\mathcal{I}}(\mathbb{R}^{\boldsymbol{d}})}$$

where$\boldsymbol{j}^{(s)} = (j_k)_{k \in I_s}$, $0 \le j_k \le d_k$and for each$s$, there are ns non-zero choices. A Riesz transform in direction 0 is understood as the identity.

For $p = 2$ and $\boldsymbol{n} = \boldsymbol{1}$this is the equivalence (i) $\Longleftrightarrow$ (ii) and for$\boldsymbol{n} = (|I_1|, \dots, |I_1|)$it is the equivalence(i) $\Longleftrightarrow$ (iii) from Theorem(4.2.16).

Our main focus on a two-sided estimate when$\boldsymbol{n} = (|I_1|, \dots, |I_1|)$ n when the tensor product is a paraproduct-free Journé operator:

Corollary (4.2.9)[115]: Let $\boldsymbol{j} = (j_1, \dots, j_t)$with$1 \le j_k \le d_k$and let for each$1 \le s \le l$, $\boldsymbol{j}^{(s)} = (j_k)_{k \in I_s}$be associated a tensor product of Riesz transforms$\boldsymbol{R}_{s,\boldsymbol{j}^{(s)}} = \bigotimes_{k \in I_s} R_{k,j_k}$; here the $R_{k,j_k}$are j_kth Riesz transforms acting on functions defined on the kth variable. We have the two-sided estimate

$$\|b\|_{BMO_{\mathcal{I}}(\mathbb{R}^{\boldsymbol{d}})} \lesssim \sup_{j} \left\| \left[\boldsymbol{R}_{1,j^{(1)}}, \dots, \left[\boldsymbol{R}_{l,j^{(l)}}, b \right] \dots \right] \right\|_{L^p(\mathbb{R}^{\boldsymbol{d}})\circlearrowleft} \lesssim \|b\|_{BMO_{\mathcal{I}}(\mathbb{R}^{\boldsymbol{d}})}.$$

The statements above also serve as the statement of the general case for products of Hilbert transforms. In fact, when any $d_k = 1$just replace the Riesz transforms by the Hilbert transform in that variable. We consider the case $d_k \ge 2$ for $1 \le k \le t$and thus iterated commutators with tensor products of Riesz transforms only. The special case when $d_k = 1$ for some k is easier but requires extra care for notation.

The proof in the Hilbert transform case relied heavily on analytic projections and orthogonal spaces, a feature that we do not have when working with Riesz transforms. We are going to simulate the one-dimensional case by a two-step passage via intermediary Calderón–Zygmund operators whose multiplier symbols are adapted to cones.

In dimension$d_k \ge 2$, a cone$C \subset \mathbb{R}^d$with cubic base is given by the data(ξ, Q) where $\xi \in \mathbb{S}^{d-1}$is the direction of the cone and the cube$Q \subset \xi^{\perp}$centered at the origin is its aperture. The cone consists of all vectorsθthat take the form$\left(\theta_\xi \xi, \theta_\perp\right)$ where$\theta_\xi = \langle \theta, \xi \rangle$and$\theta_\perp \in \theta_\xi Q$. By λC we mean the dilated cone with data$(\xi, \lambda Q)$.

A cone D with ball base has data$(\xi, r)$for$0 < r < \pi/2$and$\xi \in \mathbb{S}^{d-1}$and consists of the vectors$\{\eta \in \mathbb{R}^d : d(\xi, \eta/\|\eta\|) \le r\}$where d is the geodesic distance (with distance of antipodal points being π).

Given any cone C or$D$, we consider its Fourier projection operator defined via$\widehat{P_C f} = \chi_C \hat{f}$. When the apertures are cubes, such operators are combinations of Fourier projections onto half spaces and as such admit uniformL^pbounds. Among others, this fact made cubic cones necessary in the considerations in[111] and[108] that we are going to need. For further technical reasons in the proof these operators are not quite good enough, mainly because they are not of Calderón–Zygmund type. For a given cone$C$, consider a Calderón–

Zygmund operator T_C with a kernel K_C whose Fourier symbol $\widehat{K_C} \in C^\infty$ and satisfies the estimate $\chi_C \le \widehat{K_C} \le \chi_{(1+\tau)C}$. This is accomplished by mollifying the symbol χ_C of the cone projection associated to cone C on $\mathbb{S}^{d-1}$ and then extending radially. We use the same definition for T_D.
Given a collection of cones $C = (C_k)$ we denote by T_C, P_C the corresponding tensor product operators.
In[111]it has been proved that Calderón–Zygmund operators adapted to certain cones of cubic aperture classify product BMO via commutators. As part of the argument, it was observed that test functions with opposing Fourier supports made the commutator large. In[108] a refinement was proven, that will be helpful to us. We prefer to work with cones with round base. Lower bounds for such commutators can be deduced from the assertion of the main theorem in[108], but we need to preserve the information on the Fourier support of the test function in order to succeed with our argument. Information on this test function is instrumental to our argument: it reduces the terms arising in the commutator to those resembling Hankel operators. We have the following lemma, very similar to that in[111], [108], the only difference being that the cones are based on balls instead of cubes.
Lemma (4.2.10)[115]: For every parameter $1 \le k \le t$ there exist a finite set of directions $\Upsilon_k \in \mathbb{S}^{d_k-1}$ and an aperture $0 < r_k < \pi/2$ so that, for every symbol b belonging to product BMO, there exist cones $D_k = D(\xi_k, r_k)$ with $\xi_k \in \Upsilon_k$ as well as a normalized test function $f = \otimes_{k=1}^t f_k$ whose components have Fourier support in the opposing cones

$$\left\|\left[T_{1,D_1} \dots, \left[T_{t,D_t}, b\right] \dots\right] f\right\|_2 \gtrsim \|b\|_{BMO_{(1)\dots(t)}\left(\mathbb{R}^d\right)}.$$

The stress is on the fact that the collection is finite, somewhat specific and serves all admissible product BMO functions.
Proof: The lemma in[108] supplies us with the sets of directions Υ_k as well as cones of cubic aperture Q_k and a test function f supported in the opposing cones. Now choose the aperture r_k large enough so that $(1+\tau)C(\xi_k, Q_k) \subset D(\xi_k, r_k)$. Then we have the commutator estimate

$$\left\|\left[T_{1,D_1} \dots, \left[T_{t,D_t}, b\right] \dots\right] f\right\|_2 \gtrsim \|b\|_{BMO_{(1)\dots(t)}\left(\mathbb{R}^d\right)}.$$

In fact, both commutators with cones C and D are L^2 bounded and reduce to $\|T_D(bf)\|_2$ or $\|T_C(bf)\|_2$ respectively thanks to the opposing Fourier support of f. Observe that $T_C(bf) = T_D(bf) = T_D\big(T_C(bf)\big)$. With $\|T_C\|_{2\to 2} \le 1$, we see that $\|T_D(bf)\|_2 \ge \|T_C(bf)\|_2$.
Using this a priori lower estimate, we are going to prove the lemma below.
Lemma (4.2.11)[115]: Let D_k for $1 \le k \le t$ denote any cones with respect to the kth variable. Let T_{D_k} denote the adapted Calderón–Zygmund operators. Let K be any proper subset of $\{k: 1 \le k \le t\}$, let $D_K = \otimes_{k\in K} D_k$ and T_{D_k} the associated tensor product of Calderón–Zygmund operators. Let $P^\sigma_{D_K}$ be a tensor product of projection operators on cones $D(\xi_k, r_k)$ or opposing cones $D(-\xi_k, r_k)$. Let $j \notin K$. Then

$$\left\|T_{D_k} T_{D_j} b P^\sigma_{D_K} P_{D_j}\right\|_{L^2(\mathbb{R}^d)\circlearrowleft} = \left\|T_{D_k} b P^\sigma_{D_K}\right\|_{L^2(\mathbb{R}^d)\circlearrowleft}.$$

Proof: We will establish this by composing some unilateral shift operators and studying their Fourier transform in the j variable. Let ξ_j denote the direction of the cone D_j, for any l define the shift operator

$$S_l g(x_j) = \int_{\mathbb{R}^{d_j}} \hat{g}(\eta_j) e^{2\pi i(l\xi_j+\eta_j)x_j}\, d\eta_j.$$

S_l is a translation operator on the Fourier side along the direction ξ_j of the cone D_j. It is not hard to observe that $S_l^* = S_{-l}$. Now define

$$A_l = S_{-l} T_{D_K} T_{D_j} b P^{\sigma}_{D_K} P_{D_j} S_l, \text{ and } B = T_{D_K} b P^{\sigma}_{D_K}.$$

We will prove that as $l \to +\infty$, $A_l \to B$ in the strong operator topology. As in the argument in Lemma (4.2.6), this together with the fact that S_l is an isometry will complete the proof. To see the convergence, let's first remember that S_l only acts on the j variable, and one always has the identities

$$S_l S_{-l} = Id \text{ and } S_{-l} b S_l = b.$$

This implies

$$A_l = T_{D_K}\left(S_{-l} T_{D_j} S_l\right)(S_{-l} b S_l) P^{\sigma}_{D_K}\left(S_{-l} P_{D_j} S_l\right)$$
$$= T_{D_K}\left(S_{-l} T_{D_j} S_l\right) b P^{\sigma}_{D_K}\left(S_{-l} P_{D_j} S_l\right).$$

We claim that both $S_{-l} T_{D_j} S_l$ and $S_{-l} P_{D_j} S_l$ converge to the identity operator in the strong operator topology, which then implies that $A_l \to B$ as $l \to \infty$. We will only prove $S_{-l} T_{D_j} S_l \to Id$ as the second limit is almost identical. Observe that $\left\| S_{-l} T_{D_j} S_l f - f \right\| = \left\| \left(T_{D_j} - I\right) S_l f \right\|$. Given any L^2 function f and any fixed large $l \geq 0$. Consider the f with frequencies supported in $\mathbb{R}^{d_1} \times \ldots \times (D_j - l\xi_j) \times \ldots \times \mathbb{R}^{d_t}$. In this case, $S_l f$ has Fourier support in $\mathbb{R}^{d_1} \times \ldots \times D_j \times \ldots \times \mathbb{R}^{d_t}$ where the symbol of T_{D_j} equals 1. Thus, for such f, we have $S_{-l} T_{D_j} S_l f = f$. The sets $\mathbb{R}^{d_1} \times \ldots \times (D_j - l\xi_j) \times \ldots \times \mathbb{R}^{d_t}$ exhaust the frequency space. With $\left\| T_{D_j} - I \right\|_{2\to 2} \leq 1$ the operators $S_{-l} T_{D_j} S_l$ converge to the Identity in the strong operator topology, and the lemma is proved. Observe that the aperture of the cone D_j is not relevant to the proof.

We proceed with the proof of the lower estimate for cone transforms.

Lemma (4.2.12)[115]: Let us suppose we are in $\mathbb{R}^d$ with $\boldsymbol{d} = (d_1, \ldots, t)$ and a partition $\mathcal{J} = (I_s)_{1 \leq s \leq l}$. For every $1 \leq k \leq t$ there exists a finite set of directions $\Upsilon_k \in \mathbb{S}^{d_k - 1}$ and an aperture rk so that the following hold for all $b \in BMO_{\mathcal{J}}(\mathbb{R}^{\boldsymbol{d}})$:

(i) For every $1 \leq s \leq l$ there exists a coordinate $v_s \in I_s$ and a direction $\xi_{v_s} \in \Upsilon_{v_s}$ and so that with the choice of cone $D_{v_s} = D(\xi_{v_s}, r_{v_s})$ and arbitrary D_k for coordinates $k \in I_s \backslash \{v_s\}$ and if D_s denotes their tensor product, then we have

$$\left\| \left[T_{1,D_1} \ldots, \left[T_{l,D_l}, b\right] \ldots\right] \right\|_{2\to 2} \gtrsim \|b\|_{BMO_{\mathcal{J}(\mathbb{R}^d)}}.$$

(ii) The test function $f = \otimes_{k=1}^{t} f_k$ which gives us a large L^2 norm in (i) has Fourier supports of the f_k contained in $D(-\xi_k, r_k)$ when $k = v_s$ and in D_k otherwise. Before we can begin with the proof of Lemma (4.2.12), we will need a real variable version of the facts on Toeplitz operators used earlier.

Proof. For a given symbol $b \in BMO_{\mathcal{J}}$, there exist for all $1 \leq s \leq l$ coordinates $\boldsymbol{v} = (v_s), v_s \in I_s$ and a choice of variables not indexed by v_s, $\boldsymbol{x}^0_{\hat{\boldsymbol{v}}}$ so that up to an arbitrarily small error

$$\|b\|_{BMO_{\mathcal{J}}} = \left\| b\left(\boldsymbol{x}^0_{\hat{\boldsymbol{v}}}\right) \right\|_{BMO_{(v_1)\ldots(v_l)}}.$$

By Lemma(4.2.10), there exist cones $D_{v_s} = D(\xi_{v_s}, r_{v_s})$ with directions $\xi_{v_s} \in \Upsilon_{v_s}$ and a normalized test function f_H in variables v_s with opposing Fourier support such that we have the lower estimate

$$\left\| \left[T_{v_1, D_{v_1}}, \dots, \left[T_{v_l, D_{v_l}}, b(\boldsymbol{x}_{\hat{\boldsymbol{v}}}^0) \right] \dots \right] (f_H) \right\|_{L^2(\mathbb{R}^{d_v})} \gtrsim \| b(\boldsymbol{x}_{\hat{\boldsymbol{v}}}^0) \|_{BMO_{(v_1)\dots(v_l)}}$$

where $\mathbb{R}^{d_v} = \mathbb{R}^{d_{v_1}} \times \mathbb{R}^{d_{v_l}}$.

We now consider the commutator with the same cones but with full symbol $b = b(.,\dots,.)$. Due to the lack of action on the variables not indexed by v_s, in the commutator, we have

$$\left[T_{v_1, D_{v_1}}, \dots, \left[T_{v_l, D_{v_l}}, b \right] \dots \right] (f_H g) = g. \left[T_{v_1, D_{v_1}}, \dots, \left[T_{v_l, D_{v_l}}, b \right] \dots \right] (f_H)$$

for g that only depends upon variables not indexed by v_s. Again using that multiplication operators in L^2 have norms equal to the L^∞ norm of their symbol, for the 'worst' L^2-normalized g we have

$$\begin{aligned} &\left\| \left[T_{v_1, D_{v_1}}, \dots, \left[T_{v_l, D_{v_l}}, b \right] \dots \right] (f_H g) \right\|_{L^2(\mathbb{R}^d)} \\ &= \sup_{\boldsymbol{x}_{\hat{\boldsymbol{v}}}} \left\| \left[T_{v_1, D_{v_1}}, \dots, \left[T_{v_l, D_{v_l}}, b(\boldsymbol{x}_{\hat{\boldsymbol{v}}}^0) \right] \dots \right] (f_H) \right\|_{L^2(\mathbb{R}^{d_v})} \\ &\geq \left\| \left[T_{v_1, D_{v_1}}, \dots, \left[T_{v_l, D_{v_l}}, b(\boldsymbol{x}_{\hat{\boldsymbol{v}}}^0) \right] \dots \right] (f_H) \right\|_{L^2(\mathbb{R}^{d_v})} \\ &\gtrsim \| b(\boldsymbol{x}_{\hat{\boldsymbol{v}}}^0) \|_{BMO_{(v_1)\dots(v_l)}(\mathbb{R}^{d_v})} = \| b \|_{BMO_{\mathcal{J}}(\mathbb{R}^d)}. \end{aligned}$$

Note that the test function g can be chosen with well distributed Fourier transform. Take any cones in the variables not indexed by v_s and let $\boldsymbol{D}$ denote the tensor product of their projections. $f_T = P_{\boldsymbol{D}} g$. Notice that

$$\left\| \left[T_{v_1, D_{v_1}}, \dots, \left[T_{v_l, D_{v_l}}, b \right] \dots \right] (f_H f_T) \right\| \gtrsim \left\| \left[T_{v_1, D_{v_1}}, \dots, \left[T_{v_l, D_{v_l}}, b \right] \dots \right] (f_H g) \right\|$$

with constants depending upon how small the aperture of the chosen cones is. Notice that the test function $f := f_H f_T$ has the Fourier support as required in part (2) of the statement of Lemma(4.2.12).

Now build cones D_s from the D_{v_s} and the other chosen cones D_k as well as operators $T_{s, \boldsymbol{D}_s}$. Notice that the commutators $\left[T_{v_1, D_{v_1}}, \dots, \left[T_{v_l, D_{v_l}}, b \right] \dots \right]$ and $\left[T_{1, \boldsymbol{D}_1}, \dots, \left[T_{l, \boldsymbol{D}_l}, b \right] \dots \right]$ reduce significantly when applied to a test function f with Fourier support like ours. When the operators $T_{v_s, D_{v_s}}$ or any tensor product T_{s, D_s} fall directly on f, the contribution is zero due to opposing Fourier supports of the test function and the symbols of the operators. The only terms left in the commutators $\left[T_{1, \boldsymbol{D}_1}, \dots, \left[T_{l, \boldsymbol{D}_{v_l}}, b \right] \dots \right] (f)$ and $\left[T_{v_1, D_{v_1}}, \dots, \left[T_{v_l, D_{v_l}}, b \right] \dots \right] (f)$ have the form $\bigotimes_s T_{s, \boldsymbol{D}_s}(bf)$ and $\bigotimes_s T_{v_s, D_{v_s}}(bf)$ respectively.

By repeated use of Lemma(4.2.11) we have the operator norm estimates for any symbol b, valid on the subspace of functions with Fourier support as described for f $\left\| \bigotimes_s T_{s, \boldsymbol{D}_s} b \right\|_{L^2 \to L^2} = \left\| \bigotimes_s T_{v_s, D_{v_s}} b \right\|_{L^2 \to L^2}$. We conclude that a normalized test function f with Fourier support as described in the statement (ii) of Lemma (4.2.12) exists, so that $\left\| \bigotimes_s T_{s, \boldsymbol{D}_s}(bf) \right\|_2 \gtrsim \| b \|_{BMO_{\mathcal{J}}(\mathbb{R}^d)}$. In particular, we get the desired estimate in(i).

We pass directly to a lower commutator estimate for tensor products of Riesz transforms from that for tensor products of cone operators. Just using tensor products of operators adapted to cones merely gives us some lower bound where we are unable to control that a Riesz transform does appear in every variable such as required in(iii) of Theorem(4.2.16). The reason for this will become clear as we advance in the argument. Instead of using operators $T_{s,\boldsymbol{D}_s}$directly, we will build upon them more general multi-parameter Journé type cone operators not of tensor product type that we now describe.

We explain the multiplier we need for i copies of$\mathbb{S}^{d-1}$when all dimensions are the same. We will explain how to pass to the case of i copies of varying dimension $d_k$below. A picture illustrating a base case, a product of two 1-spheres.

For $0 < b < a < 1$, let $\varphi: [-1,1] \to [-1,1]$be a smooth function with $\varphi(x) = 1$ when $a \leq x \leq 1$ and $\varphi(x) = 0$ when $b \geq x \geq 0$. And let φ be odd, meaning antisymmetric about $t = 0$. The function φ gives rise to a zonal function with pole $\xi_1$on the first copy of $\mathbb{S}^{d-1}$, denoted by$C_1(\xi_1; \eta_1)$. This is the multiplier of a one-parameter Calderón– Zygmund operator adapted to a cone $D(\xi_1, r)$ for $r = \pi/2(1-a)$. For $i > 1$ we define $C_k(\xi_1, \dots, \xi_k; \eta_1, \dots, \eta_k)$ for $1 < k \leq i$ inductively. In what follows, expectation is taken with respect to traces of surface measure. When $\eta_i = \pm\xi_i$, then conditional expectation is over a one-point set.

$$C_k(\xi_1, \dots, \xi_k; \eta_1, \dots, \eta_k)$$
$$= \mathbb{E}_{a_{k-1}}\big(C_{k-1}(\xi_1, \dots, a_{k-1}; \eta_1, \dots, \eta_{k-1}) | d(a_{k-1}, \xi_{k-1}) = d(\eta_k, \xi_k)\big).$$

If the dimensions are not equal take $d = \max_{d_j}$ and imbed $\mathbb{S}^{d_j-1}$ into $\mathbb{S}^{d-1}$by the map $\xi\left(\xi_1, \dots, \xi_{d_j}\right) \mapsto \left(\xi_1, \dots, \xi_{d_j}, 0, \dots, 0\right)$. Obtain in this manner the function C_i and then restrict to the original number of variables when the dimension is smaller than d.

The multiplier $\boldsymbol{J} = C_i(\boldsymbol{\xi}; .)$gives rise to a multi-parameter Calderón–Zygmund operator of convolution type (but not of tensor product type), $\boldsymbol{T}_J = \boldsymbol{T}_{C_i(\xi;.)}$. In fact, it is defined through principal value convolution against a kernel $\boldsymbol{K}_J = \boldsymbol{K}_{C_i(\xi;.)}(x_1, \dots, x_i)$such that

$$\forall l: \int_{\alpha < |x_l| < \beta} \boldsymbol{K}_J(x_1, \dots, x_i) dx_l = 0, \forall 0 < \alpha < \beta, x_j \in \mathbb{R}^{d_j} \text{fixed} \forall j \neq l$$
$$\left| \frac{\partial^{|n|}}{\partial x_1^{n_1} \dots \partial x_i^{n_i}} \boldsymbol{K}_J(x_1, \dots, x_i) \right| \leq A_n |x_1|^{-d_1-n_1} \dots |x_i|^{-d_i-n_i}, n_j \geq 0.$$

This kind of operator is a special case of the more general, non-convolution type discussed. It has many other nice features that will facilitate our passage to Riesz transforms. One of them is its very special representation in terms of homogeneous polynomials, the other one a lower commutator estimate in terms of the$BMO_{\mathcal{J}}$norm.

In order to proceed with the proof of these lemmas, we will use some well known facts about zonal harmonics. Fix a pol$\xi \in \mathbb{S}^{d-1}$. The zonal harmonic with poleξ of degree n is written as$Z_\xi^{(n)}(\eta)$. With$t = \langle \xi, \eta \rangle \in [-1,1]$, one writes$Z_\xi^{(n)}(\eta) = P_n(t)$where$P_n$is the Legendre polynomial of degree$n$. It is common to suppress the dependence on d in the notation for$Z_\xi^{(n)}$andP_n.

$Z_\xi^{(n)}$are reproducing for spherical harmonics of degree$n, Y^{(n)}$. When$Y^{(n)}$is harmonic and homogeneous of degree n with$Y^{(n)}(\xi) = 1$ and$Y^{(n)}(R\eta) = Y^{(n)}(\eta)$for any rotation$R \in \mathcal{O}(d)$ with $R\xi = \xi$, then$Y^{(n)} = Z_\xi^{(n)}$.

The lemma below will aid us in understanding the special form of the functionsC_i.

Lemma (4.2.13)[115]: Let$\xi_1, \xi_2 \in \mathbb{S}^{d-1}$. We have

$$Z_{\xi_1}^{(n)}(\eta_1)Z_{\xi_2}^{(n)}(\eta_2) = \mathbb{E}_{a_1}\left(Z_{\xi_1}^{(n)}(a_1)|d(\xi_1,a_1) = d(\xi_2,\eta_2)\right)$$
$$= \mathbb{E}_{a_2}\left(Z_{\eta_2}^{(n)}(a_2)|d(\xi_2,a_2) = d(\xi_1,\eta_1)\right).$$

Proof: The first equality is a change of variable, thanks to symmetry of the zonal harmonic in its variables and invariance with respect to action of the measure preserving elements of the orthogonal group fixing poles$\xi_1$orξ_2, that we now detail. By a rotation in one of the spheres, assume$\xi_1,\xi_2 = \xi$. Take a small ball

$$B_{\xi,\eta_1}(a_2^0;\varepsilon_2) = \{a_2: d(a_2,a_2^0) < \varepsilon_2\} \cap \{a_2: d(a_2,\xi) = d(\eta_1,\xi)\}.$$

Note$\{a_2: d(a_2,\xi) = d(\eta_1,\xi)\} \sim \mathbb{S}^{d-2}$. Every$a_2 \in B_{\xi,\eta_1}(a_2^0;\varepsilon_2)$gives rise to a canonical orthogonal mapσ_{a_2}along geodesics in a scaled copy of$\mathbb{S}^{d-2}$. Lifted to$\mathbb{S}^{d-1}$, these are orthogonal maps fixing ξ. Let σ^0 fix ξ and map a_2^0 to η_1. Let $a_1^0 = \sigma^0(\eta_2)$. We observe that $\{\sigma^0\sigma_{a_2}(\eta_2): a_2 \in B_{\xi,\eta_1}(a_2^0;\varepsilon_2)\} = B_{\xi,\eta_2}(a_1^0;\varepsilon_1)$with ε_1 so that

$$\mathbb{P}(d(a_2,a_2^0) < \varepsilon_2|d(\xi,a_2) = d(\xi,\eta_1)) = \mathbb{P}(d(a_1,a_1^0) < \varepsilon_1|d(\xi,a_1) = d(\xi,\eta_2))$$

Together with the symmetry and the rotation property $Z_\eta^{(n)}(a) = Z_a^{(n)}(\eta) = Z_{\sigma(a)}^{(n)}(\sigma(\eta))$ we obtain the first equality.

For fixed a_1, the function $Z_{\eta_1}^{(n)}(a_1) = Z_{a_1}^{(n)}(\eta_1)$ is a function harmonic in $\mathbb{R}^d$, n-homogeneous. These properties are preserved when taking expectation in a_1. So the expression $\mathbb{E}(Z_{\eta_1}^{(n)}(a_1)|d(\xi_1,a_1) = d(\xi_2,\eta_2))$ remains harmonic (regarded as a function in $\mathbb{R}^d$), n-homogeneous. From the form $\mathbb{E}(Z_{\eta_2}^{(n)}(a_2)|d(\xi_2,a_2) = d(\xi_1,\eta_1))$, we learn that its restriction to $\mathbb{S}^{d-1}$ depends only upon$d(\xi_1,\eta_1)$. This implies that it is a constant multiple of the zonal harmonic with pole ξ_1. Exchanging the roles of η_1 and $\eta_2$gives

$$\mathbb{E}\left(Z_{\eta_1}^{(n)}(a_1)\middle|d(\xi_1,a_1) = d(\xi_2,\eta_2)\right) = c_n Z_{\xi_1}^{(n)}(\eta_1)Z_{\xi_2}^{(n)}(\eta_2)$$

When assuming the normalization $Z_\xi^{(n)}(\xi) = 1$ then $c_n = 1$.

This is a generalization of the classical symmetrizing of the cosinus sum formula

$$1/2\,(\cos(x+y) + \cos(x-y)) = \cos(x)\cos(y)$$

Lemma (4.2.14)[115]: Let$C_i$be a multiplier in$\bigotimes_{k=1}^{i}\mathbb{R}^{d_k}$as described above, with any fixed direction and aperture. Let m be an integer of order$d = \max_{d_k}$. For any$\delta > 0$, the function Cihas an approximation by a polynomialC_i^Nin the$\prod_{k=1}^{i} d_k$variables $\{\Pi_{k:1\le k\le i}\eta_k, j_k|1 \le j_k \le d_k\}$so that$\|C_i - C_i^N\|_{C^m(\mathbb{S}^{d_k-1})} < \delta$in each variable separately.

C^mindexes the norm of uniform convergence on functions that are m times continuously differentiable. On the space side, C_i^Ncorresponds to an operator that is a polynomial in Riesz transforms of the variables$\bigotimes_k R_{k,j_k}$.

Proof. It is well known that zonal harmonic series have convergence proper-ties when representing smooth zonal functions similar to that of the Fourier transform. For any given m and sufficiently smooth φ of the type described above, then

$$C_1(\xi_1;\eta_1) = \sum_n \varphi_n Z_{\xi_1}^{(n)}(\eta_1)$$

where the convergence is C^m-uniform. The degree of smoothness required for φ to obtain convergence in the C^m in the above expression depends upon m and the dimension d. For our purpose, we choose $m \ge d$.

Let us denote this function's representation of degree N by a series of zonal harmonics by $C_1^{(N)}(\xi_1;\eta_1)$.

$$C_1^{(N)}(\xi_1;\eta_1) = \sum_{n\leq N} \varphi_n Z_{\xi_1}^{(n)}(\eta_1)$$

For every $\delta > 0$ there exists N so that we have the estimate

$$\left\| C_1^{(N)}(\xi_1;\eta_1) - C_1(\xi_1;\eta_1) \right\|_{C^m(\mathbb{S}^{d_1-1})} < \delta$$

In the case of icopies of spheres, we define $C_i^{(N)}$inductively in the same manner as C_i. Let us for the moment make all dimensions equal using the argument discussed above. So we set

$$C_k^{(N)}(\xi_1, \dots, \xi_{k;}\eta_1, \dots, \eta_k)$$
$$= \mathbb{E}_{\alpha_k-1}(C_{k-1}^{(N)}(\xi_1, \dots, a_{k-1;}\eta_1, \dots, \eta_{k-1})|d(a_{k-1}, \xi_{k-1}) = d(\eta_k, \xi_k))$$

We claim the identity

$$C_i^{(N)}(\xi;\eta_1, \dots, \eta_i) = \sum_{n\leq N} \varphi_n \prod_{k=1}^{i} Z_{\xi_k}^{(n)}(\eta_k) \qquad (14)$$

This is trivially true for $i = 1$. For $i > 1$induct on the number of parameters:

$$C_i^{(N)}(\xi;\eta_1, \dots, \eta_i)$$
$$= \mathbb{E}_{\alpha_i-1}(C_{i-1}(\xi_1, \dots, a_{i-1;}\eta_1, \dots, \eta_{i-1})|d(a_{i-1}, \xi_{i-1}) = d(\eta_i, \xi_i))$$
$$= \mathbb{E}_{\alpha_i-1}(\sum_{n\leq N} \varphi_n \prod_{k=1}^{i} Z_{\xi_k}^{(n)}(\eta_k)\ |d(a_{i-1}, \xi_{i-1}) = d(\eta_i, \xi_i))$$
$$= \sum_{n\leq N} \varphi_n \prod_{k=1}^{i} Z_{\xi_k}^{(n)}(\eta_k)\ \mathbb{E}_{\alpha_i-1}(Z_{\xi_{i-1}}^{(n)}|d(a_{i-1}, \xi_{i-1}) = d(\eta_i, \xi_i))$$
$$= \sum_{n\leq N} \varphi_n \prod_{k=1}^{i} Z_{\xi_k}^{(n)}(\eta_k)$$

The first equality is the definition of $C_i^{(N)}$, the second one is the induction hypothesis and the last an application of Lemma8.

It follows that neither C_i nor $C_i^{(N)}$depend on the order chosen in their definition and

$$C_i(\xi;\eta_1, \dots, \eta_i) = \sum_{n\leq N} \varphi_n \prod_{k=1}^{i} Z_{\xi_k}^{(n)}(\eta_k)$$

where the convergence is in C^m in each variable.

Next, we study the terms arising in multipliers of the form $C_i^{(N)}$. When all dimen-sions are equal, indeed, $\sum_{n\leq N} \varphi_n \prod_{k=1}^{i} Z_{\xi_k}^{(n)}(\eta_k)$ has the important property that, as a product of n-homogeneous polynomials, has only terms of the form

$$\prod_{k=1}^{i} \eta_k^{\alpha_k} = \prod_{k=1}^{i} (\prod_{jk=1}^{d} \eta_{k,j_k}^{\alpha_{k,j_k}})$$

where $\eta_k \in \mathbb{S}^{d-1}$and $\alpha_k = \alpha_{k,j_k}$ are multi-indices with $|\alpha_k| = \sum_{j_k} \alpha_{k,j_k}$ for all k. This form is inherited by $C_i^{(N)}$ with varying n. It shows that $C_i^{(N)}$ is indeed a polynomial in the variables$\prod_{k=1}^{i} \eta_{k,k_k}$. When the dimensions d_k are not equal, observe that by restricting back to the original number of variables, we certainly lose harmonicity of the polynomials, but not n-homogeneity or the required form of our polynomials.

Lemma (4.2.15)[115]: We are in$\mathbb{R}^{\boldsymbol{d}}$with partition$\mathcal{J} = (I_s)_{1\le s\le l}$. Let$\Upsilon$consist of vectors$\boldsymbol{\xi} = (\xi)_{k=1}^{t}$with$\xi_k \in \Upsilon_k$. Let$\Upsilon^{(s)}$consist of$\boldsymbol{\xi}^{(s)}$=$(\xi_k)_{k\in I_s}$. Let us consider the class of Journé type cone multipliers$\boldsymbol{J}_s = C_{i_s}(\boldsymbol{\xi}^{(s)};.)$of aperture rs with associated multi-parameter Calderón–Zygmund operators$\boldsymbol{T}_{s,\boldsymbol{J}_s}$. Then we have the two-sided estimate

$$\|b\|_{BMO_{\mathcal{J}}(\mathbb{R}^d)} \lesssim \sup_{\boldsymbol{\xi}\in\boldsymbol{\Upsilon}} \left\| \left[\boldsymbol{T}_{1,\boldsymbol{J}_1}, \dots, \left[\boldsymbol{T}_{l,\boldsymbol{J}_{l'}}, b\right] \dots\right]\right\|_{L^p(\mathbb{R}^{\boldsymbol{d}})\circlearrowleft} \lesssim \|b\|_{BMO_{\mathcal{J}}(\mathbb{R}^{\boldsymbol{d}})}.$$

Proof. By Lemma (4.2.12) we know that for each parameter $1 \le s \le l$ there exists a tensor product of cones $D_s = \otimes_{k\in I_s} D(\xi_k, \tau_k)$ with $\tau_k := \sum_{k\in I_s} \tau_k < \pi/2$ and $\xi_k \in \Upsilon_k$ and test functions fssupported as described in Lemma(4.2.12) part (ii)so that

$$\left\|\left[T_{1D_1}, \dots, \left[T_{lD_l}, b\right] \dots\right](f)\right\|_2 \gtrsim \|b\|_{BMO_I(\mathbb{R}^d)}$$

where $f = \otimes_{s=1}^{l} f_s$. We make a remark about the apertures r_s. Let $d(\cdot,\cdot)$ denote geodesic distance on $\mathbb{S}^{d-1}$, where antipodal points have distance π. Let $\boldsymbol{\xi}^{(s)}$ be the set of directions of $\boldsymbol{D_s}$. Remember that according to Lemma(4.2.12), one component had a specific direction $\xi_v^{(s)} \in \Upsilon_v$ and possibly large aperture with$(1+\tau)\tau_v^{(s)} < \pi/2$. Let us choose the other directions arbitrarily but with apertures $\tau_{\hat{v}}^{(s)}$small enough so that $(1+\tau)(\tau_v^{(s)} + (i-1)\tau_{\hat{v}}^{(s)}) < \pi/2$. Now choose an aperture $r_s < \pi/2$ so that $(1+\tau)(\tau_v^{(s)} + (i-1)\tau_{\hat{v}}^{(s)}) < r_s < \pi/2$.

Writing $i_s = |I_s|$, we find Journé type cone multipliers $\boldsymbol{J_s} = \boldsymbol{C}_{\boldsymbol{i_s}}(\boldsymbol{\xi}^{(s)};\cdot)$ ̈qaccording to the construction above with center $\boldsymbol{\xi}^{(s)}$ and aperture r_s.

We are going to observe that $\boldsymbol{J_s} = 1$on $spt(\boldsymbol{D_s})$ and $\boldsymbol{J_s} = -1$ on the Fourier support of f_s. Let us drop the dependence on s for the moment. We see in an inductive manner that $C_i(\boldsymbol{\xi};\cdot)$ takes the value 1 in a certain ℓ^1 ball of radius $r < \pi/2$centered at $\boldsymbol{\xi}$. We show that

$$\prod_k d(\xi_k, \eta_k) < r \Longrightarrow C_i(\xi; \eta_1, \dots, \eta_i) = 1$$

When $i = 1$, the assertion is obviously true: $d(\xi_1, \eta_1) < r \Longrightarrow C_1(\xi_1; \eta_1) = 1$ by the choice of φ, rand definition of C_1. For $i > 1$, we proceed by induction. As-sume that $\sum_{k=1}^{i-1} d(\xi_k, \eta_k) < r$ implies $C_{i-1}(\xi_1, \dots, \xi_{i-1;}\eta_1, \dots, \eta_{i-1}) = 1$. Let us assume that $\sum_{k=1}^{i} d(\xi_k, \eta_k) < r$. Remembering the definition of $C_i(\boldsymbol{\xi};\cdot)$ we assumed$d(a_{i-1}, \xi_{i-1}) = d(\eta_i, \xi_i)$. By the triangle inequality $\sum_{k=1}^{i-1} d(\xi_k, \eta_k) + d(a_{i-1}, \eta_{i-1}) \le \sum_{k=1}^{i-2} d(\xi_k, \eta_k) + d(a_{i-1}, \xi_{i-1}) + d(\xi_{i-1}, \eta_{i-1}) = \sum_{k=1}^{i} d(\xi_k, \eta_k) < r$ So

$$C_{i-1}(\xi_1, \dots, a_{i-1;}\eta_1, \dots, \eta_{i-1}) = 1$$

for all a_{i-1} relevant to the conditional expectation in the definition of $C_i(\boldsymbol{\xi};\cdot)$. The statement for I follows.

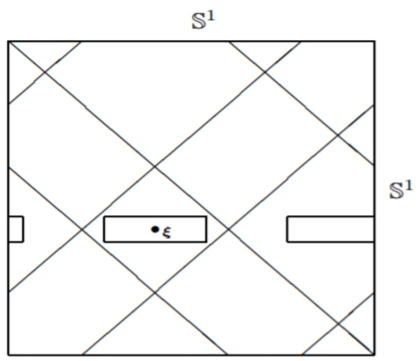
Since $C_i(\boldsymbol{\xi};\cdot)$ does not depend upon the order of the variables in its construction, we are also able to see exactly as done above that when $\sigma_k = -1$for exactly one choice of k, then$\sum_k d(\sigma_k \xi_k, \eta_k) < r \Longrightarrow C_i(\xi, \eta_1, \dots, \eta_i) = -1$. Consider associated multi-parameter

Calderón–Zygmund operators $\boldsymbol{T}_{s,\boldsymbol{J}_s}$ and $Id_s = \otimes_{k\in I_s} Id_k$and Id_k the identity on the kthvariable. Now

$$[T_{1,J_1}, \dots [T_{1,J_1}, b] \dots](f) = [T_{1,J_1} + Id_1, \dots, [T_{l,J_l} + Id_l, b] \dots](f)$$
$$= \otimes_{s=1}^{l} (T_{s,J_s} + Id_s)(bf).$$

With $\left\| \otimes_{s=1}^{l} (T_{s,J_s} + Id_s)(bf) \right\|_2 \geq \left\| \otimes_{s=1}^{l} T_{s,D_s}(bf) \right\|_2$ and $\otimes_{s=1}^{l} T_{s,D_s}(bf) = [T_{1,D_1}, \dots [T_{1,D_1}, b] \dots](f)$ we get the desired lower bound on the Journé commuta-tor as claimed.

Let us illustrate the base case $(\mathbb{S}^1)^2$by a picture. The picture is simplified in the sense that the odd function φ above is replaced by an indicator function of an interval.



Cone functions based on the oblique strips containing $\boldsymbol{\xi}$ are averaged. In the two-dimensional case, $\mathbb{S}^1$, expectation is over a one or two point set only. The rectangle around $\boldsymbol{\xi}$ with sides parallel to the axes representing $\mathbb{S}^1$illustrates the support of the tensor product of cone operators with direction $\boldsymbol{\xi}$. The longer side is the aperture that arises from the Hankel part. The short sides can be chosen freely as they arise from the Toeplitz part and are chosen small so that the rectangle fits into the oblique square. The other small rectangle corresponds to the Fourier support of the test function f.

Theorem (4.2.16)[115]: The following are equivalent with linear dependence of the respective norms.

(i) $b \in BMO_{\mathcal{J}}(\mathbb{R}^{\boldsymbol{d}})$.

(ii) All commutators of the form$\left[R_{k_1,j_{k_1}}, \dots, \left[R_{k_l,j_{k_l}}, b\right] \dots\right]$are bounded in $L^2(\mathbb{R}^{\boldsymbol{d}})$where$k_s \in I_s$ and$R_{k_s,j_{k_s}}$is the one-parameter Riesz transform in direction j_{k_s}.

(iii) All commutators of the form$\left[\boldsymbol{R}_{1,j^{(1)}}, \dots, \left[\boldsymbol{R}_{l,j^{(l)}}, b\right] \dots\right]$are bounded in $L^2(\mathbb{R}^{\boldsymbol{d}})$ where$\boldsymbol{j}^{(s)} = (j_k)_{k\in I_s}, 1 \leq j_k \leq d_k$and the operators$\boldsymbol{R}_{s,\boldsymbol{j}^{(s)}}$are a tensor product of Riesz transforms$\boldsymbol{R}_{s,\boldsymbol{j}^{(s)}} = \otimes_{k\in I_s} R_{k,j_k}$.

Proof. In contrast to the Hilbert transform case, both lower bounds require separate proofs. This is a notable difference that stems from the loss of orthogonal subspaces in conjunction with the special form of the Hilbert transform only seen in one variable. It does not seem possible to get a lower estimate (iii)⟺(ii)directly.

(i)⟺(ii). The upper bound (i)⟹(ii)is an easy consequence of the upper estimates of iterated commutators of single Riesz transforms. The lower bound (ii)⟹(i)follows from a standard fact on multipliers in combination with the main result in [111], the two-sided estimate for iterated commutators with Riesz transforms, similar to the first arguments used in 4.

(i)⟺(iii). The upper bound (i)⟹(iii)follows from the tensor product structure and use of the little product BMO norm. The lower bound (iii)⟹(i)uses the considerations leading up to this proof: Suppressing again the dependence on s, we see that the multiplier C_i is an odd, smooth, bounded function in each η_k when the other variables are fixed. Furthermore, since φ, written as a function of $t = \langle \xi, \eta \rangle$ is odd with respect to $t = 0$, the above series has $\varphi_n \neq 0$ at most when n is odd and thus $Z_\xi^{(n)}$is odd. So $C_i^{(N)}$is as a sum of odd functions odd.

We see that $\boldsymbol{T_J}$, the Journé operator associated to the cone $\boldsymbol{J} = C_i(\boldsymbol{\xi}; \cdot)$ as well as the operator associated to $C_i^{(N)}(\boldsymbol{\xi}; \cdot)$ are paraproduct free. In fact, applied to a test function $f = \otimes_k f_k$ with f_k acting on the kth variable and where $f_l = 1$for some lgives $\boldsymbol{T_J}(f) = 0$. To see this, apply the multiplier $C_i^{(N)}(\boldsymbol{\xi}; \cdot)$ in the lvariable (acting on 1) first, leaving the other Fourier variables fixed. The multiplier function

$$\eta_l \longmapsto C_i^{(N)}(\boldsymbol{\xi};\, \eta_1, \dots, \eta_i) = \sum_{n \leq N} \varphi_n Z_{\xi_l}^{(n)}(\eta_l) \prod_{k \neq l, k=1}^{i} Z_{\xi_k}^{(n)}(\eta_k)$$

is, as a sum of odd functions, odd on $\mathbb{S}^{d_l - 1}$, bounded by 1 and uniformly smooth for all choices of η_k with $k \neq l$. As such it gives rise to a paraproduct free convolution type Calderón–Zygmund operator in the lth variable whose values are multi-parameter multiplier operators.

Due to the convergence properties proved above, the difference

$$C_i(\boldsymbol{\xi}; \cdot) - C_i^{(N)}(\boldsymbol{\xi}; \cdot)$$

gives rise to a paraproduct free Journé operator with Calderón–Zygmund norm depending on N. This is seen by an application of an appropriate version of the Marcinkievicz multiplier theorem.

By our stability result on Journé commutators, Corollary(4.2.20), there exist for all $1 \leq s \leq l$ integers N_s so that $C_s^{(N_s)}(\boldsymbol{\xi_s}; \cdot)$ with $\xi_k \in \Upsilon_k$ is a characterizing set of operators via commutators for $BMO_I(\mathbb{R}^d)$. This is a finite set of possibilities because of the universal choice of the rsand finiteness of the set $\boldsymbol{\Upsilon}$. Using the multi-parameter ana-log of the observation $[AB, b] = A[B, b] + [A, b]B$ and the special form of the $C_s^{(N_s)}(\boldsymbol{\xi}; \cdot)$ leaves us with the desired lower bound: Observe that when $[AB, b]$ has large L^2 norm then either $[A, b]$ or $[B, b]$ has a fair share of the norm. We use this argument finitely many times in a row for operators that are polynomials in tensor products of Riesz transforms $\otimes_{k \in I_s} R_{k, j_k}$. This finishes (iii)⇒(i).

We remark that there are two cases of dimension greater than 1, where the proof simplifies. In the case of arbitrarily many copies of $\mathbb{R}^2$, one can work with the multi-plicative structure of complex numbers and avoid the symmetrizing procedure to obtain cone functions with the appropriate polynomial approximations. If the dimensions are arbitrary, but only tensor products of two Riesz transforms arise, one can avoid part of the construction above by using the addition formula for zonal harmonics.

We are interested in upper bounds for commutator norms by means of little product BMO norms of the symbol. In the case of the Hilbert transform, we have seen that these estimates, even in the iterated case, are straightforward. Other streamlined proofs exist for Hilbert or Riesz transforms when considering dyadic shifts of complexity one, see [126], [127] and [112]. When considering more general Calderón–Zygmund operators, the arguments required are more difficult, in each case. The first classical upper bound goes back to [3], where an estimate for one-parameter commutators with convolution type Calderón–Zygmund operators is given. Next, [111] includes a technical estimate for the multi-parameter case for such Calderón–Zygmund operators with a high enough degree of smoothness. This smoothness assumption was removed in [106]thanks to an approach using the representation formula for Calderón–Zygmund operators by means of infinite complexity dyadic shifts [110]. This last proof also gives a control on the norm of the commutators which depends on the Calderón–Zygmund norm of the operators themselves, a fact we will employ later. Below, we give an estimate by little product BMO when the Calderón–Zygmund operators are of Journé type and cannot be written as a tensor product. While this estimate is interesting in its own right, remember that it is also essential for our characterization result, the lower estimate. The first generation of multi-parameter singular integrals that are not of tensor product type goes back to Fefferman [120]and was generalized by Journé in [16]to the non-convolution type in the framework of his $T_q^{p_1}$ theorem in this setting. Much later, Journé's $T_q^{p_1}$ theorem was revisited, for example in [123], [124], [125]. See also [114] for some difficulties related to this subject. The references [123]in the bi-parameter case and [125]in the general multi-parameter case include a representation formula by means of adapted, infinite complexity dyadic shifts. While these representation formulae look complicated, they have a feature very useful to us. 'Locally', in a dyadic sense, they look as if they were of tensor product type, a feature we will exploit in the argument below. We start with the simplest bi-parameter case with no iterations and make comments about the generalization.

The class of bi-parameter singular integral operators treated is that of any paraproduct free Journé type operator (not necessarily a tensor product and not necessarily of convolution type) satisfying a certain weak boundedness property, which we define as follows:

Definition (4.2.17)[115]: A continuous linear mapping $T: C_0^\infty(\mathbb{R}^n) \otimes C_0^\infty(\mathbb{R}^m) \to [C_0^\infty(\mathbb{R}^n) \otimes C_0^\infty(\mathbb{R}^m)]'$is called a *paraproduct free bi-parameter Calderón–Zygmund operator* if the following conditions are satisfied:

(i) T is a Journé type bi-parameter δ-singular integral operator, i.e. there exists a pair (K_1, K_2) of $\delta CZ\text{–}\,\delta$-standard kernels so that, for all $f_1, g_1 \in C_0^\infty(\mathbb{R}^n)$ and $f_2, g_2 \in C_0^\infty(\mathbb{R}^m)$

$$\langle T(f_1 \otimes f_2), g_1 \otimes g_2\rangle = \int f_1(y_1)\langle K_1(x_1, y_1) f_2, g_2\rangle g_1(x_1) dx_1 dy_1$$

when $sptf_1 \cap sptg_1 = \emptyset$,

$$\langle T(f_1 \otimes f_2), g_1 \otimes g_2\rangle = \int f_2(y_2)\langle K_2(x_2, y_2) f_1, g_1\rangle g_2(x_2) dx_2 dy_2$$

when $sptf_2 \cap sptg_2 = \emptyset$.

(ii) T satisfies the weak boundedness property $|\langle T(\chi_I \otimes \chi_J), \chi_I \otimes \chi_J\rangle| \lesssim |I||J|$, for any cubes $I \subset \mathbb{R}^n$, $J \in \mathbb{R}^m$.

(iii) T is paraproduct free in the sense that $T(1 \otimes\cdot) = T(\cdot\otimes 1) = T^*(1 \otimes\cdot) = T^*(\cdot\otimes 1) = 0$.

Recall that a $\delta CZ\text{–}\,\delta$-standard kernel is a vector valued standard kernel taking values in the Banach space consisting of all Calderón–Zygmund operators. It is easy to see that an operator defined as above satisfies all the characterizing conditions in Martikainen [123], hence is L^2 bounded and can be represented as an average of bi-parameter dyadic shift operators together with dyadic paraproducts. Moreover, since Tis paraprod-uct free, one can conclude from observing the proof of Martikainen's theorem, that all the dyadic shifts in the representation are cancellative.

The base case from which we pass to the general case below, is the following:

Theorem (4.2.18)[115]: *Let T be a paraproduct free bi-parameter Calderón–Zygmund operator, and b be a little bmo function, there holds*

$$\|[b,T]\|_{L^2(\mathbb{R}^n\times\mathbb{R}^m)} \lesssim \|b\|_{bmo(\mathbb{R}^n\times\mathbb{R}^m)}$$

where the underlying constant depends only on the characterizing constants of T.

Proof: According to the discussion above, for any sufficiently nice functions f, g, one has the following representation:

$$\langle Tf,g\rangle = C\mathbb{E}_{\omega_1}\mathbb{E}_{\omega_2}\sum_{i_1,j_1=0}^{\infty}\sum_{i_2,j_2=0}^{\infty} 2^{-\max(i_1,j_1)}2^{-\max(i_2,j_2)}\langle S^{i_1j_1i_2j_2}f,g\rangle$$

where expectation is with respect to a certain parameter of the dyadic grids. The dyadic shifts $S^{i_1j_1i_2j_2}$ are defined as

$$S^{i_1j_1i_2j_2}f := \sum_{K_1\in D_1}\sum_{\substack{I_1,J_1\subset K_1, I_1,J_1\subset D_1\\ \ell(I_1)=2^{-i_1}\ell(K_1)\\ \ell(J_1)=2^{-j_1}\ell(K_1)}}\sum_{K_2\in D_2}\sum_{\substack{I_2,J_2\subset K_2, I_2,J_2\subset D_2\\ \ell(I_2)=2^{-i_2}\ell(K_2)\\ \ell(J_2)=2^{-j_2}\ell(K_2)}} a_{I_1J_1K_1I_2J_2K_2}\langle f,h_{I_1}\otimes h_{I_2}\rangle\, h_{J_1}\otimes h_{J_2}$$
$$=\sum_{K_1}\sum_{I_1,J_1\subset K_1}\sum_{K_2}\sum_{I_2,J_2\subset K_2} a_{I_1J_1K_1I_2J_2K_2}\langle f,h_{I_1}\otimes h_{I_2}\rangle\, h_{J_1}\otimes h_{J_2}$$

The coefficients above satisfy $a_{I_1J_1K_1I_2J_2K_2} \le \frac{\sqrt{|I_1||J_1||I_2||J_2|}}{|K_1||K_2|}$, which also guarantees the normalization $\|S^{i_1j_1i_2j_2}\|_{L^2\to L^2} \le 1$. Moreover, since T is paraproduct free, all the Haar functions appearing above are cancellative.

It thus suffices to show that for any dyadic grids D_1, $D_2$and fixed i_1, j_1, i_2, $j_2 \in \mathbb{N}$, one has

$$\left\|[b,S^{i_1j_1i_2j_2}]f\right\|_{L^2} \lesssim (1+\max(i_1,j_1))(1+\max(i_2,j_2))\|b\|_{bmo}\|f\|_{L^2}$$

as the decay factor $2^{-\max(i_1,j_1)}2^{-\max(i_2,j_2)}$in (iii)will guarantee the convergence of the series.

To see [2], one decomposes band a L^2test function fusing Haar bases:

$$[b,S^{i_1j_1i_2j_2}]f = \sum_{I_1,I_2}\sum_{J_1,J_2}\langle b,h_{I_1}\otimes h_{I_2}\rangle\langle f,h_{J_1}\otimes h_{J_2}\rangle[h_{I_1}\otimes h_{I_2},S^{i_1j_1i_2j_2}]\,h_{J_1}\otimes h_{J_2}$$

A similar argument to that in [106], implies that $[h_{I_1}\otimes h_{I_2},S^{i_1j_1i_2j_2}]h_{J_1}\otimes h_{J_2}$ is nonzero only if $I_1\subset J_1^{(i_1)}$or $I_2\subset J_2^{(i_2)}$, where $J_1^{(i_1)}$denotes the i_1 th dyadic ancestor of J_1, similarly for $J_2^{(i_2)}$. Hence, the sum can be decomposed into three parts: $I_1\subset J_1^{(i_1)}$ and $I_2\subset J_2^{(i_2)}$ (regular), $I_1 \supsetneq J_1^{(i_1)}$and $I_2\supsetneq J_2^{(i_2)}$, and $I_2\subset J_2^{(i_2)}$ (mixed).

Regular case (1): Following [106], one can decompose the arising sum into sums of classical bi-parameter dyadic paraproducts $B_0(b,f)$ and its slightly revised version $B_{kl}(b,f)$: for any integers $k,l\ge 0$, B_{kl} is the *bi*-parameter dyadic paraproduct defined as

$$B_{kl}(b,f)=\sum_{I,J}\beta_{IJ}\langle b,h_{I(k)}\otimes u_{J(l)}\rangle\langle f,h_I^{\epsilon_1}\otimes u_J^{\epsilon_2}\rangle h_I^{\epsilon_1'}\otimes u_J^{\epsilon_2'}\left|I^{(k)}\right|^{-1/2}\left|J^{(l)}\right|^{-1/2}$$

where β_{IJ} is a sequence satisfying $|\beta_{IJ}|\le 1$. When $k>0$, all Haar functions in the first variable are cancellative, while when $k=0$, there is at most one of $h_I^{\epsilon_1}$, $h_I^{\epsilon_1'}$ being noncancellative. The same assumption goes for the second variable. Observe that when $k=l=0$, B_{kl} becomes the classical paraproduct B_0. It is proved in [106]in Lemma4.1 that

$$\|B_{kl}(b,f)\|_{L^2}\lesssim\|b\|_{BMO}\|f\|_{L^2}$$

with a constant independent of k, land the product BMO norm on the right hand side.

Then since little BMO functions are contained in product BMO, this part can be controlled. More specifically, write

$$\begin{aligned}&[b,S^{i_1j_1i_2j_2}]f\\&=\sum_{I_1,I_2}\sum_{J_1,J_2}\langle b,h_{I_1}\otimes h_{I_2}\rangle\langle f,h_{J_1}\otimes h_{J_2}\rangle\, h_{I_1}\otimes h_{I_2}S^{i_1j_1i_2j_2}(h_{J_1}\otimes h_{J_2})\\&-\sum_{I_1,I_2}\sum_{J_1,J_2}\langle b,h_{I_1}\otimes h_{I_2}\rangle\langle f,h_{J_1}\otimes h_{J_2}\rangle\, S^{i_1j_1i_2j_2}(h_{I_1}h_{J_1}\otimes h_{I_2}h_{J_2})=I+II\end{aligned}$$

then one can estimate term I and II separately. According to the definition of dyadic shifts, term I is equal to

$$\sum_{J_1,J_2}\sum_{I_1:I_2\subset J_1^{(i_1)}}\sum_{I_2:I_2\subset J_2^{(i_1)}}\langle b,h_{I_1}\otimes h_{I_2}\rangle\langle f,h_{J_1}\otimes h_{J_2}\rangle\, h_{I_1}\otimes h_{I_2}.$$

$$\left(\sum_{\substack{J_1':J_1'\subset J_1^{(i_1)}\\ \ell(J_1')=2^{i_1-j_1}\ell(J_1)}}\ \sum_{\substack{J_2':J_2'\subset J_2^{(i_2)}\\ \ell(J_2')=2^{i_2-j_2}\ell(J_2)}} a_{J_1,J_1',J_1^{(i_1)},J_2,J_2',J_2^{(i_2)}}\, h_{J_1'}\otimes h_{J_2'}\right)$$

$$=\sum_{K_1,K_2}\ \sum_{J_1:J_1\subset K_1}^{(i_1)}\ \sum_{J_2:J_2\subset K_2}^{(i_2)}\ \sum_{I_1:I_1\subset K_1}\ \sum_{I_2:I_2\subset K_2}\langle b,h_{I_1}\otimes h_{I_2}\rangle\langle f,h_{J_1}\otimes h_{J_2}\rangle h_{I_1}\otimes h_{I_2}.$$

$$\left(\sum_{J_1':J_1'\subset K_1}^{(j_1)}\ \sum_{J_2':J_2'\subset K_2}^{(j_2)} a_{J_1,J_1',K_1,J_2,J_2',K_2}\, h_{J_1'}\otimes h_{J_2'}\right)$$

$$=\sum_{I_1,I_2}\langle b,h_{I_1}\otimes h_{I_2}\rangle h_{I_1}\otimes h_{I_2}\sum_{\substack{K_1\supset I_1\\K_2\supset I_2}}\ \sum_{J_1,J_1'\subset K_1}^{(i_1,j_1)}\ \sum_{J_1,J_2'\subset K_2}^{(i_2,j_2)} a_{J_1,J_1',K_1,J_2,J_2',K_2}\langle f,h_{J_1}\otimes h_{J_2}\rangle h_{J_1'}\otimes h_{J_2'}$$

$$=\sum_{I_1,I_2}\langle b,h_{I_1}\otimes h_{I_2}\rangle h_{I_1}\otimes h_{I_2}\sum_{J_1':J_1'\subset I_1}\ \sum_{J_2':J_2'\subset I_2}\langle S^{i_1j_1i_2j_2}f,h_{J_1'}\otimes h_{J_2'}\rangle\, h_{J_1'}\otimes h_{J_2'}$$

Because of the supports of Haar functions, the inner sum above can be further decomposed into four parts, where

$$I=\sum_{I_1,I_2}\sum_{J_1'\supsetneq I_1}\sum_{J_2'\supsetneq I_2}$$

$$II = \sum_{I_1,I_2} \sum_{J_1' \supsetneq I_1} \sum_{J_2':J_2' \subset I_2 \subset J_2^{(j_2)}}$$

$$III = \sum_{I_1,I_2} \sum_{J_1':J_1' \subset I_1 \subset J_1^{(j_1)}} \sum_{J_2' \supsetneq I_2}$$

$$IV = \sum_{I_1,I_2} \sum_{J_1' \subset I_1 \subset J_1^{(j_1)}} \sum_{J_2':J_2' \subset I_2 \subset J_2^{(j_2)}}$$

Hence, using the same technique as in [106], one has

$$I = \sum_{I_1,I_2} \langle b, h_{I_1} \otimes h_{I_2} \rangle \langle S^{i_1 j_1 i_2 j_2} f, h_{J_1'}^1 \otimes h_{J_2'}^2 \rangle h_{I_1} \otimes h_{I_2} |I_1|^{-1/2} |I_2|^{-1/2}$$

which is a bi-parameter paraproduct $B_0(b,f)$. Moreover, one has

$$II = \sum_{I_1,I_2} \langle b, h_{I_1} \otimes h_{I_2} \rangle h_{I_1} \otimes h_{I_2} \sum_{J_2':J_2' \subset I_2 \subset J_2^{(j_2)}} \langle S^{i_1 j_1 i_2 j_2} f, h_{I_1}^1 \otimes h_{J_2'} \rangle |I_1|^{-1/2} h_{J_2'}$$

where constants $\beta_{J_2'} \in \{1,-1\}$, and B_{0l} are the generalized bi-parameter paraproducts of type $(0,l)$ whose $L^2 \to L^2$ operator norm is uniformly bounded by $\|b\|_{\mathrm{BMO}}$ product BMO. Similarly, one can show that

$$III = \sum_{k=0}^{j_1} B_{k0}(b, S^{i_1 j_1 i_2 j_2} f) \quad IV = \sum_{k=0}^{j_1} \sum_{l=0}^{j_2} B_{kl}(b, S^{i_1 j_1 i_2 j_2} f)$$

Since $\|b\|_{\mathrm{BMO}} \lesssim \|b\|_{\mathrm{bmo}}$, all the forms above are L^2 bounded. This completes the discussion of term I.

To get an estimate of term II, we need to decompose it into finite linear combinations of $S^{i_1 j_1 i_2 j_2} B_{kl}(b,f)$. By linearity, one can write $S^{i_1 j_1 i_2 j_2}$ on the outside from the begin-ning, and we will only look at the inside sum. One splits for example the sum regarding the first variable into three parts: . If we split the second variabl$I_1 \subsetneq J_1, I_1 = J_1 , J_1 \subsetneq I_1 \subset J_1^{(i_1)}$e as well, there are nine mixed parts, and it's not hard to show that each of them can be represented as a finite sum of $B_{kl}(b,f)$. We omit the details.

Let's call the second and the third 'mixed' parts, and as the two are symmetric, it suffices to look at the second one, i.e. $I_1 \subset J_1^{(i_1)}$, $I_2 \supsetneq J_2^{(i_2)}$. In the first variable, we still have the old case $I_1 \subset J_1^{(i_1)}$ that appeared in [106], so morally speaking, we only need to nicely play around with the stronger little BMO norm to handle the second variable. For any fixed I_1, J_1, I_2, J_2, since $I_2 \supsetneq J_2^{(i_2)}$, the definition of dyadic shifts implies that

$$h_{I_1} \otimes h_{I_2} S^{i_1 j_1 i_2 j_2}\big(h_{J_1} \otimes h_{J_2}\big) = h_{I_1} S^{i_1 j_1 i_2 j_2}\big(h_{J_1} \otimes h_{I_2} h_{J_2}\big)$$

and

$$S^{i_1 j_1 i_2 j_2}\big(h_{I_1} h_{J_1} \otimes h_{I_2} h_{J_2}\big) = h_{I_2} S^{i_1 j_1 i_2 j_2}\big(h_{I_1} h_{J_1} \otimes h_{J_2}\big)$$

Hence, we still have cancellation in the second variable, which converts the mixed case to

$$\sum_{I_1 \subset J_1^{(i_1)}} \sum_{I_2 \supsetneq J_2^{(i_2)}} \langle b, h_{I_1} \otimes h_{I_2} \rangle \langle f, h_{J_1} \otimes h_{J_2} \rangle \big[h_{I_1}, S^{i_1 j_1 i_2 j_2}\big] \big(h_{J_1} \otimes h_{I_2} h_{J_2}\big)$$

$$= \sum_{I_1 \subset J_1^{(i_1)}} \sum_{J_2} \langle f, h_{J_1} \otimes h_{J_2} \rangle \left[h_{I_1}, S^{i_1 j_1 i_2 j_2} \right] \left(h_{J_1} \sum_{I_2 \supsetneq J_2^{(i_2)}} \langle b, h_{I_1} \otimes h_{I_2} \rangle \otimes h_{I_2} h_{J_2} \right)$$

$$= \sum_{I_1 \subset J_1^{(i_1)}} \sum_{J_2} \langle f, h_{J_1} \otimes h_{J_2} \rangle \left[h_{I_1}, S^{i_1 j_1 i_2 j_2} \right] \left(h_{J_1} \otimes \langle b, h_{I_1} \otimes h^1_{J_2^{(j_2)}} \rangle \otimes h^1_{J_2^{(j_2)}} h_{J_2} \right)$$

$$= \sum_{I_1 \subset J_1^{(i_1)}} \sum_{J_2} \langle b, h_{I_1} \otimes h_{I_2} \rangle \left| J_2^{(j_2)} \right|^{-1/2} \langle f, h_{J_1} \otimes h_{J_2} \rangle \left[h_{I_1}, S^{i_1 j_1 i_2 j_2} \right] \left(h_{J_1} \otimes h_{J_2} \right)$$

$$= \sum_{I_1 \subset J_1^{(i_1)}} \sum_{J_2} \langle \langle b \rangle_{J_2^{(j_2)}}, h_{I_1} \rangle_1 \langle f, h_{J_1} \otimes h_{J_2} \rangle \left[h_{I_1}, S^{i_1 j_1 i_2 j_2} \right] \left(h_{J_1} \otimes h_{J_2} \right)$$

where $\langle b \rangle_{J_2^{(i_2)}}$denotes the average value of bon $J_2^{(i_2)}$, which is a function of only the first variable.

In the following, we will once again estimate the first term and second term of the commutator separately, and the L^2norm of each of them will be proved to be bounded by $\|b\|_{\mathrm{bmo}} \|f\|_{L^2}$.

a) First term.

By definition of the dyadic shift, the first term is equal to

$$\sum_{I_1 \subset J_1^{(i_1)}} \sum_{J_2} \langle \langle b \rangle_{J_2^{(i_2)}}, h_{I_1} \rangle_1 \, h_{I_1} \langle f, h_{J_1} \otimes h_{J_2} \rangle$$

$$\Bigg(\sum_{\substack{J_1' : J_1' \subset J_1^{(i_1)} \\ \ell(J_1') = 2^{i_1 - j_1} \ell(J_1)}} \sum_{\substack{J_2' : J_2' \subset J_2^{(i_2)} \\ \ell(J_2') = 2^{i_2 - j_2} \ell(J_2)}} a_{J_1, J_1', J_1^{(i_1)}, J_2, J_2', J_2^{(i_2)}} h_{J_1'} \otimes h_{J_2'} \Bigg)$$

which by reindexing $K_1 := J_1^{(i_1)}$is the same as

$$\sum_{I_1, J_2} \langle \langle b \rangle_{J_2^{(i_2)}}, h_{I_1} \rangle_1 \, h_{I_1} \cdot \sum_{K_1 : K_1 \supset I_1} \sum_{J_1 \subset K_1}^{(i_1)} \sum_{J_1' \subset K_1}^{(j_1)} \sum_{J_2' \subset J_2^{(i_2)}}^{(j_2)} a_{J_1, J_1', K_1, J_2, J_2', J_2^{(i_2)}} \langle f, h_{J_1} \otimes h_{J_2} \rangle h_{J_1'} \otimes h_{J_2'}$$

$$= \sum_{I_1, J_2} \langle \langle b \rangle_{J_2^{(i_2)}}, h_{I_1} \rangle_1 \, h_{I_1} \sum_{J_1' : J_1'^{(j_1)} \supset I_1} h_{J_1'} \otimes \langle S^{i_1 j_1 i_2 j_2} \left(\langle f, h_{J_2} \rangle_2 \otimes h_{J_2} \right), h_{J_1'} \rangle_2$$

where the inner sum is the orthogonal projection of the image of $\langle f, h_{J_2} \rangle_2 \otimes h_{J_2}$under $S^{i_1 j_1 i_2 j_2}$ onto the span of $\{h_{J_1'}\}$such that $J_1'^{(j_1)} \supset I_1$. Taking into account the supports of the Haar functions in the first variable, one can further split the sum into two parts where

$$I = \sum_{J_2} \sum_{I_1 \subsetneq J_1'} \qquad II = \sum_{J_2} \sum_{J_1' \subset I_1 \subset J_1'^{(i_1)}}$$

Summing over J_1' first implies that

$$I = \sum_{J_2} \sum_{I_1} \langle \langle b \rangle_{J_2^{(i_2)}}, h_{I_1} \rangle_1 \, h_{I_1} \langle S^{i_1 j_1 i_2 j_2} \left(\langle f, h_{J_2} \rangle_2 \otimes h_{J_2} \right), h_{J_1'} \rangle_2$$

$$= \sum_{J_2} B_0 (\langle b \rangle_{J_2^{(i_2)}}, S^{i_1 j_1 i_2 j_2} \left(\langle f, h_{J_2} \rangle_2 \otimes h_{J_2} \right))$$

where $B_o(b,f) \coloneqq \sum_I \langle b, h_I\rangle \langle f, h_{\frac{1}{I}}\rangle h_I |I|^{-1/2}$is a classical one-parameter paraproduct in the first variable. Note that its L^2 norm is bounded by $\|b\|_{\mathrm{BMO}} \|f\|_{L^2}$. Moreover, according to the definition of $S^{i_1 j_1 i_2 j_2}$, for any fixed J_2

$$S^{i_1 j_1 i_2 j_2}\big(\langle f, h_{J_2}\rangle_2 \otimes h_{J_2}\big) = \sum_{J_2' : J_2'^{(j_2)} = J_2^{(j_2)}} \langle S^{i_1 j_1 i_2 j_2}\big(\langle f, h_{J_2}\rangle_2 \otimes h_{J_2}\big), h_{J_1'}\rangle_2 \otimes h_{J_2'}$$

In other words, $S^{i_1 j_1 i_2 j_2}\big(\langle f, h_{J_2}\rangle_2 \otimes h_{J_2}\big)$ only lives on the span of $\left\{h_{J_2'} : J_2'^{(j_2)} = J_2^{(i_2)}\right\}$. Hence, by linearity there holds

$$I = \sum_{J_2} \sum_{J_2' : J_2'^{(j_2)} = J_2^{(j_2)}} B_0\left(\langle b\rangle_{J_2^{(i_2)}}, \langle S^{i_1 j_1 i_2 j_2}\big(\langle f, h_{J_2}\rangle_2 \otimes h_{J_2}\big), h_{J_1'}\rangle_2\right) \otimes h_{J_2'}$$

$$\sum_{J_2} (B_0(\langle b\rangle_{J_2^{(i_2)}}, \langle S^{i_1 j_1 i_2 j_2}(\sum_{J_2' : J_2'^{(j_2)} = J_2^{(j_2)}} \big(\langle f, h_{J_2}\rangle_2 \otimes h_{J_2}\big), h_{J_1'}\rangle_2) \otimes h_{J_2'}$$

Thus, orthogonality in the second variable implies that

$$\|I\|^2_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}$$

$$= \sum_{J_2} \left\| (B_0(\langle b\rangle_{J_2^{(i_2)}}, \langle S^{i_1 j_1 i_2 j_2}(\sum_{J_2' : J_2'^{(j_2)} = J_2^{(j_2)}} \big(\langle f, h_{J_2}\rangle_2 \otimes h_{J_2}\big), h_{J_1'}\rangle_2) \right\|^2_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}$$

$$\lesssim \sum_{J_2} \left\| B_0(\langle b\rangle_{J_2^{(i_2)}} \right\|^2_{BMO(\mathbb{R}^n)} \left\| \langle S^{i_1 j_1 i_2 j_2}(\sum_{J_2' : J_2'^{(j_2)} = J_2^{(j_2)}} \big(\langle f, h_{J_2}\rangle_2 \otimes h_{J_2}\big), h_{J_1'}\rangle_2) \right\|^2_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}$$

Observing that $\left\|\langle b\rangle_{J_2'^{(j_2)}}\right\|_{BMO(\mathbb{R}^n)} \le \langle \|b\|_{BMO(\mathbb{R}^n)}\rangle_{J_2'^{(j_2)}} \le \|b\|_{bmo}$, one has

$$\le \|b\|^2_{bmo} \sum_{J_2'} \left\| \langle S^{i_1 j_1 i_2 j_2}\left(\sum_{J_2 : J_2^{i_2} = J_2'^{(j_2)}} \langle f, h_{J_2}\rangle_2 \otimes h_{J_2} \right), h_{J_2'}\rangle_2 \right\|^2_{L^2(\mathbb{R}^n)}$$

$$= \|b\|^2_{bmo} \left\| \sum_{J_2'} \langle S^{i_1 j_1 i_2 j_2}\left(\sum_{J_2 : J_2^{i_2} = J_2'^{(j_2)}} \langle f, h_{J_2}\rangle_2 \otimes h_{J_2} \right), h_{J_2'}\rangle_2 \right\|^2_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}$$

Note that the sum in the L^2 norm is in fact very simple:

$$\sum_{J_2'} \langle S^{i_1 j_1 i_2 j_2}\left(\sum_{J_2 : J_2^{i_2} = J_2'^{(j_2)}} \langle f, h_{J_2}\rangle_2 \otimes h_{J_2} \right), h_{J_2'}\rangle_2 \otimes h_{J_2'}$$

$$= \sum_{J_2} \sum_{J_2': J_2'^{(j_2)} = J_2'^{(j_2)}} \langle S^{i_1 j_1 i_2 j_2}\big(\langle f, h_{J_2}\rangle_2 \otimes h_{J_2}\big), h_{J_2'}\rangle_2 \otimes h_{J_2'}$$

$$= \sum_{J_2} S^{i_1 j_1 i_2 j_2}\big(\langle f, h_{J_2}\rangle_2 \otimes h_{J_2}\big) = S^{i_1 j_1 i_2 j_2}(f)$$

Hence, the uniform boundedness of the $L^2 \to L^2$ operator norm of dyadic shifts implies that

$$\|I\|^2_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} \lesssim \|b\|^2_{bmo} \|f\|^2_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}$$

In order to handle II, we split it into a finite sum depending on the levels of I_1upon J_1', which leads to

$$II = \sum_{k=0}^{j_1} \sum_{j_2} \sum_{j_1'} \langle \langle b \rangle_{J_2^{(i_2)}}, h_{J_1'^{(k)}}\rangle_1 \, h_{J_1'^{(k)}} h_{J_1'} \otimes \langle S^{i_1 j_1 i_2 j_2}\big(\langle f, h_{J_2}\rangle_2 \otimes h_{J_2}\big), h_{J_1'}\rangle_1$$

$$= \sum_{k=0}^{j_1} \sum_{j_2} \sum_{j_1'} \beta_{J_1',k} \left|J_1'^{(k)}\right|^{-1/2} \langle \langle b \rangle_{J_2^{(i_2)}}, h_{J_1'^{(k)}}\rangle_1 \, h_{J_1'^{(k)}} h_{J_1'} \otimes \langle S^{i_1 j_1 i_2 j_2}\big(\langle f, h_{J_2}\rangle_2 \otimes h_{J_2}\big), h_{J_1'}\rangle_1$$

$$\sum_{k=0}^{j_1} \sum_{j_2} B_k(\langle b \rangle_{J_2^{(i_2)}}, S^{i_1 j_1 i_2 j_2}\big(\langle f, h_{J_2}\rangle_2 \otimes h_{J_2}\big))$$

where $B_k(b, f) \coloneqq \sum_I \beta_{I,k} \langle b, h_{I(k)}\rangle \langle f, h_I\rangle h_I \left|I^{(k)}\right|^{-1/2}$ is a generalized one-parameter paraproduct studied in [106], whose L^2 norm is uniformly bounded by $\|b\|_{\mathrm{BMO}}\ \|f\|_{L^2}$, independent of k and the coefficients $\beta_{I,k} \in \{-1,1\}$. Then one can proceed as in part I to conclude that

$$\|II\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} \lesssim (1 + J_1)\|b\|_{bmo} \|f\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}$$

which together with the estimate for part I implies that

$$\|first\ term\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)} \lesssim (1 + j_1)\|b\|_{bmo} \|f\|_{L^2(\mathbb{R}^n \times \mathbb{R}^m)}$$

b) Second term.

As the second term by linearity is the same as

$$S^{i_1 j_1 i_2 j_2}\Big(\sum_{J_2} \sum_{I_1 \subset J_1^{(i_1)}} \langle \langle b \rangle_{J_2^{(i_2)}}, h_{I_1}\rangle_1 \, \langle f, h_{J_1} \otimes h_{J_2}\rangle \, h_{I_1} h_{J_1} \otimes h_{J_2}\Big)$$

the $L^2 \to L^2$boundedness of the shift implies that it suffices to estimate the L^2 norm of the term inside the parentheses. Since $I_1 \cap J_1 \neq \emptyset$, one can further split the sum into two parts:

$$I \coloneqq \sum_{J_2} \sum_{I_1 \subsetneq J_1}, \qquad II \coloneqq \sum_{J_2} \sum_{I_1 \subset J_1 \subset J_1^{(i_1)}}$$

Summing over J_1first implies that

$$I = \sum_{J_2} \sum_{I_1} \langle \langle b \rangle_{J_2^{(i_2)}}, h_{I_1}\rangle_1 \, \langle f, h^1_{I_1} \otimes h_{J_2}\rangle \, h_{I_1} h^1_{I_1} \otimes h_{J_2} =: \sum_{J_2} B_0\left(\langle b \rangle_{J_2^{(i_2)}}, \langle f, h_{J_2}\rangle_2\right) \otimes h_{J_2}$$

where $B_o(b, f) \coloneqq \sum_I \langle b, h_I\rangle \, \langle f, h_{\frac{1}{I}}\rangle \, h_I |I|^{-1/2}$is a classical one-parameter paraproduct in the first variable. Hence,

$$\|I\|^2_{L^2(\mathbb{R}^n\times\mathbb{R}^m)} = \sum_{J_2}\left\|B_0\left(\langle b\rangle_{J_2^{(i_2)}}, \langle f, h_{J_2}\rangle_2\right)\right\|^2_{L^2(\mathbb{R}^n)}$$
$$\lesssim \sum_{J_2}\left\|\langle b\rangle_{J_2^{(i_2)}}\right\|^2_{BMO(\mathbb{R}^n)}\|\langle f, h_{J_2}\rangle_2\|^2_{L^2(\mathbb{R}^n)} \le \|b\|^2_{bmo}\sum_{J_2}\|\langle f, h_{J_2}\rangle_2\|^2_{L^2(\mathbb{R}^n)}$$
$$= \|b\|^2_{bmo}\|f\|^2_{L^2(\mathbb{R}^n\times\mathbb{R}^m)}$$

For part II, note that it can be decomposed as

$$II = \sum_{k=0}^{i_1}\sum_{J_2}\sum_{J_1}\langle\langle b\rangle_{J_2^{(i_2)}}, h_{J_1^{(k)}}\rangle_1 \langle f, h_{J_1}\otimes h_{J_2}\rangle\, h_{J_1^{(k)}} h_{J_1}\otimes h_{J_2}$$
$$\sum_{k=0}^{i_1}\sum_{J_2}\sum_{J_1}\beta_{J_1,k}\left|J_1^{(k)}\right|^{-1/2}\langle\langle b\rangle_{J_2^{(i_2)}}, h_{J_1^{(k)}}\rangle_1 \langle\langle f, h_{J_2}\rangle_2, h_{J_1}\rangle_1\, h_{J_1}\otimes h_{J_2}$$
$$\sum_{k=0}^{i_1}\sum_{J_2} B_{,k}(\langle b\rangle_{J_2^{(i_2)}}, \langle f, h_{J_2}\rangle_2)\otimes h_{J_2}$$

where coefficients $\beta_{J_1k}\in\{1,-1\}$ and the L^2 norm of the generalized paraproduct B_k is uniformly bounded as mentioned before. Therefore, the same argument as for part I shows that

$$\|II\|_{L^2(\mathbb{R}^n\times\mathbb{R}^m)} \lesssim (1+i_1)\|b\|_{bmo}\|f\|_{L^2(\mathbb{R}^n\times\mathbb{R}^m)}$$

which completes the discussion of the second term, and thus proves that the mixed case is bounded.

The upper bound result we just proved can be extended to $\mathbb{R}^d$, to arbitrarily many parameters and an arbitrary number of iterates in the commutator. To do this, consider multi-parameter singular integral operators studied in [125], which satisfy a weak boundedness property and are paraproduct free, meaning that any partial adjoint of T is zero if acting on some tensor product of functions with one of the components being 1. And consider a little product BMO function $b\in BMO_{\mathcal{J}}(\mathbb{R}^d)$. One can then prove

Theorem (4.2.19)[115]: Let us consider $\mathbb{R}^{\boldsymbol{d}}, \boldsymbol{d}=(d_1,\dots,d_t)$ with a partition $\mathcal{I}=(I_s)_{1\le s\le l}$ of $\{1,\dots,t\}$ as discussed before. Let $b\in BMO_{\mathcal{J}}(\mathbb{R}^d)$ and let T_s denote a multi-parameter paraproduct free Journé operator acting on functions defined $\otimes_{k\in I_s}\mathbb{R}^{d_k}$. Then we have the estimate below

$$\|[T_1,\dots,[T_l,b]\dots]\|_{L^2(\mathbb{R}^{\boldsymbol{d}})\circlearrowleft} \lesssim \|b\|_{BMO_{\mathcal{J}}(\mathbb{R}^{\boldsymbol{d}})}.$$

The part of the proof that targets the Journé operators proceeds exactly the same as the bi-parameter case with the multi-parameter version of the representation theorem proven in [125]. Certainly, as the number of parameters increases, more mixed cases will appear. However, if one follows the corresponding argument above for each variable in each case, it is not hard to check that eventually, the boundedness of the arising paraproducts is implied exactly by the little product BMO norm of the symbol. The difficulty of higher iterates is overcome in observing that the commutator splits into commutators with no iterates, as was done in [106]. The assumption that the operators be paraproduct free is sufficient for our lower estimate. The general case is currently under investigation.

Important to our arguments for lower bounds with Riesz transforms is the corollary below, which follows from the control on the norm of the estimate in Theorem(4.2.19) by an application of triangle inequality. It is a stability result for characterizing families of Journé operators.

Corollary (4.2.20)[115]: Let for every $1 \le s \le l$ be given a collection$\tau_s = \{T_{s,j_s}\}$of paraproduct free Journé operators on $\bigotimes_{k\in I_s} \mathbb{R}^{d_k}$ that characterize$BMO_{\mathcal{J}}(\mathbb{R}^d)$ via a two-sided commutator estimate

$$\|b\|_{BMO_{\mathcal{J}}(\mathbb{R}^d)} \lesssim \sup_j \left\| \left[T_{1,J_1}, \dots, \left[T_{l,J_l}, b\right] \dots \right] \right\|_{L^2(\mathbb{R}^d) \circlearrowleft} \lesssim \|b\|_{BMO_{\mathcal{J}}(\mathbb{R}^d)}.$$

Then there exists $\varepsilon > 0$ such that for any family of paraproduct free Journé operators$\tau'_s = \{T'_{s,j_s}\}$ with characterizing constants$\|T'_{s,j_s}\|_{CZ} \le \varepsilon$, the family$\{T_{s,j_s} + T'_{s,j_s}\}$still characterizes $BMO_{\mathcal{J}}(\mathbb{R}^d)$.

It is well known, that theorems of this form have an equivalent formulation in the language of weak factorization of Hardy spaces. We treat the model case$\mathbb{R}^d = \mathbb{R}^{(d_1,d_2,d_3)}andBMO_{(13)(2)}(\mathbb{R}^d)$only for sake of simplicity. The other statements are an obvious generalization. For the corresponding collections of Riesz transforms $\mathcal{R}_{k,j_k}$and $b \in BMO_{(13)(2)}(\mathbb{R}^d)$,$1 \le s \le 3$, by unwinding the commutator one can define the operator Π_j such that

$$\langle \left[R_{2,j_2}, \left[R_{1,j_1} R_{3,j_3}, b\right]\right] f, g \rangle_{L^2} = \langle b, \Pi_j(f,g) \rangle_{L^2}.$$

Consider the Banach space$L^2 * L^2$of all functions in$L^1(\mathbb{R}^d)$of the form$f = \sum_j \sum_i \Pi_j(\phi_i^j, \psi_i^j)$normed by

$$\|f\|_{L^2 * L^2} = \inf \left\{ \sum_j \sum_i \|\phi_i^j\|_2 \|\psi_i^j\|_2 \right\}$$

with the infimum running over all possible decompositions of$f$. Applying a duality argument and the two-sided estimate in Corollary (4.2.20) we are going to prove the following weak factorization theorem.

Theorem(4.2.21)[115]: $H^1_{Re}\left(\mathbb{R}^{(d_1,d_2)}\right) \otimes L^1(\mathbb{R}^{d_3}) + L^1(\mathbb{R}^{d_1}) \otimes H^1_{Re}\left(\mathbb{R}^{(d_2,d_3)}\right)$ coincides with the space $L^2 * L^2$. In other words, for any $f \in H^1_{Re}\left(\mathbb{R}^{(d_1,d_2)}\right) \otimes L^1(\mathbb{R}^{d_3}) + L^1(\mathbb{R}^{d_1}) \otimes H^1_{Re}\left(\mathbb{R}^{(d_2,d_3)}\right)$ there exist sequences $\phi_i^j, \psi_i^j \in L^2$ such that $\|f\| \sim \sum_j \sum_i \|\phi_i^j\|_2 \|\psi_i^j\|_2$.

Proof: Let's first show that$L^2 * L^2$is a subspace of$H^1_{Re}\left(\mathbb{R}^{(d_1,d_2)}\right) \otimes L^1(\mathbb{R}^{d_3}) + L^1(\mathbb{R}^{d_1}) \otimes H^1_{Re}\left(\mathbb{R}^{(d_2,d_3)}\right)$. Recalling the remark after Theorem(4.2.3), this is the same as to show$\forall f \in L^2 * L^2$,$f$is a bounded linear functional on$BMO_{(13)(2)}(\mathbb{R}^d)$. This follows from the upper bound on the commutators since

$$\langle b, \sum_j \sum_i \Pi_j(\phi_i^j, \psi_i^j) \rangle = \sum_j \sum_i \langle \left[R_{2,j_2}, \left[R_{1,j_1} R_{3,j_3}, b\right]\right] \phi_i^j, \psi_i^j \rangle.$$

Now we are going to show

$$\sup_{f \in L^2 * L^2} \left\{ \left| \iint fb \right| : \|f\|_{L^2 * L^2} \le 1 \right\} \sim \|b\|_{BMO_{(13)(2)}}$$

which gives the equivalence of$H^1_{Re}\left(\mathbb{R}^{(d_1,d_2)}\right) \otimes L^1(\mathbb{R}^{d_3}) + L^1(\mathbb{R}^{d_1}) \otimes H^1_{Re}\left(\mathbb{R}^{(d_2,d_3)}\right)$ norm and the$L^2 * L^2$norm, thus showing that the two spaces are the same.

To see this, note that the direction $\lesssim$ is trivial, and the direction $\gtrsim$ is implied by the lower bound of commutators. For any $b \in BMO_{(13)(2)}(\mathbb{R}^d)$, there exists j such that $\|b\|_{BMO_{(13)(2)}} \lesssim \left\| \left[R_{2,j_2}, \left[R_{1,j_1} R_{3,j_3}, b \right] \right] \right\|$. Hence, there exist $\phi, \psi \in L^2$ with norm 1 such that

$$\|b\|_{BMO_{(13)(2)}} \lesssim \left| \left\langle \left[R_{2,j_2}, \left[R_{1,j_1} R_{3,j_3}, b \right] \right] \phi, \psi \right\rangle \right| = \left| \langle b, \Pi_j(\phi, \psi) \rangle \right| \leq LHS,$$

which completes the proof.

Chapter 5
Nonlinear Piecewise Polynomial Approximation

We consider uniform approximation and approximation in the metric of the spaces$L_{q,\sigma} = L_q(Q^m; \sigma)$, where $q > 1$ and σ is a finite Borel measure. We show that the main approximation result concerns the space V_{pq}^k of smoothness $s := d\left(\frac{1}{p} - \frac{1}{q}\right) \in (0, k]$. It asserts the following: Let $f \in V_{pq}^k$ be of smoothness $s \in (0, k]$, $1 \leq p < q < \infty$ and $N \in \mathbb{N}$. There exist a family Δ_N of N dyadic subcubes of $[0,1)^d$ and a piecewise polynomial g_N over Δ_N of degree $k - 1$ such that $\|f - g\| \leq CN^{-\frac{s}{d}}|f|_{V_{pq}^k}$. This implies similar results for the above mentioned smoothness spaces, in particular, solves the going back to the 1967 Birman–Solomyak [147] problem of approximation of functions from $W_p^k([0,1)^d\ )$ in $L_q([0,1)^d)$ whenever $\frac{k}{d} = \frac{1}{p} - \frac{1}{q}$ and $q < \infty$.

Section (5.1): Functions of the Classes W_p^α

For the approximation device we use functions which become polynomials of some fixed degree $l = l(\alpha)$on each of the cubes in a suitable partitioning of Q^minto cubes. The partition according to which we construct the approximating piecewise-polynomial function is not fixed in advance; only the number of cubes in it is restricted. The partition itself is chosen (according to the formulation of the problem) as a function either of the function being approximated or of the measure σ. Naturally such a choice makes possible a better rate of approximation. Indeed, as one of the results, the rate of approximation obtained in the uniform metric for $pa > m$ is the same as that for functions which have smoothness of order α in the classical sense.

The problem investigated arose first in connection with the development of the theory of double Stieltjes operator integrals [130]-[133]. Estimates of singular numbers of integral operators acting from the space $L_{2,\sigma}$into another space $L_{2,\tau}$are of considerable value for this theory. The specific problem is that we need estimates not depending (in the usual sense) on the measures σ and τ. The method of approximation presented here was in fact developed with the aim of investigating integral operators.

This method of approximation found another application in the theory of double Stieltjes operator integrals in so-called "interpolation in smoothness" (see [133]). Here the basic approximation results can be used directly, without the corresponding theorems on integral operators. We note that new results in the multidimensional problem of multipliers inl_p spaces (see[132], [133]) are a consequence of "interpolation in smoothness".

It has also become clear that the proposed untraditional tool for approximation is also useful io other problems. First of all, with it we can find the exact order of ϵ-entropy of the unit sphere [1)] $SW_p^a(Q^m)$, $pa > m$, as a compactum in $C(Q^m)$.

It is known [134] that the ϵ-entropy of the sphere $SC^a(Q^m)$ as a compactum in $C(Q^m)$has order$\epsilon^{-m^{a-1}}$. In this case "strong" and "weak" norms are similar in nature, so that we can use relatively simple approximation methods for estimating entropy. Sometimes we are concerned with estimating the entropy of the set $SW_p^a(Q^m)$in the metric of the space $C(Q^m)$. The norms in $W_p^a(Q^m)$and in $C(Q^m)$ate different in nature, and the norm in$C(Q^m)$is essentially more "restrictive" than the norm in $L_p(Q^m)$. As a result the classical linear approximation methods do not lead to an exact result in this case. The approximation method proposed makes it possible to find the precise order of ϵ-entropy, which again turns out to be $e^{-m^{a^{-1}}}$.

All the results obtained, in the final analysis, based on one special theorem on set functions. The proof of this theorem rests on a detailed investigation of a certain concrete algorithm for partitions of the cube Q^m. When applied to approximation theory, this algorithm generates a tool for approximations which is closely linked with the peculiarities in the formulation of the problem, which makes possible a good rate of approximations.

We study the classes W_p^a. We use only a few properties of these spaces: imbedding theorems in the spaces C and L_p, and also the property of homogeneity of the principal term of the norm with respect to similarity transformations of the region. As a result of this all the results also hold for other functional classes having the same properties. In particular, this refers to the spaces H_p^a of S. M. Nikol'skii and $B_p^a$of O. V. Besov (see e.g.[135]).

In the case $m = 1$ the theorems on approximation carry over to classes of functions of bounded β-variation. These results are of definite interest for estimating singular numbers of integral operators. Besides the results relating to the theory of approximation, estimates of e-entropy are also presented. Applications of the basic results to estimating singular numbers of integral operators are treated in [136].

We prove the basic theorem on partition functions. We derive theorems on approximation of functions of the class W_p^a. We obtain analogous results for functions of bounded β-variation. We give estimates of the ϵ-entropy of the set $SW_p^a(Q^m)$in$C(Q^m)$ for $pa > m$ and in$L_p(Q^m)$for $pa < m$, $q < mp(m - pa)^{-1}$.
Results for the case $pa > m$ (though more important) were published without proof in [137]

Let R^mbe the m-dimensional Euclidean space of points (vectors)$x = (x_1, \dots, x_m)$, $|x|$the length of the vector$x$. If $k = (k_1, \dots, k_m)$is a multi-index (allk_i. are integers,$\kappa > 0$), then$x^k = \prod_{i=1}^m x_i^{k_i}$,$|k| = \sum_{i=1}^m k_i$. Let$D^k$denote the differentiation operator:

$$D^x = \frac{\partial |x|}{\partial_{x_1}^{x_1} \dots \partial_{x_m}^{x_m}}.$$

Let$\Delta \subset R^m$be a cube with edges parallel to the coordinate axes,$p > 1, a > 0$,$[a]$ the integral part of$a, \theta = a - [a]$. We introduce the spaces $L_p(\Delta), L_\infty(\Delta)$, $C(\Delta), C^a(\Delta)$in the usual manner. We use$W_p^a(\Delta)$to denote the Sobolev-Slobodeckn space (see e.g.[135]). The norm in$W_p^a(\Delta)$is defined by

$$\|u\|_{W_p^\alpha(\Delta)} = \|u\|_{L_p(\Delta)} + \|u\|_{L_p^\alpha(\Delta)}. \tag{1}$$

where for integralα

$$\|u\|_{L_p^\alpha(\Delta)}^p = \sum_{|x|=\alpha} \int_\Delta |D^x u|^p \, dx, \tag{2}$$

and for nonintegralα

$$\|u\|_{L_p^\alpha(\Delta)}^p = \sum_{|x|=\alpha} \int_\Delta \int_\Delta \frac{|D^x u(x) - D^x u(y)|^p}{|x - y|^{p\theta + m}} dxdy. \tag{3}$$

Below Q^m is the halfopen unit cube $0 \leq x_i < 1 (i = 1, \dots, m)$ of the space R^m; if $\Delta = Q^m$ the symbol Δ will be omitted in the notation for spaces and their corresponding norms. The seminorm $\|\cdot\|_{L_p^\alpha(\Delta)}$ has the property of homogeneity with respect to similarity transformations of the cube Δ. Indeed, let $\Delta = x_0 + hQ^m$, suppose $u \in W_p^\alpha$ and v is the function defined for $x \in \Delta$ by $v(x) = u\big(h^{-1}(x - x_0)\big)$. Then $v \in W_p^a(\Delta)$ and

$$\|v\|_{L_p^\alpha(\Delta)} = h^{mp^{-1}-\alpha}\|u\|_{L_p^\alpha}. \quad (4)$$

In the one-dimensional case, besides the classes $W_p^a(\Delta)$ we shall also consider classes of functions of bounded β-variation. The function u given on the (possibly infinite) interval $\Delta \subset R_1$ belongs to the class $V_\beta(\Delta)$ of functions of bounded β-variation $(\beta \geq 1)$ if the quantity

$$\|u\|_{V_\beta^o(\Delta)}^p = \sup \sum_{k=1}^{n} |u(x_k) - u(x_{k-1})|^\beta$$

is finite; here the least upper bound is taken relative to all possible finite sets of points $x_0 < x_1 < \cdots \dots < x_n$ from the interval Δ. The class $V_\beta(\Delta)$ is a Banach space relative to the norm

$$\|u\|_{V_\beta(\Delta)} = \|u\|_{V_\beta^o(\Delta)} + \sup_{x \in \Delta}|u(x)|. \quad (5)$$

Every function $u \in V_\beta(\Delta)$ has limits from the left and from the right at each point of the interval Δ. To simplify the exposition we normalize the functions of bounded β-variation, making them continuous from the right, and restrict consideration to classes $V_\beta(\Delta)$ for intervals Δ which are halfopen from the right. (In particular, if the left end of the interval Δ is infinite, then we include the improper end $x = -\infty$ in the interval Δ, using as the value $u(-\infty)$ the corresponding limit.) We note another obvious property of the classes V_β- their invariance with respect to monotonic replacements of the independent variable.

We recall the definition of e-entropy and the n-diameter of a compact set in a normed linear space (see [134], [138]). Let X be a Banach space and suppose the set $K \subset X$ is compact. Let $\mathfrak{N}_\epsilon(K; X)$ be the minimal number of elements of the ϵ-net of the set K in the metric of the space X; the ϵ-entropy of the set K in X is the quantity

$$\mathcal{H}_\varepsilon(K; X) = \log_2 \mathcal{N}_\varepsilon(K; X).$$

The number

$$d_n(K; X) = \log \sup_{x \in K} \min_{y \in L_n} \|x - y\|$$

is called the η-diameter of the set K in X; the inf is taken relative to all possible re-dimensional hyperplanes $L_n \subset X$.

If γ is a Banach space compactly imbedded in X and $K = SY$, then we write $\mathcal{H}_\epsilon(Y; X)$ instead of $\mathcal{H}_\epsilon(SY; X)$ and $d_n(Y; X)$ instead of $d_n(SY; X)$.

We use the following notation for the constants encountered in the various estimates. Constants whose values are not essential are denoted by the letter c with no subscript, and essential constants are denoted by C with subscripts. Also we write $f \asymp g$ if $f \leq cg$ and $g \leq cf$.

Let Ξ be a partition of the cube Q^m into a finite number of halfopen ro-dimensional cubes Δ_k; let $|\Xi|$ be the number of cubes in the partition Ξ. We write $\Xi = \{\Delta_k\} (k = 1, \dots, |\Xi|)$ and $\Delta_k \in \Xi$. A partition Ξ' obtained from a by dividing certain cubes $\Delta_k \in \Xi$ into 2^m different cubes is called an elementary extension of the partition Ξ. Below a basic role is played by the class $\mathfrak{R}$ of all partitions which can be obtained from the trivial partition $\Xi_0(|\Xi_0| = 1)$by a finite number of elementary extensions. The symbol$\Xi_0$will always denote the trivial partition.

Let J be a nonnegative function of half open cubes $\Delta \subset Q^m$, semiadditive from below in the following sense: if the cube $\Delta \subset Q^m$is decomposed into a finite number of disjoint cubes $\Delta_{\bar{j}}$, then$\sum_j J(\Delta_j) \leq J(\Delta)$. Let$|\Delta|$be the Euclidean volume of the cube Δ,$a > 0$ some number. Set

$$g_a(J; \Delta) = |\Delta|^a J(\Delta)(\Delta \subset Q^m)$$

and consider the following function of partitions Ξ of the cube Q^m:

$$G_a(J; \Xi) = \max_{\Delta \in \Xi} g_a(J; \Delta). \tag{6}$$

The basic goal is to estimate the quantity $\min_{|\Xi| \leq n} G_a(J; \Xi)$, depending on n. To obtain such an estimate we construct a special sequence of partitions. This sequence is constructed by the method of "successive division." Indeed, the first step is to divide the cube Q^m into 2^m different cubes. Then we again partition that cube Δ for which the quantity $g_a(J; \Delta)$ is maximal into 2^mcubes. This process is continued and we obtain a sequence of partitions for which we can give a good estimate of the rate of decrease of the quantity (6). The considerations presented below follow basically from this procedure- There is a difference, however, in that we allow simultaneous division of several cubes of one partition.

Thus, let J be a given function semiadditive from below and with it associate a sequence of partitions $\{\Xi_i\}_0^\infty$which is constructed as follows. We start with the trivial partition Ξ_0. Suppose the partition Ξ_i has already been constructed and let $\Delta_j \in \Xi_i (j = 1, \dots, s_i)$ be those cubes of the partition Ξ_i for which

$$g_a(J; \Delta_j) \geq 2^{-ma} G_a(J; \Xi_i)(j = 1, \dots, s_i). \tag{7}$$

Then as Ξ_{i+1} we take the elementary extension of the partition Ξ_i obtained by dividing these cubes. Thus s_i is the number of cubes in the partition Ξ_i which were divided in passing to Ξ_{i+1}. It is clear that $\Xi_i \in \mathfrak{R}(i = 0,1, \dots)$.

The sequence of partitions$\Xi_i$obtained by this construction is denoted as follows:

$$\{\Xi_i\}_0^\infty = T_a(J).$$

For the quantities characterizing the sequence$T_a(J)$, we use the notation

$$n_i = n_i(J; a) = |\Xi_i| (\Xi_i \in T_a(J)), \tag{8}$$

$$\delta_i = \delta_i(J; a) = G_a(J; \Xi_i) = \max_{\Delta \in \Xi_i} |\Delta|^a J(\Delta) \left(\Xi_i \in T_a(J)\right). \tag{9}$$

It is clear that $n_0 = 1$and

$$n_{i+1} \leq 2^m n_i (i = 0,1, \dots). \tag{10}$$

Theorem (5.1.1)[129]: For every function J semiadditive from below and for each natural number$n$ there is a partition $\Xi \in \mathfrak{R}$of the cube Q^msuch that $|\Xi| \leq n$and

$$G_a(J; \Xi) \leq C_1 n^{-(a+1)} J(Q^m), \tag{11}$$

where the constant $C_1 = C_2(a, m)$does not depend on the function$J$.

The validity of Theorem (5.1.1) stems from the following assertion.

Lemma (5.1.2)[129]: Suppose the cube $\Delta \subset Q^m$is divided into 2^m different cubes $\Delta_j (j = 1, \dots, 2^m)$. Then

$$\max_j g_a(J; \Delta_j) \leq 2^{-ma} g_a(J; \Delta).$$

Lemma (5.1.3)[129]: Let s be a natural number and let $x_j > 0$, $\gamma_j > 0 (j = 1, \dots, s)$be numbers satisfying the relationships

$$\sum_{j=1}^{s} x_i \leq 1, \sum_{j=1}^{s} y_i \leq 1, x_i y_i^a \geq b (j = 1, \dots, s).$$

for some $a > 0$, $b > 0$. Then $b \geq s^{-(a+1)}$.

Lemma (5.1.2) is obvious and Lemma (5.1.3) is proved with the help of elementary work with extrema.

Theorem (5.1.4)[129]: For every function J semiadditive from below the quantities $n_i(J;a)$ and $\delta_i(J;a)$are related by the inequality

$$\delta_i \leq C_2 n_i^{-(a+1)} (i = 0,1, \dots), \tag{12}$$

where the constant $C_2 = C_2(a,m)$ does not depend on J.

Proof. Without loss of generality we assume $J(Q^m) \leq 1$. We investigate certain properties of the sequences(8) and(9). It follows from Lemma(5.1.2) and the inequality(7) that

$$\delta_{i+1} \leq 2^{-ma} \delta_i (i = 0,1, \dots), \tag{13}$$

Another inequality for the quantities δ_i follows from Lemma (5.1.3). Namely, setting $x_j = J(\Delta_j), y_j = |\Delta_j| \left(j = 1, \dots, s_j\right)$and taking account of (7), we find that the conditions of the lemma are satisfied for $b = 2^{-ma}\delta_i$. Hence

$$\delta_i \leq 2^{ma} s_i^{-(a+1)} \quad (i = 0,1, \dots) \tag{14}$$

We note the obvious relationships

$$n_0 = 1, \;\; s_i \leq n_i, n_{i+1} - n_i = (2^m - 1)s_i,$$
$$n_i \leq 2^m \sum_{j=0}^{i-1} s_j \qquad (i = 1, 2, \dots). \tag{15}$$

Let $k \geq i \geq 0$; from (13) and (14) we obtain that

$$\delta_k \leq 2^{-(k-i-1)ma} s_i^{-(a+1)}$$

Hence for every i $(0 \leq i \leq k)$

$$s_i \leq 2^{-(k-i-1)ma(a+1)^{-1}} \delta_k^{-(a+1)^{-1}} \tag{16}$$

Further, for $k \geq 1$, taking account of (15) and (16), we find that

$$n_k \leq 2^m \delta_k^{-(a+1)^{-1}} \sum_{i=0}^{k-1} 2^{-(k-i-1)ma(a+1)^{-1}} = 2^m \delta_k^{-(a+1)^{-1}} \sum_{j=0}^{k-1} 2^{-jma(a+1)^{-1}}$$
$$< 2^m \left[1 - 2^{-ma(a+1)^{-1}}\right]^{-1} \delta_k^{-(a+1)^{-1}}.$$

Thus for $k \geq 1$

$$\delta_k \leq C_2 n_k^{-(a+1)}, \tag{17}$$

where the constant

$$C_2 = 2^{m(a+1)} \left[1 - 2^{-ma(a+1)^{-1}}\right]^{-(a+1)}$$

does not depend on J. It is also obvious that (17) holds for $k = 0$ too. Thus Theorem (5.1.4)has been proved.

The considerations are of an "entropic" nature. We let $\mathfrak{J}$ denote the set of all functions J semiadditive from below which satisfy the condition $J(Q^m) \leq 1$. We combine the functions $J \in \mathfrak{J}$ which are close in a certain sense into a class and estimate the number of such classes. Together with the sequences (8), (9) we also consider the sequence of numbers:

$$\tilde{\delta}_i = \tilde{\delta}_i(J;a) = C_2 \min_{0 \leq j \leq i} \left[2^{-am(i-j)} n_j^{-(a+1)}\right] \quad (i = 0, 1, \dots). \tag{18}$$

It follows from (13) and (12) that

$$\delta_i \leq \tilde{\delta}_i \qquad (i = 0,1, \dots). \tag{19}$$

It is also clear that

$$\tilde{\delta}_i \leq C_2 n_j^{-(a+1)} \quad (i = 0, 1, \dots). \tag{20}$$

Thus the sequence (18) majorizes (9) and satisfies an inequality analogous to (12). Together with the sequence $\{\tilde{\delta}_i\}$ behaves more regularly than the sequence $\{\delta_i\}$: the following inequalities hold for it:

$$2^{-(a+1)m}\tilde{\delta}_i \leq \tilde{\delta}_{i+1} \leq 2^{-am}\tilde{\delta}_i \,. \tag{21}$$

Indeed,

$$\tilde{\delta}_{i+1} = C_2 \min_{0\leq j\leq i+1} \left[2^{-am(i-j+1)} n_j^{-(a+1)}\right] = \min\left[2^{-am}\tilde{\delta}_i; C_2 n_{i+1}^{-(a+1)}\right].$$

The right inequality in (21) now follows immediately; it remains to refer to (10) and (20) to derive the left inequality.

Now let η be a fixed number $(0 < \eta \leq C_2)$. Let T_a^η denote the interval $\{\Xi_i\}_0^k$ of the sequence $T_a(J)$, where the number k is determined by the conditions

$$\tilde{\delta}_k < \eta \leq \tilde{\delta}_{k-1}. \tag{22}$$

We shall assume that the functions $J, J' \in \mathfrak{J}$ belong to the same class if and only if

$$T_a^\eta(J) = T_a^\eta(J').$$

The number of classes into which the set $\mathfrak{J}$ can be separated here is denoted by $N(a;\, \eta)$.

Lemma (5.1.5)[129]: The estimate

$$\log_2 N(a;\, \eta) \leq C_3 \eta^{-(a+1)^{-1}}, \qquad C_3 = C_3(a, m). \tag{23}$$

holds for all values of $\eta, 0 < \eta \leq C_2$.

Proof: Let $\{\Xi_i\}_0^k$ be a finite sequence of partitions such that $\{\Xi_i\}_0^k = T_a^\eta(J)$ for at least one function $J \in \mathfrak{J}$ .Then according to (10), (20) and (22) we obtain

$$n_k \leq 2^m\, n_{k-1} \leq 2^m \left(C_2 \tilde{\delta}_{k-1}^{-1}\right)^{(a+1)^{-1}} \leq (C_2 \eta^{-1})^{(a+1)^{-1}}. \tag{24}$$

First of all we estimate the number of different sequences $n_i = |\Xi_i|$ $(i = 1, \dots, k)$ whose last terms satisfy (24). We use n^* to denote the integral part of the number $2^m (C_2\eta^{-1})^{(a+1)^{-1}}$. Writing n^* in the form

$$n^* = 1 + \sum_{i=1}^{k} (n_i - n_{i-1}) + (n^* - n_k),$$

we see that the number of such sequences does not exceed the number of representations of the number $n^* - 1$ in the form of a sum of positive integral terms; here representations which differ in the order of terms are considered different. The number of such representations is equal to 2^{n^*-2} (see for example [139]).

Let $\{n_i\}_0^k$ $(n_k \leq n^*)$ be a fixed sequence of the form under consideration. We estimate the number of all possible sequences of partitions $\{\Xi_i\}_0^k = T_a^\eta(J)$ for which $|\Xi_i| = n_i$. For this we note that if the partition Ξ_i $(i = 0, 1, \dots, k-1)$ is already fixed, then the partition Ξ_{i+1} is uniquely determined by which $s_i = (n_{i+1} - n_i)(2^m - 1)^{-1}$ of the cubes of the partition Ξ_i (from the overall number of cubes n_i) are decomposed in passing to Ξ_{i+1}. The number of possible variants here is equal to $\binom{n_i}{s_i} < 2^{n_i}$. Hence the number of all sequences of partitions of the form under consideration with a fixed sequence of numbers $\{n_i\}_0^k$ is less than

$$2^{n_0 + n_1 + \cdots + n_{k-1}}.$$

We note that from the definition (18) of the number $\tilde{\delta}_{k-1}$ we obtain the inequalities

$$n_i \leq \left(C_2\tilde{\delta}_{k-1}^{-1}\right)^{(a+1)^{-1}} 2^{-(k-1-i)ma(a+1)^{-1}} \quad (i = 0,1,\dots,k-1).$$

Hence we find that

$$\sum_{i=0}^{k-1} n_i \leq \left(C_2\tilde{\delta}_{k-1}^{-1}\right)^{(a+1)^{-1}} \sum_{i=0}^{k-1} 2^{-ima(a+1)^{-1}} < C_3'\tilde{\delta}_{k-1}^{-(a+1)^{-1}}, \tag{25}$$

where $C_3' = C_2^{(a+1)^{-1}}\left[1-2^{-ma(a+1)^{-1}}\right]^{-1}$ Combining the estimates and taking account of (22), we find that

$$\log_2 N(a;\,\eta) \leq n^* - 2 + C_3'\tilde{\delta}_{k-1}^{-(a+1)^{-1}} < \left[C_3' + 2^m C_2^{(a+1)^{-1}}\right]\eta^{-(a+1)^{-1}}.$$

Thus inequality (23) is obtained for $C_3 = C_3' + 2^m C_2^{(a+1)^{-1}}$. The lemma is proved.

We turn to a discussion of Theorem (5.1.1). The condition $a > 0$ in it is essential; in fact if, for example, J is a point load type of function, then the estimate (11) is not true for $a = 0$. In the one- dimensional case, a modification of Theorem (5.1.1), valid also for $a = 0$, is possible. This modification will be needed in studying functions of bounded β-variation. Here, however, we cannot even restrict ourselves to partitions of the class $\mathfrak{R}$. As a result of this the considerations.

On the other hand, failure of the condition $\Xi \in \mathfrak{R}$ leads to an improvement of the constant in (11): an analogous estimate holds for $C_1 = 1$ and is attained for the function $J(\Delta) = |\Delta|$.We note that for $m = 1$ we cannot obtain the estimate (11) with $C_1 = 1$ even by passing to a wider class of partitions.

Thus, let $m = 1$ and $Q^1 = [0,1)$. We write $J[x',x'')$ instead of $J([x',x''))$ for every interval $[x',x'') \subset Q^1$. In view of the condition of semiadditivity, the function $\psi(t) = J[t,x'')$ does not increase on $(x\,,x'')$, is bounded, and consequently has a finite limit as $t \to x'+0$ which we denote by $\tilde{J}[x',x'')$. Obviously $\tilde{J}[x',x'') \leq J[x',x'')$.

Theorem (5.1.6)[129]: Suppose the nonnegative function J, semiadditive from below, of half op en intervals $\Delta \subset Q^1$ is continuous from the left:

$$J[x',t) \to J[x',x'') \quad as \quad t \to x''-0.$$

For every such function and arbitrary $a \geq 0$ for each natural number η there is a partition Ξ of the interval Q^1 such that $|\Xi| \leq n$ and

$$G_a(\tilde{J};\Xi) \leq n^{-(a+1)}J(Q^1). \tag{26}$$

The proof is by induction, assuming $J[0,1) = 1$. For $n = 1$ the inequality (26) is obvious. Suppose the assertion of the theorem is true for some $n \geq 1$; we shall show that then it also holds for $n+1$. First of all we note that if we take $[0,x_0)$ for the basic interval, then (26) becomes

$$\tilde{G}_a(\tilde{J};\Xi) \leq n^{-(a+1)}x_0^a J[0,x_0).$$

We introduce the notation $\phi(x) = J[0,x)$ and consider the function

$$\varphi(x) - \left(\frac{n}{n+1}\right)^{a+1} x^{-a}.$$

Since this function is continuous from the left, does not increase and changes sign in the interval $(0,1)$, there is a point $x_0 \in (0,1)$ such that

$$\varphi(x_0) \leq \left(\frac{n}{n+1}\right)^{a+1} x_0^{-a} \leq \varphi(x_0+0).$$

In accord with the induction hypothesis, we can divide the interval $[0, x_0)$ into halfopen intervals $\Delta_1, \dots, \Delta_k (k \le n)$ so that

$$\max_{i=1,..,k} g_a(\tilde{J}; \Delta_i) \le n^{-(a+1)} x_0^a \varphi(x_0) \le (n+1)^{-(a+1)}. \quad (27)$$

Further, from the inequality

$$\phi(x) + J[x, 1) \le 1,$$

passing to the limit as $x \to x_0 + 0$, we find that

$$\varphi(x_0 + 0) + \tilde{J}[x_0, 1) \le 1.$$

Hence

$$\tilde{J}[x_0, 1) \le 1 - \varphi(x_0 + 0) \le 1 - \left(\frac{n}{n+1}\right)^{a+1} x_0^{-a}.$$

It is an elementary matter to verify that for $x_0 \in (0, 1)$ the right side of the last inequality does not exceed $(n+1)^{-(a+1)}(1 - x_0)^{-a}$, and consequently

$$(1 - x_0)^{-a} \tilde{J}[x_0, 1) \le (n+1)^{-(a+1)}. \quad (28)$$

The inequalities (27) and (28) show that the partition of the interval $[0, 1)$ into intervals $\Delta_1, \dots, \Delta_k, [x_0, 1)$ is the desired one. Thus the induction has been verified and the theorem proved.

We note that under the conditions of the theorem it is possible to relax the requirement that the function J be continuous from the left. An inequality of the form (26) remains valid but with a factor $c > 1$ on the right-hand side.

The basic difference between the inequalities (26) and (11) is that the function J is replaced by J on the left-hand side. As an example of a point load type function shows, for $a = 0$ this is essential.

We investigate the rate of approximation of functions of the class W_p^a by piecewise polynomial functions. The degree l of the approximating polynomials is fixed, with $l = a - 1$ for integral a and $l = [a]$ for nonintegral a. Below we use the notation $\omega = am^{-1}$ and, when $p\omega \le 1, q^* = p(1 - \rho\omega)^{-1}$. Here q^* is the so-called limit exponent in the theorem of imbedding of the space W_p^a into the space L_q. As usual, we set $q^* = \infty$ when $\rho\omega = 1$.

Let $\Delta \subset Q^m$ be a cube. With every function $u \in W_p^a(\Delta)$ we associate a polynomial r of degree l satisfying the conditions

$$\int_\Delta x^x r(x) dx = \int_\Delta x^x u(x) dx \quad (|x| \le l). \quad (29)$$

Conditions (29) obviously determine r uniquely. Set $r = P_\Delta u$; thus P_Δ is a linear projection operator mapping the space $W_p^a(\Delta)$ onto the finite-dimensional space of polynomials of degree l in m variables. The dimensionality of this space is denoted by $\nu = v(m, l)$.

We note the following simple assertions.

Lemma (5.1.7)[129]: When $\rho\omega > 1$ for every function $u \in W_p^a(\Delta)$ the following inequality is satisfied:

$$\|u - P_\Delta u\|_{C(\Delta)} \le C_4 |\Delta|^{\omega - p^{-1}} \|u\|_{L_p^a(\Delta)}, \quad (30)$$

where the constant $C_4 = C_4(p, a, m)$ does not depend on Δ.

Lemma (5.1.8)[129]: When $\rho\omega \le 1$ and $q < q^*$ for every function $u \in W_p^a(\Delta)$ the following inequality is satisfied:

$$\|u - P_\Delta u\|_{L_q(\Delta)} \leq C_5 |\Delta|^{q^{-1}-q^{*-1}} \|u\|_{L_p^a(\Delta)}, \quad (31)$$

where the constant $C_5 = C_5(p, q, a, m)$ does not depend on Δ.

For the proof of both assertions we first consider the case in which $\Delta = Q^m$. We introduce a new norming in the space W_p^a:

$$\|u\|_{W_p^a} = \sum_{|x|\leq l} \left| \int_{Q^m} x^x u(x) dx \right| + \|u\|_{L_p^a}.$$

Equivalence of the norms $|||\cdot|||_{W_p^a}$ and $\|\cdot\|_{W_p^a}$ follows from considerations of S. L. Sobolev [140] for integral a and can be proved quite analogously for nonintegral a. It follows from conditions (29) that

$$\left|\left|\left|u - P_{Q^m} u\right|\right|\right|_{W_p^a} = \|u\|_{L_p^a}.$$

The theorem of imbedding of the space W_p^a into the space C (for $\rho\omega > 1$) and into L_q (for $\rho\omega \leq 1$) shows that the inequality (30) or, respectively, (31), is satisfied in the cube $\Delta = Q^m$. To pass to an arbitrary cube Δ we need to implement the similarity transformation and for this use the property of homogeneity (1.4) of the seminorm $\|\cdot\|_{L_p^a(\Delta)}$ .Thus Lemmas (5.1.7) and (5.1.8) are proved.

Let Ξ be a partition of the (halfopen) cube Q^m into halfopen cubes. We use $\mathcal{P}(\Xi;\ l)$ to denote the linear set of all functions whose restriction to each of the cubes $\Delta \in \Xi$ is a polynomial of degree l. We introduce the projection operator P_Ξ defined as follows: $v = P_\Xi u$ is the function of the class $\mathcal{P}(\Xi;\ l)$ coinciding with the polynomial $P_\Delta u$ on each cube $\Delta \in \Xi$.

Theorem (5.1.9)[129]: For every function $u \in W_p^a(\rho\omega > 1)$ and for every natural n there is a partition $\Xi \in \mathfrak{R}$ of the cube Q^m such that $|\Xi| \leq n$ and

$$|||u - P_\Xi u|||_{L_\infty} \leq C_6 n^{-\omega} \|u\|_{L_p^\alpha}, C_6 = C_6(p, \alpha, m). \quad (32)$$

Proof: Let Ξ be an arbitrary partition of Q^m into cubes and $v = P_\Xi u$. Then, according to (30),

$$\sup_{x \in Q^m} |u(x) - v(x)| \leq C_4 \left[\max_{\Delta \in \Xi} |\Delta|^{p\omega - 1} \|u\|_{L_p^a(\Delta)}^p\right]^{p^{-1}}. \quad (33)$$

We consider the following function $J_u(\Delta)$ of cubes $\Delta \subset Q^m$:

$$J_u(\Delta) = \|u\|_{L_p^a(\Delta)}^p. \quad (34)$$

The function J_u is semiadditive from below and additive [1)] for integral a.

In the square brackets on the right-hand side of (33) we have the partition function $G_{p\omega-1}(J_u; \Xi)$ constructed from the function (34). By Theorem (5.1.1) there exists a partition $\Xi \in \mathfrak{R}$ of the cube Q^m for which $|\Xi| \leq n$ and

$$G_{p\omega-1}(J_u; \Xi) \leq C_1 n^{-p\omega} J_u(Q^m).$$

The last inequality together with (33) leads to the estimate (34) with constant $C_6 = C_4 C_1^{p-1}$. The theorem is proved.

Theorem (5.1.10)[129]: Let $\rho\omega \leq 1, q < q^*$. For every function $u \in W_p^a$ and every natural number n there is a partition $\Xi \in \mathfrak{R}$ of the cube Q^m such that $|\Xi| \leq n$ and

$$|||u - P_\Xi u|||_{L_q} \leq C_7 n^{-\omega} \|u\|_{L_p^\alpha}, \quad C_7 = C_7(p, q, \alpha, m). \quad (35)$$

Proof. Consider the partition function $G_a(J_u; \Xi)$ with $a = p(q^{-1} - q^{*-1})$. According to Theorem (5.1.1) there is a partition $\Xi \in \mathfrak{R}$ for which $|\Xi| \leq n$ and

$$G_a(J_u;\ \Xi) = \max_{\Delta\in\ \Xi}|\Delta|^{p(q^{-1}-q^{*-1})}\|u\|^p_{L^\alpha_{p(\Delta)}} \le\ C_1 n^{-p(\omega+q^{-1})}\|u\|^p_{L^\alpha_p}.$$

Taking Lemma (5.1.8) into account for this partition Ξ and the function $v = P_\Xi u$, we find that

$$\|u-v\|^p_{L^\alpha_p} = \sum_{\Delta\in\Xi}\|u-v\|^p_{L_q(\Delta)} \le\ |\Xi|C_5^q[G_a(J_u;\ \Xi)]^{qp^{-1}}$$
$$\le |\Xi|C_5^q C_1^{qp^{-1}} n^{-(q\omega+1)}\|u\|^p_{L^\alpha_p} \le \left[C_7 n^{-\omega}\|u\|_{L^\alpha_p}\right]^q,$$

where $C_7 =\ C_1^{p-1}C_5$. The theorem is proved.

Now let σ be a finite Borel measure defined on subsets of Q^m. We consider approximation of functions from W^α_p in the metric of $L_{q,\sigma} = L_q(Q^m;\ \sigma), q \ge 1$. In contrast to Theorem (5.1.9) and (5.1.10), we are here concerned with the rate of approximation which can be attained by choosing partitions not depending on the function u (but depending, generally speaking, on the measure σ).

Let $\mathfrak{M}$ be the set of all finite Borel measures on Q^m satisfying the condition $\sigma(Q^m) \le 1$. By $\mathfrak{M}_\lambda(1 \le \lambda < \infty)$ we denote the set of all absolutely continuous measures σ on Q^m whose density $d\sigma/dx$ belongs to the space L_λ and satisfies the condition

$$\int_{Q^m}\left(\frac{d\sigma}{dx}\right)^\lambda dx \le 1 \qquad (36)$$

Theorem (5.1.11)[129]: Let $p\omega > 1$. For every Borel measure $\sigma \in \mathfrak{M}$ and every natural n there exists a partition $\Xi \in \mathfrak{R}$ of the cube Q^m such that following inequality is satisfied:

$$\|u - P_\Xi u\|_{L_{q,\sigma}} \le\ C_8 n^{-\gamma}\|u\|_{L^\alpha_p}, \qquad C_8 = C_8(p,q,\alpha,m), \qquad (37)$$

where $\gamma =\ \omega$ when $p \ge q$ and $\gamma = \omega - p^{-1} + q^{-1}$ when $p < q$. The consent C_8 does not depend on σ.

Proof: It is suffices to give the reasoning for the case $q \ge p$ since the validity of the assertion of the theorem for $q < p$ follows in an obvious way from the assertion for $q = p$.

Thus let Ξ be an arbitrary partition of Q^m into (half open) cubes). For every function $u \in W^\alpha_p$ and the function $v = P_\Xi u$ by Lemma (5.1.7) we have

$$\|u-v\|^q_{L_{q,\sigma}} \le \sum_{\Delta\in\Xi}\sup_{x\in\Delta}|u-v|^q\sigma(\Delta) \le C_4^q\sum_{\Delta\in\Xi}|\Delta|^{(\omega-p^{-1})q}\,\|u\|^q_{L^\alpha_p(\Delta)}\sigma(\Delta).$$

When $p \le q$ we have

$$\sum_{\Delta\in\Xi}\|u\|^q_{L^\alpha_p(\Delta)} \le \left[\sum_{\Delta\in\Xi}\|u\|^q_{L^\alpha_p(\Delta)}\right]^{qp^{-1}} \le \|u\|^q_{L^\alpha_p}. \qquad (38)$$

Consequently

$$\|u-v\|^q_{L_{q,\sigma}} \le\ C_4^q\|u\|^q_{L^\alpha_p}\cdot\max_{\Delta\in\ \Xi}\{|\Delta|^{(\omega-p^{-1})q}\sigma(\Delta)\}.$$

Now we apply Theorem (5.1.1) to the function $G_a(\sigma;\ \Xi)$ with $a = (\omega - p^{-1})q$, which leads to (37). The theorem is proved.

The following theorem extended our result to the case $p\omega \le 1$. Naturally here inequality (37) may not be true for arbitrary measures. It remains valid, however, for absolutely continuous measures whose density is summable when raised to a sufficiently high degree.

Theorem (5.1.12)[129]: Let $p\omega \le 1$ and

$$\lambda^{-1} + qq^{*-1} < 1. \qquad (39)$$

Then for every measure $\sigma \in \mathfrak{M}_\lambda$ and every natural n there exists a partition $\Xi \in \mathfrak{R}$ of the cube Q^m such that $|\Xi| \le n$ and for every function $u \in W_p^\alpha$ the inequality (37) is satisfied with some constant $C_8 = C_8'(p, q, \lambda, \alpha, m)$ not depending on σ.

Proof: For every cube $\Delta \subset Q^m$ and for $v = P_\Delta u$ we find, by Lemma (5.1.8), that

$$\int_\Delta |u - v|^q d\sigma = \int_\Delta |u - v|^q \frac{d\sigma}{dx} \le \left\{\int_\Delta |u - v|^{q(1-\lambda^{-1})^{-1}} dx\right\}^{1-\lambda^{-1}} \left\{\int_\Delta \left(\frac{d\sigma}{dx}\right)^\lambda dx\right\}^{\lambda-1}$$

$$\le c|\Delta|^{1-\lambda^{-1}-qq^{*-1}} \|u\|^q_{L_p^\alpha(\Delta)} \left\{\int_\Delta \left(\frac{d\sigma}{dx}\right)^\lambda dx\right\}^{\lambda-1}.$$

Hence for every partition Ξ of Q^m into cubes, for an arbitrary function $u \in W_p^\alpha$ and for the function $v = P_\Xi u$ we obtain

$$\|u - v\|^q_{L_{q,\sigma}} \le c \sum_{\Delta \in \Xi} |\Delta|^{1-\lambda^{-1}-qq^{*-1}} \|u\|^q_{L_p^\alpha(\Delta)} \left\{\int_\Delta \left(\frac{d\sigma}{dx}\right)^\lambda dx\right\}^{\lambda-1}. \tag{40}$$

Now assume that $q \le p$. We note for $q \le p$

$$\sum_{\Delta \in \Xi} \|u\|^q_{L_p^\alpha(\Delta)} \le |\Xi|^{1-qp^{-1}} \left\{\sum_{\Delta \in \Xi} \|u\|^q_{L_p^\alpha(\Delta)}\right\}^{qp^{-1}} \le |\Xi|^{1-qp^{-1}} \|u\|^q_{L_p^\alpha}. \tag{41}$$

Together with (40) this leads to the inequality

$$\|u - v\|^q_{L_{q,\sigma}} \le c\, |\Xi|^{1-qp^{-1}} \|u\|^q_{L_p^\alpha} \max_{\Delta \in \Xi} \left\{|\Delta|^{1-\lambda^{-1}-qq^{*-1}} \int_\Delta \left(\frac{d\sigma}{dx}\right)^\lambda dx\right\}^{\lambda-1} \tag{42}$$

This choice of partition Ξ is made according to Theorem (42) leads directly to (37) if we take into account (36) and the fact that $|\Xi| \le n$.

The case $q > p$ is considered analogously. The inequality (38) is to be used in place of (41). the theorem is proved.

We make several remarks.

Remark (5.1.13)[129]: If under conditions of Theorems (5.1.11) and (5.1.12) we allow independence of partition Ξ from the function u, then we can replace (37) by the inequality

$$\|u - P_\Xi u\|_{L_{q,\sigma}} \le cn^{-\omega} \|u\|_{L_p^\alpha}. \tag{43}$$

Under the conditions of Theorem (5.1.12) this follows directly from (32); under the condition of Theorem (5.1.12) it can be derived easily from (35). A comparison of (43) and (37) shows that passing to a method of partitions which does not depend on the function u yields a worse resut only for $q > p$.

Theorems on approximation of functions of the classes $V_\beta (\beta \ge 1)$ (Theorems (5.1.15) – (5.1.17)) are proved by the same procedure, on the basis of Theorem (5.1.15). We note that the first two theorems (Theorems (5.1.15), (5.1.16)) are easy to prove directly without using theorems on partition functions.

We restrict consideration to functions $u \in V_\beta$ which are normalized to continuity from the right, and we assume the basis interval (denoted by X) to be halfopen.

With every function $u \in V_\beta(X)$ we associate a function I_u of halfopen intervals $\Delta \in X$ by the rule

$$I_u(\Delta) = \|u\|^\beta_{\overset{\circ}{V}_{\beta(\Delta)}}. \tag{44}$$

This function plays the same role in the investigation of the classes V_β as the function (34) for the classes W_p^α. The function (44) is obviously semiadditive from below. In view of the assumed normalization of the function $u \in V_\beta$ the following assertion is also true.

Lemma (5.1.14)[129]: Suppose the function $u \in V_\beta(X)$ is continuous from the right and I_u is continuous from the left, and the function $\tilde{I}_u$ generated by it according to the definition coincides with I_u.

Proof: Let $\Delta = [x', x'') \subset X$. For a given $\epsilon > 0$ there is a system of points $x' \le x_0 < x_1 < \cdots < x_n < x''$ for which

$$\sum_{k=1}^{n} |u(x_k) - u(x_{k-1})|^\beta > I_u(\Delta) - \epsilon. \tag{45}$$

In view of the continuity from the right of the function u, we can assume that $x_0 > x'$ in (45). For every $x \in (x_m, x'')$ it follows from (45) that

$$I_u(\Delta) - \varepsilon < \|u\|^\beta_{\overset{\circ}{V}_\beta([x_0, x_n])} \le I_u[x', x) \le I_u(\Delta).$$

Analogously for $x \in (x', x_0]$ we find that

$$I_u(\Delta) - \varepsilon < \|u\|^\beta_{\overset{\circ}{V}_\beta([x_0, x_n])} \le I_u[x, x'') \le I_u(\Delta).$$

Both assertions of the lemma follows from these inequalities.

1. Let $\Delta = [x', x'') \subset X$ be an interval. The role of inequality (30) for the functions $u \in V_\beta(X)$ is played by the obvious inequality

$$\sup_{x \in \Delta} |u(x) - u(x')| \le \|u\|_{\overset{\circ}{V}_\beta(\Delta)}. \tag{46}$$

Let Ξ be a partition of the basis interval X into intervals $\Delta_i = [x_{i-1}, x_i)$ $(i = 1, \dots, |\Xi|)$. Let P_Ξ denote the operator which associates with the function $u \in V_\beta(X)$ the piecewise-constant function $v \in \mathcal{P}(\Xi; 0)$ assuming the constant value equal to $u(x_{i-1})$ in each interval $\Delta_i \in \Xi$.

Theorem (5.1.15)[129]: Let $X = [b', b'')$ be a finite of infinite interval. For every function $u \in V_\beta(X)$ continuous from the right and every natural number n there is a partition Ξ of the interval X into half open intervals such that $|\Xi| \le n$ and

$$\|u - P_\Xi u\|_{L_\infty(\Delta)} \le n^{-\beta^{-1}} \|u\|_{\overset{\circ}{V}_\beta(L)}.$$

Proof: As a result of the invariance of the classes V_β with respect to monotonic replacement of the independent variable, we assume that $X = Q^1 = [0, 1)$. Let Ξ be a partition of the interval Q^1 into halfopen intervals. From inequality (46) for every function $u \in V_\beta$ and the function $v = P_\Xi u$ we obtain that

$$\sup_{x \in X} |u(x) - v(x)| \le \left[\max_{\Delta \in \Xi} I_u(\Delta)\right]^{\beta - 1} = [G_0(I_u; \Xi)]^{\beta - 1}. \tag{47}$$

Lemma (5.1.14) shows that Theorem (5.1.6) applies to the function I_u. Indeed, there exists a partition Ξ of the interval Q^1 into halfopen intervals such that $|\Xi| \le n$ and

$$G_0(\tilde{I}_u; \Xi) = G_0(I_u; \Xi) \le n^{-1} I_u(Q^1).$$

The last inequality together with (47) proves the theorem.

Now let σ be a Borel measure defined on subsets of the interval X and satisfying the condition $\sigma(X) \le 1$. If the measure σ is considered on halfopen intervals $\Delta = [x', x'') \subset X$, then in view of the complete additivity of the measure, the function $\sigma(\Delta) = \sigma[x', x'')$ is continuous from the left. It is also clear that the corresponding function $\tilde{\sigma}(\Delta)$ coincides with the measure of the interval (x', x''), i.e. $\tilde{\sigma}[x', x'') = \sigma(x', x'')$.

In the following theorem we estimate the rate of approximation of functions of the class $V_\beta(X)$ by piecewise-constant functions in the metric of the space $L_{q,\sigma} = L_q(X;\, \sigma), q \geq 1$.

Theorem (5.1.16)[129]: Let $X = [b', b'')$ be a finite or infinite interval. For every Borel measure σ $(\sigma(X) \leq 1)$ and every natural number n there exists a partition Ξ of the interval X into halfopen intervals such that $|\Xi| \leq n$ and for every function $u \in V_\beta(X)$ continuous from the right we have

$$\|u - P_\Xi u\|_{L_{q,\sigma}} \leq\ n^{-****}\|u\|_{\overset{\circ}{V}_\beta(\mathrm{X})}, **** = \min(\beta^{-1}, q^{-1}).$$

Proof: As in the proof of the preceding theorem, it suffices to consider the case $X = Q^1$. Let Ξ be some partition of the interval Q^1. For every function $u \in V_\beta$ and the function $v = P_\Xi u$ form (46) we find that

$$\|u - v\|_{L_{q,\sigma}}^q \leq \sum_{\Delta \in \Xi} \sup_{\mathrm{x}\,\in\,\Delta} |u(x) - v(x)|^q \tilde{\sigma}(\Delta) \leq \sum_{\Delta \in \Xi} \|u\|_{\overset{\circ}{V}_{\beta(\Delta)}}^q\ \tilde{\sigma}(\Delta). \qquad (48)$$

(We note that in (48) it was possible to replace $\sigma(\Delta)$ by $\tilde{\sigma}(\Delta)$ because the function $u - v$ vanishes at the left end interval $\Delta \in \Xi$.) From (48) we find that when $q \leq\ \beta$

$$\|u - v\|_{L_{q,\sigma}}^q \leq \left[\max_{\Delta \in \Xi} \tilde{\sigma}(\Delta)\right] |\Xi|^{1 - q\beta^{-1}} \left[\sum_{\Delta \in \Xi} \|u\|_{\overset{\circ}{V}_{\beta(\Delta)}}^\beta\right]^{q\beta^{-1}}$$

$$\leq \left[\max_{\Delta \in \Xi} \tilde{\sigma}(\Delta)\right]\ |\Xi|^{1 - q\beta^{-1}} \|u\|_{\overset{\circ}{V}_{\beta(\Delta)}}^q$$

And when $q > \beta$

$$\|u - v\|_{L_{q,\sigma}}^q \leq \left[\max_{\Delta \in \Xi} \tilde{\sigma}(\Delta)\right] \left[\sum_{\Delta \in \Xi} \|u\|_{\overset{\circ}{V}_{\beta(\Delta)}}^\beta\right]^{q\beta^{-1}} \leq\ \left[\max_{\Delta \in \Xi} \tilde{\sigma}(\Delta)\right] \|u\|_{\overset{\circ}{V}_{\beta(\Delta)}}^q.$$

Now we have only to note that by Theorem (5.1.6) applied (with $a = 0$) to the function $J = \sigma$, there is a partition $\mathit{\Xi}(|\mathit{\Xi}| \leq n)$ such that

$$\max_{\Delta \in \Xi} \tilde{\sigma}(\Delta) \leq\ n^{-1}.$$

The theorem is proved.

For the function $u \in V_\beta(X) \cap C^\mu(X)$ the assertion of Theorem (5.1.16) can be somewhat strengthened in the case where the interval X is finite. To simplify the statement we assume that $X = Q^1$. As a result of the obvious imbedding $C^\mu \subset V_\beta$ for $\beta = \max(1, \mu^{-1})$, it makes sense to consider the given problem only under the condition $0 < \mu < \beta^{-1}$.

Theorem (5.1.17)[129]: Suppose $X = Q^1$ and the exponents μ, β and q satisfy the condition $1 \leq \beta < q, 0 < \mu < \beta^{-1}$. For every Borel measure $\sigma(\sigma(Q^1) \leq 1)$ and every natural number n there is a partition Ξ of the interval Q^1 into halfopen intervals so that $|\Xi| \leq n$ and for every function $u \in V_\beta \cap C^\mu$ we have

$$\|u - P_\Xi\|_{L_{q,\sigma}} \leq n^{-\rho} L^{1 - \beta q^{-1}} \|u\|_{\overset{\circ}{V}_\beta}^{\beta q^{-1}}, \rho = \mu(1 - \beta q^{-1}) + q^{-1},$$

where L is the Hölder constant of the function u.

Proof: Let Ξ be a partition of the interval Q^1 into intervals $\Delta_k = [x_{k-1}, x_k), 0 = x_0 < x_1 < \cdots < x_{|\Xi|} = 1$. We have

$$\|u - P_\Xi u\|_{L_{q,\sigma}}^q \le \sum_{k=1}^{|\Xi|} \sup_{x \in \Delta_k} |u(x) - u(x_{k-1})|^q \tilde{\sigma}(\Delta_k)$$

$$\le \sum_{k=1}^{|\Xi|} \sup_{x \in \Delta_k} |u(x) - u(x_{k-1})|^\beta L^{q-\beta} |\Delta_k|^{\mu(q-\beta)} \tilde{\sigma}(\Delta_k) \le$$

$$\le L^{q-\beta} \sum_{k=1}^{|\Xi|} \|u\|_{\overset{\circ}{V}_{\beta(\Delta)}}^\beta |\Delta_k|^{\mu(q-\beta)} \tilde{\sigma}(\Delta_k) \le L^{q-\beta} \left[\max_k |\Delta_k|^{\mu(q-\beta)} \tilde{\sigma}(\Delta_k)\right] \cdot \|u\|_{\overset{\circ}{V}_\beta}^\beta .$$

Now it remains to choose the partition Ξ according to Theorem (5.1.6), which is to be applied to the function $J = \sigma$ with $a = \mu(q - \beta)$. The theorem is proved.

The class of function $\mathcal{P}(\Xi; l)$ is a linear set of dimentionality $|\Xi| \cdot v$. As a result of this, Theorem (5.1.11) and (5.1.12) can be interpreted in terms of n-diameters. Indeed, we have the following assertion.

Theorem (5.1.18)[129]: Under the conditions of Theorem (5.1.11) or (5.1.12) the n-diameters d_n of the set SW_p^a in the metric of the space $L_{q,\sigma}$ satisfy the inequality

$$d_n(W_p^a; L_{q,\sigma}) \le cn^{-\gamma}, \tag{49}$$

Where the exponent γ is the same as in (37) and the constant c does not depend on σ.

We also note that in Theorem (5.1.11) and (5.1.12) a linear approximation operator (the operator P_Ξ) is constructed for which (49) is realized. Thus (49) is actually valid for linear n-diameters of SW_p^a in the metric of $L_{q,\sigma}$.

All of the above also holds for the classes H_p^a and B_p^a.

The approximation method used for the proof of Theorems (5.1.9) and (5.1.10) is different: there the partition Ξ depends on the function being approximated, and consequently the class of functions used for approximation is nonlinear. As a consequence of this C or L_q. However, an analysis of the method of proof of these two theorems allows us to estimate another metric characteristic of the set SW_p^a − its ϵ-entropy.

Theorem (5.1.19)[129]: For the ϵ-entropy of the set SW_p^a in the metric of L_q we have (for sufficiently small $\epsilon > 0$) the estimate

$$\mathcal{H}_\varepsilon(W_p^a;\ L_q) \le c\varepsilon^{-\omega^{-1}}. \tag{50}$$

Here $1 \le q \le \infty$ for $p\omega > 1$ and $1 \le q \le q^*$ for $p\omega \le 1$.

First of all we explain the general plan of the proof. The set SW_p^a is first divided into classes, with each class comprising those functions whose approximation with a given accuracy requires, according to Theorem (5.1.9) or (5.1.10), the same sequence of partitions. The number of classes is estimated on the basis Lemma (5.1.5), after which we want to estimate the ϵ-entropy of each class. Suppose a sequence of partitions $\{\Xi_i\}_0^k$ corresponds to certain class. A crude method for calculating ϵ-entropies (approximation of the function u by the function $P_\Xi u$ and estimation of the ϵ-entropy of the unit sphere of the finite-dimentional space $\mathcal{P}(\Xi; L)$) leads to an axcessive estimate. The method does not take into account that the polynomials $P_{\Delta'} u$ and $P_{\Delta''} u$ for neighbouring cubes $\Delta', \Delta'' \in \Xi_k$ cannot differ form each other very strongly. We will take into account the closeness of such polynomials $P_{\Delta'} u$ and $P_{\Delta''} u$ as follows. We consider all piecewise-polynomial approximations $P_{\Xi_i} u (i = 0,1, \dots, k)$, and in passing form the number i to the number $i + 1$ we make use of the fact that for cubes $\Delta', \Delta'' \in \Xi_{i+1}$ contained in the same cube $\Delta \in \Xi_i$, both polynomials $P_{\Delta' u}$ and $P_{\Delta'' u}$ differ little from $P_\Delta u$.

In the proof of Theorem (5.1.19) we require preliminary estimates in a special metric related to a fixed partition. Let Ξ be some partition of the cube Q^m. For the function $u \in \mathrm{L_q}$, along with the usual norm of the space L_q, we consider the norm

$$\|u\|_{q,\Xi} \leq \|u\|_{L_q} \leq |\Xi|^{q^{-1}} \|u\|_{q,\Xi}, \tag{51}$$

Which become equalities when $q = \infty$. If Ξ' is an extension of Ξ then

$$\|u\|_{q,\Xi'} \leq \|u\|_{q,\Xi}.$$

We establish two auxiliary assertions.

Lemma (5.1.20)[129]: Let Ξ be a partition of the cube Q^m and $\tilde{\mathcal{P}} = \tilde{\mathcal{P}}(\Xi; l, M)$ be the set of functions $v \in \mathcal{P}(\Xi; l)$ satisfy the condition

$$\|v\|_{q,\Xi} \leq M. \tag{52}$$

Then for every $\epsilon \leq M$ we have

$$\mathcal{N}_\varepsilon(\tilde{\mathfrak{P}};\ L_{q,\Xi}) \leq\ C_9^{|\Xi|}(M\varepsilon^{-1})^{\nu|\Xi|}, \tag{53}$$

Where the constant $C_9 =\ C_9(m, l, q)$ does not depend on Ξ.

Proof: Let $\mathcal{R} = \mathcal{R}(\Delta; l, M)$ be the set of polynomials r of degree m variables satisfy the condition

$$\|r\|_{L_q(\Delta)} \leq M. \tag{54}$$

For every $\epsilon \leq M$ we have

$$\mathcal{N}_\varepsilon\left(\mathcal{R};\ L_q(\Delta)\right) \leq\ C_9(M\varepsilon^{-1})^\nu. \tag{55}$$

Indeed, for the case $\Delta = Q^m$ inequality (55) follows from the fact that the norm $\|\cdot\|_{L_q(\Delta)}$ on the finite-dimensional (ν-dimensional) space of polynomials is equivalent to the Euclidian norm on the space of coefficients. To pass to an arbitrary cube it suffices to male a transformation of the independent variables, which does not affect the size of the constant C_9.

Condition (52) obviously implies condition (54) for polynomials r obtained by restricting the functions $v \in \tilde{\mathcal{P}}$ to some cube $\Delta \in \Xi$. Here estimate (53) is easily obtained from (55). indeed, the required ϵ-net is the set $\tilde{\mathcal{P}}$ can be formed by means of all possible "pasting together" of elements of ϵ-nets constructed in the set $\mathcal{R}(\Delta; l, M)$ for all cubes $\Delta \in \Xi$. The lemma is proved.

Assume that $\{\Xi_i\}_0^k \subset \mathfrak{R}$ is a sequence of partitions of the cube Q^m where each partition Ξ_i is an extinction of the preceding partition Ξ_{i-1}. As usual, we write $|\Xi_i| = n_{i^*}$. let ζ_i be numbers satisfying the conditions

$$b\zeta_i \leq \zeta_{i+1} \leq \zeta_i\ \ (b > 0; i = 0, 1, \dots, k-1), \tag{56}$$

And let $\hat{\mathcal{P}}_i \subset \mathcal{P}(\Xi_i; l)$ $(i = 0, 1, \dots, k)$ be certain sets of piecewise-polynomial functions, where $\hat{\mathcal{P}}_i$ is the $2\zeta_i$- net for the set $\hat{\mathcal{P}}_{i+1}$ in the metric of $L_{q,\Xi_{i+1}}$ $(i = 0, 1, \dots, k-1)$.

Lemma (5.1.21)[129]: Under the above assumptions we have

$$\mathcal{N}_{\zeta_i}(\hat{\mathcal{P}}_i; L_{q,\Xi_i}) \leq\ C_{10}^{n_1+\dots+n_i} \mathcal{N}_{\zeta_i}\left(\hat{\mathcal{P}}_0; L_q\right)\ (i = 1, \dots, k). \tag{57}$$

The constant $C_{10} = C_{10}(m, l, q, b)$ does not depend on the sequence of partitions being considered or on the numbers ζ_i.

Proof: For every element $v \in \hat{\mathcal{P}}_{i+1}$ there is an element $\tilde{v} \in V_i$ we obtain

$$\|v - \tilde{v}\|_{q,\Xi_{i+1}} \leq \|v - v'\|_{q,\Xi_{i+1}} + \|v' - \tilde{v}\|_{q,\Xi_i} \leq 3\zeta_i.$$

On the strength of Lemma (5.1.20) we can construct a ζ_{i+1}- net Z_{i+1}, in the set $\tilde{\mathcal{P}}(\Xi_{i+1}; l, 3\zeta_i)$ whose cardinality does not exceed the equality

$$C_9^{n_{i+1}}(3\zeta_i\zeta_{i+1}^{-1})^{\ \nu n_{i+1}} \leq\ C_{10}^{n_{i+1}}\ \ (C_{10} = C_9 3^\nu b^{-\nu}).$$

Since $v-\tilde{v} \in \tilde{\mathcal{P}}(\Xi_{i+1}; l, 3\zeta_i)$ for some element $z \in Z_{i+1}$ we have

$$\|v-\tilde{v}-z\|_{q,\Xi_{i+1}} \leq \zeta_{i+1}.$$

All possible elements of the form $\tilde{v}+z$, where $\tilde{v} \in V_i$, $z \in Z_{i+1}$, form ζ_{i+1}- net for the set $\hat{\mathcal{P}}_{i+1}$ relative to the metric of the space $L_{q,\Xi_{i+1}}$. The cardinality of this net is estimated, obviously, by the quantity

$$C_{10}^{n_{i+1}} \mathcal{N}_{\zeta_i}\big(\hat{\mathcal{P}}_i;\ L_{q,\Xi_i}\big).$$

Thus we have obtained the estimate

$$\mathcal{N}_{\zeta_{i+1}}\big(\hat{\mathcal{P}}_{i+1};\ L_{q,\Xi_{i+1}}\big) \leq C_{10}^{n_{i+1}}\ \mathcal{N}_{\zeta_i}\big(\hat{\mathcal{P}}_i;\ L_{q,\Xi_i}\big),$$

From which (57) follows. The lemma is proved.

We proceed to the proof of Theorem (5.1.19). We use the notation $\mathcal{T}_a^{\eta}(J)$, (8), (9), (18), and relate it to the partition functions $J=J_u$ defined by the formula (34). In the case $p\omega>1$ it obviously suffices to prove the theorem for $q=\infty$.

Let η be a fixed number, $0<\eta\leq C_2$. Set $a=p\omega-1$ for $p\omega>1$ and $a=p(q^{-1}-q^{*-1})$ for $p\omega\leq 1. q<q^*$. Divide the set SW_p^a into classes, associating two functions $u_1, u_2 \in SW_p^a$ with the same if and only if $\mathcal{T}_a^{\eta}\big(J_{u_1}\big)=\mathcal{T}_a^{\eta}\big(J_{u_2}\big)$. The number of distinct classes obviously does not exceed the number $N(a,\eta)$. on the strength of Lemma (5.1.5), the latter can be estimated as follows:

$$\log_2 N(a,\eta) \leq C_3\eta^{-(a+1)^{-1}} \quad (0<\eta\leq C_2) \tag{58}$$

Now we estimate the entropy of each of the classes. We use $\hat{S}$ to denote some class. With all functions $u \in \hat{S}$ we associate the same sequence of partitions $\{\Xi_i\}_0^k=\mathcal{T}_a^{\eta}(J)$ and also the same sequence of numbers $\tilde{\delta}_i=\tilde{\delta}_i(J_u;a)$ $(i=0,1,\dots,k)$. The number k is determined by (22).

Using the notation (9), (34) we observe that we can write the assertions of Lemmas (5.1.7) (the case $q=\infty$) and (5.1.8) (the case $q<q^*$) in the form

$$\big\|u-P_{\Xi_i}u\big\|_{q,\Xi_i} \leq C_{11}\delta_i^{p^{-1}} (i=0,1,\dots,k). \tag{59}$$

Here $C_{11}=C_4$ for $q=\infty$ and $C_{11}=C_5$ for $q<q^*$. Taking (19) into account, we obtain the relationships

$$\big\|u-P_{\Xi_i}u\big\|_{q,\Xi_i} \leq \zeta_i\ \left(\zeta_i=C_{11}\delta_i^{p^{-1}}; i=0,1,\dots,k\right) \tag{60}$$

Which are somewhat cruder but will be more convenient below. We set $\hat{\mathcal{P}}_i=P_{\Xi_i}\hat{S}$ and note that the sets $\hat{\mathcal{P}}_i$ and the numbers ζ_i satisfy the conditions of Lemma (5.1.21). Indeed, from (60) we find that

$$\begin{aligned}\big\|P_{\Xi_i}u-P_{\Xi_{i+1}}u\big\|_{q,\Xi_{i+1}} &\leq \big\|u-P_{\Xi_{i+1}}u\big\|_{q,\Xi_{i+1}}+\big\|u-P_{\Xi_i}u\big\|_{q,\Xi_i}\\ &\leq \zeta_{i+1}+\zeta_i \leq 2\zeta_i.\end{aligned}$$

The inequalities (56) are obviously satisfied in the view of (21).

We estimate the quantity $\mathfrak{N}_{\zeta_0}\big(\hat{\mathcal{P}}_0;\ L_q\big)$. Since $P_{\Xi_0}=P_{Q^m}$ and $\tilde{\delta}_0=C_2$, from (60) and the imbedding theorem we obtain

$$\big\|P_{Q^m}u\big\|_{L_q} \leq C_{11}C_3^{p^{-1}}+\|u\|_{SW_p^a} \leq C_{11}C_2^{p^{-1}}+C_{11}\|u\|_{SW_p^a} \leq C_{11}\left(C_2^{p^{-1}}+1\right).$$

It follows from (55) that

$$\mathcal{N}_{2\zeta_k}\big(\hat{S};\ L_{q,\Xi_k}\big) \leq \mathcal{N}_{\zeta_k}\big(\hat{\mathcal{P}};\ L_{q,\Xi_k}\big) \leq C_{12}C_{10}^{n_1+\dots+n_k}. \tag{61}$$

According to (25), (21) and (22)

$$n_1+\dots+n_k \leq C_3'\tilde{\delta}_k^{-(a+1)^{-1}} \leq 2^m C_3'\tilde{\delta}_{k-1}^{-(a+1)^{-1}} \leq 2^m C_3'\eta^{-(a+1)^{-1}}.$$

Hence and from (61) we obtain an estimate of the form

$$\mathcal{H}_{2\zeta_k}(\hat{S};\ L_{q,\Xi_k}) \le C_{13}\eta^{-(a+1)^{-1}}. \tag{62}$$

To obtain the final result we have to pass to estimating the entropy in the original metric, i.e. in the metric of L_q. Setting $\epsilon_k = 2\zeta_k n_k^{q-1}$ and comparing (62) and (51), we obtain

$$\mathcal{H}_{\varepsilon_k}(\hat{S};\ L_q) \le C_{13}\eta^{-(a+1)^{-1}}.$$

Further, it follows from (24) and (22) that

$$\varepsilon_k \le c\tilde{\delta}_k^{p^{-1}}\eta^{-(a+1)^{-1}} \le c\eta^{p^{-1}-q^{-1}(a+1)^{-1}}.$$

Thus

$$\mathcal{H}_{\varepsilon}(\hat{S};\ L_q) \le C_{13}\eta^{-(a+1)^{-1}}, \tag{63}$$

where

$$\varepsilon = c\eta^{p^{-1}-q^{-1}(a+1)^{-1}}.$$

Finally, combining the inequalities (63) and (58), we arrive at the estimate

$$\mathcal{H}_{\varepsilon}(SW_p^a;\ L_q) \le c\varepsilon^{[p^{-1}-q^{-1}(a+1)^{-1}]^{-1}}.$$

It is easy to see that the last inequality coincides with the estimate (50). Indeed, the relationship $p^{-1}(a+1)^{-1} - q^{-1} = \omega$ holds for both $q = \infty, a = p\omega - 1$ and for $q < q^*, a = p(q^{-1} - q^{*-1}), q^* = p(1 - p\omega)^{-1}$. The theorem is proved.

We note that we can establish the estimates

$$\mathcal{H}_{\varepsilon}(H_p^a;\ L_q) \le c\varepsilon^{-\omega^{-1}}, \mathcal{H}_{\varepsilon}(B_p^a;\ L_q) \le c\varepsilon^{-\omega^{-1}}$$

is exactly the same way.

Corollary(5.1.22)[129]: When $p\omega > 1$ the relationship $\mathcal{H}_{\varepsilon}(W_p^a;\ C) \asymp \epsilon^{-\omega^{-1}}$ holds.

Indeed, the estimate from above obviously coincides with the estimate (50) for $q = \infty$. The estimate from below for integral a follows from the inclusion $C^a \subset W_p^a$ and the inequality obtained in [134]:

$$\mathcal{H}_{\varepsilon}(C^a; C) \ge c\varepsilon^{-\omega^{-1}}. \tag{64}$$

For nonintegral a the class C^a is not in W_p^a. However, in this case we can also obtain the required estimate below of $\mathcal{H}_{\varepsilon}(W_p^a;\ C)$ with the help of the system of functions which was used in [134] to obtain (64).

In conclusion we make some remarks about our estimates for e-entropy and rcdiameters. For simplicity we restrict consideration to the case of the sphere $SW_2^1(Q^1)$considered in the metric of$c(Q^1)$.

The inclusions

$$Sc^1 \subset SW_2^1 \subset Sc^{1/2}$$

imply the inequalities

$$\mathcal{H}_{\varepsilon}(C^1; C) \le \mathcal{H}_{\varepsilon}(W_2^1; C) \le \mathcal{H}_{\varepsilon}\left(C^{1/2}; C\right),$$

$$d_n(C^1; C) \le d_n(W_2^1; C) \le d_n\left(C^{1/2}; C\right).$$

Above we saw that the quantity$\mathcal{H}_{\varepsilon}(W_2^1; C)$has the same order of magnitude as $\mathcal{H}_{\varepsilon}(C^1; C)$:

$$\mathcal{H}_{\varepsilon}(W_2^1; C) \asymp \mathcal{H}_{\varepsilon}(C^1; C) \asymp \varepsilon^{-1}.$$

As for the re-diameters, the precise order of$d_n(W_2^1; C)$ is unknown. The most natural assumption is

$$d_n(W_2^1; C) \asymp d_n\left(C^{1/2}; C\right) \asymp n^{-1/2}. \tag{65}$$

This would mean that from the point of view of approximation by linear sets in the metric of the space C the sphere$SW_2^1$is not better than the widest set$SC^{1/2}$. At the same time, from the point of view of ϵ-entropy, the sphere SW_2^1 is constructed essentially like the narrowest set SC^1. We note that for several other metric characteristics-n-diameters in the sense of I. M. Gelf and (see [141])- relationships of the form (65) are indeed valid.

Section (5.2): Multivariate BV Spaces of a Wiener–L. Young

In the first, the 1967 pioneering [147], asserts the following:

Theorem (5.2.1)[142]: Given $f \in W_p^k([0,1)^d), N \in \mathbb{N}$ and $1 \le p < q < \infty$ satisfying

$$\frac{k}{d} > \frac{1}{p} - \frac{1}{q}$$

there exist a partition Δ_N of $[0,1)^d$ into at most N dyadic subcubes and a piecewise polynomial g_N on Δ_N of degree $k-1$ such that

$$\|f - g_N\|_q \le CN^{-k/d} \sup_{|\alpha|=k} \|D^\alpha f\|_p ; \tag{66}$$

the constant $C > 0$ is independent of f and N and $C \to \infty$ as q tends to the Sobolev limiting exponent $q^* := \left(\frac{1}{p} - \frac{k}{d}\right)^{-1}$.

Using the compactness argument from [147] one can prove that validity of (66) for $q = q^*$ implies (incorrect) compactness of embedding $W_p^k \subset L_{q^*}$ .This leads to the following:

For the special case $k = p = 1, d = 2$ the answer was given in [151] by A. Cohen, DeVore, Petrushev and Hong Xu; the case $d > 2$ was than proved by Wojtaszczyk [156]. The result states:

Theorem (5.2.2)[142]: Given $f \in W_1^1([0,1)^d), d > 2$, and $N \in \mathbb{N}$ there exist a partition Δ_N of $[0,1)^d$ into at most N d-rings (differences of two dyadic subcubes) and a piecewise constant function g_N on Δ_N such that

$$\|f - g_N\|_{q^*} \le C(d)N^{-1/d} \sup_{|\alpha|=k} \|D^\alpha f\|_1 ; \tag{67}$$

Hereafter $c(x, y, ...)$ denotes a positive constant depending only on the parameters in the parentheses.

We achieved by using the BV spaces of integrable on $[0,1)^d$functions of arbitrary smoothness introduced in [149]. To motivate the definition of the corresponding space denoted by V_{pq}^k we begin with a model case, the Wiener–L. Young space V_p, whose associated seminorm is presented in the following equivalent form:

$$\mathrm{var}_p f := \sup_\Delta \left(\sum_{I\in\Delta} osc(f;I)^p\right)^{1/p} \tag{68}$$

where Δ runs over disjoint families of intervals $I = [a,b) \subset [0,1)$ and

$$osc(f;I) := \sup_{x,y\in I} |f(x) - f(y)|. \tag{69}$$

To obtain the required seminorm of V_{pq}^kdenoted by $\mathrm{var}_p^k(\cdot; L_q)$ we replace in (68) intervals by cubes $Q \subset [0,1)^d$, in (69) the first difference by the k-th one, and the uniform norm by L_q norm. This gives the following:

Definition (5.2.3)[142]: The seminorm $f \mapsto \mathrm{var}_p^k(f; L_q)$ is a function on $L_q([0,1)^d)$given by

$$\mathrm{var}_p^k(f;L_q):=\sup_{\Delta}\left\{\sum_{Q\in\Delta} osc_q^k(f;Q)^p\right\}^{1/p} \tag{70}$$

where Δ runs over disjoint families of cubes $Q\subset[0,1)^d$ and

$$osc_q^k(f;Q):=\sup_{h\in\mathbb{R}^d}\left\|\Delta_h^k\right\|_{L_q(Q_{kh})}; \tag{71}$$

here $\Delta_h^k:=\sum_{j=0}^k(-1)^{n-j}\binom{k}{j}\delta_{jh}$ and $Q_{kh}:=\{x\in\mathbb{R}^d; x+jh\in Q,\ j=0,1,\ldots,k\}$.

The important characteristic of the space $V_{pq}^k$is its smoothness introduced by the following:

Definition (5.2.4)[142]: Smoothness of the space V_{pq}^kdenoted by $s(V_{pq}^k)$ is a real number given by

$$s(V_{pq}^k):=d\left(\frac{1}{p}-\frac{1}{q}\right). \tag{72}$$

This concept is closely related to differential and approximation properties of V_{pq}^k functions. In fact, a function with $s(V_{pq}^k)=s$ belongs to the Taylor class $T_q^s(x)$ a.e. if $0<s<k$ and $t_q^k(x)$ a.e. if $s=k$, see [27]. Moreover, as we will see, its order of N-term approximation in $L_q([0,1)^d)$ by piecewise polynomial is $N^{-s/d}$ for $0<s\le k$.

In particular, the proof of Theorem (5.2.23) is based on the equality

$$V_{pq^*}^k=W_p^k$$

and the fact that

$$s(V_{pq^*}^k):=d\left(\frac{1}{p}-\frac{1}{q^*}\right)=k$$

that allow to derive it directly from the corresponding result for $V_{pq^*}^k$.

The main result, Theorem (5.2.23), is formulated along with its consequences describing similar approximation results for classical smoothness spaces.

We prove properties of V_{pq}^k spaces essential for the proof of Theorem (5.3.2). The first result asserts that a function $f\in V_{pq}^k$ can be weakly approximated in L_q by C^∞ functions whose (k,p)-variations are bounded by $\mathrm{var}_p^k(f;L_q$ For the special case of the space BV $([0,1)^d)$ $(=V_{1\,d/d-1}^1)$, see, e.g., [157].

The second result estimates polynomial approximation of order $k-1$ for $f\in V_{pq}^k(Q'\backslash Q'')$ via its (k,p)-variation; here $Q''\subsetneq Q'$are dyadic cubes.

The latter result essentially uses a cover lemma proved in collaboration with V. Dolnikov; its proof is presented.

The main result, Theorem (5.2.22), and its consequences for the classical smoothness spaces. The approximation algorithm used in the construction of the family of dyadic cubes for Theorem (5.2.22).

Its primary version was developed to prove the similar to Theorem (5.2.23) result for the N-term approximation of functions from Besov spaces by B–splines; the result is announced in [145] and proved in [28].

The special case of Theorem (5.2.22) for functions with absolutely continuous (k,p)-variation was proved a long time ago and announced in [150]. This result allows to prove all consequences of Theorem (5.2.22) presented but only much later he derive Theorem (5.2.22) from this special case.

A cube denoted by Q, Q', K etc. is a set of $\mathbb{R}^d$ homothetic to the (half–open) unit cube

$$Q^d := [0,1)^d. \quad (73)$$

$\mathcal{D}(Q)$ denotes the family of dyadic cubes of Q, i.e., cubes of the form

$$K := 2^{-j}(Q + \alpha) \quad (74)$$

where $j \in \mathbb{Z}_+ := \{0,1,2,\ldots,\}$ and $\alpha \in \mathbb{Z}^d$.

Further, $\mathcal{P}_l = \mathcal{P}_l(\mathbb{R}^d)$ is the space of polynomials in $x := (x_1, x_2, \ldots, x_d)$ of degree l while $\mathcal{P}_l(\Delta)$ denotes the space of piecewise polynomials on a set $\Delta \subset \mathcal{D}(Q)$ of degree l.

In other words,

$$\mathcal{P}_l(\Delta) := \left\{ f \in L_\infty(Q); f = \sum_{K \in \Delta} P_K . 1_K \right\} \quad (75)$$

where $\{P_K\}_{K\in\Delta} \subset \mathcal{P}_l$.

Stipulation (5.2.5)[142]: We drop the symbol Q^d from the next notations writing, e.g., $\mathcal{D}$, V_{pq}^k, L_q instead of $\mathcal{D}(Q^d)$, $V_{pq}^k(Q^d)$, $L_q(Q^d)$, if it does not lead to misunderstanding.

The first consequence of the main result, Theorem (5.2.23), immediately follows from Theorem(5.2.22)(a) and the inequality

$$\operatorname{var}_p^k(f; L_{q^*}) \le c \begin{cases} |f|_{W_p^k} & \text{if } p > 1 \\ |f|_{BV^k} & \text{if } p = 1, \end{cases} \quad (76)$$

here $c = c(k, d, q^*)$ and $q^* := \left(\frac{k}{d} - \frac{1}{p}\right)^{-1}$.

This and analogous embedding results for Besov spaces.

Let now $\dot{B}_p^{\lambda\theta} := \dot{B}_p^{\lambda\theta}(Q^d)$ be the homogeneous Besov space defined by the seminorm

$$|f|_{B_p^{\lambda\theta}} := \left\{ \int_0^1 \left(\frac{\omega_k(t; f; L_p)}{t^\lambda} \right)^\theta \frac{dt}{t} \right\}^{1/\theta} \quad (77)$$

where $k = k(\lambda) := \min\{n \in \mathbb{N}; n > \lambda\}$ and $\omega_k(\cdot; f; L_p)$is the k-th modulus of continuity of $f \in L_p$, see e.g., [154] or [152] for its definition.

The first result concerns the "diagonal" Besov space $\dot{B}_p^\lambda := \dot{B}_p^{\lambda p}$, $1 \le p < \infty$.

Definition (5.2.6)[142]: (k,p)-variation of a function $f \in L_q^{loc}(\mathbb{R}^d)$is a set-function on subsets $S \subset \mathbb{R}^d$with nonempty interior given by

$$\operatorname{var}_p^k(f; S; L_q) := \sup_\Delta \left\{ \sum_{Q\in\Delta} E_k(f; Q; L_q)^p \right\}^{1/p} \quad (78)$$

where Δ runs over all disjoint families of cubes $Q \subset S$.

Equivalence of this definition to the previous one follows from the main result of [148] implying, e.g., the next two-sided inequality with constants depending only on k:

$$osc_p^k(f; Q; L_q) \approx E_k(f; Q; L_q).$$

It should be pointed out that in what follows all definitions and results involving the space V_{pq}^kuse Definition (5.2.6). In particular, the associated seminorm of this space is

$$\|f\|_{V_{pq}^k} := \sup_\Delta \left\{ \sum_{Q\in\Delta} E_k(f; Q; L_q)^p \right\}^{1/p}$$

where Δ runs over all disjoint families of $Q \subset Q^d$.

We present some basic properties of (k, p)-variation starting with those following directly from Definition (5.2.6).

Proposition (5.2.7)[142]: (Subadditivity). Let $\{S_i\}$ be a disjoint families of measurable sets with nonempty interiors. Then

$$\left\{\sum_i \operatorname{var}_p^k(f; S_i; L_q)^p\right\}^{1/p} \le \operatorname{var}_p^q\left(f; \bigcup_i S_i; L_q\right). \tag{79}$$

(Lower semicontinuity) If $\{f_j\}$ converges in L_q to a function f, then

$$\operatorname{var}_p^k(f; S; L_q) \le \varliminf_{j\to\infty} \operatorname{var}_p^k(f_j; S; L_q) \quad . \tag{80}$$

Proof: Let $\Delta \coloneqq \{Q\}$ be a disjoint family of cubes and

$$\operatorname{var}_p^k(f; \Delta; L_q) := \left\{\sum_{Q\in\Delta} E_k\big(f; Q; L_q\big)^p\right\}^{1/p} . \tag{81}$$

If $\{\Delta_i\}$ is disjoint and $\bigcup_{Q\in\Delta_i} Q \subset S_i$, then

$$\sum_i \operatorname{var}_p^k(f; \Delta_i; L_q)^p = \sum_i \operatorname{var}_p^k\left(f; \bigcup \Delta_i; L_q\right)^p \le \operatorname{var}_p^q\left(f; \bigcup_i S_i; L_q\right)$$

and it remains to take supremum over each Δ_i to prove (79). The property (80) is proved similarly.

A more substantive property of (k, p)-variation gives the next result.

Proposition (5.2.8)[142]: Let a C^∞ function f belong to the space V_{pq}^k of smoothness $s \le k$. Then uniformly in S

$$\lim_{|S|\to 0} \operatorname{var}_p^k(f; S; L_q) \tag{82}$$

Hereafter $|S|$ denotes the Lebesgue d-measure of S.

Proof: Let Δ be a disjoint family of cubes $Q \subset S$. By the Taylor formula

$$E_k\big(f; Q; L_q\big) \le |Q|^{1/q} E_k(f; Q; C) \le c(k,d)|Q|^{1/q+k/d} \max_{|\alpha|=k} \max_Q |D^\alpha f|\,;$$

this implies

$$\operatorname{var}_p^k\big(f; \Delta; L_q\big) \le c(k,d)\left\{\sum_{Q\in\Delta} |Q|^{p\left(\frac{1}{q}+\frac{k}{d}\right)}\right\}^{1/p} |f|_{C^k(Q^d)}.$$

Since $p\left(\frac{1}{q}+\frac{k}{d}\right) \ge p\left(\frac{1}{q}+\frac{s}{d}\right) = 1$, the sum here is bounded by $\left\{\sum_{Q\in\Delta}|Q|\right\}^{1/p} \le |S|^{1/p}$, and therefore

$$\operatorname{var}_p^k\big(f; S; L_q\big) \coloneqq \sup_\Delta \operatorname{var}_p^k\big(f; \Delta; L_q\big) \to 0 \ as\ |S| \to 0.$$

Since the space V_{pq}^k is, in general, nonseparable, C^∞ approximated functions form a proper subspace of V_{pq}^k. However, a weaker form of C^∞approximation is true.

Theorem (5.2.9)[142]: Let a function f belong to V_{pq}^k if $q < \infty$ and to $V_{p\infty}^k \cap C(\mathbb{R}^d)$, other-wise. Assume that Q is a subcube of Q^d such that

$$dist(Q; \mathbb{R}\backslash Q^d) > 0. \tag{83}$$

Then there exists a sequence $\{f_j\} \subset C^\infty(\mathbb{R}^d)$ such that

$$\lim_{j\to\infty} f_j = f\ \left(convergence\ in\ L_q(Q)\right) \tag{84}$$

and, moreover,

$$\sup_j \mathrm{var}_p^k(f_j; Q; L_q) \le \mathrm{var}_p^k(f; L_q).$$

Proof: Let f_ε be a regularizer of f given by

$$f_\varepsilon(x) := \int_{Q_\varepsilon} f(x - \varepsilon y)\varphi(y)dy, \quad x \in Q, \tag{85}$$

where $\varphi \in C^\infty(\mathbb{R}^d)$ is a test function, i.e.,

$$\varphi \ge 0, \int \varphi dx = 1 \text{ and } supp\ \varphi \subseteq [-1,1]^d; \tag{86}$$

here $\varepsilon > 0$ is such that

$$Q_\varepsilon := Q + [-\varepsilon, \varepsilon]^d \subseteq Q^d, \tag{87}$$

see (83).

Now (86) and the Minkowski inequality yield

$$\|f - f_\varepsilon\|_{L_q(Q)} \le \sup_{|y|\le 1} \|f(\cdot - \varepsilon y) - f\|_{L_q(Q)}.$$

Since the right-hand side tends to 0 as $\varepsilon \to 0$ for $q < \infty$ and for $q = \infty$if $f \in C(\mathbb{R}^d)$, (84) follows.

To proceed we need the following:

Lemma (5.2.10)[142]: It is true that

$$E_k(f_\varepsilon; Q; L_q) \le \int_{|y|\le 1} E_k(f_\varepsilon; Q - \varepsilon y; L_q)\varphi(y)dy. \tag{88}$$

Proof. It suffices to prove (88) for $q < \infty$ and then pass q to $+\infty$.

Let $q < \infty$ and q' denote the conjugate exponents. By $\mathcal{P}_{k-1}^{\perp}(Q)$ we denote the set of functions $g \in L_{q'}(\mathbb{R}^d)$ such that

$$\|g\|_{L_{q'}} = 1, supp\ g \subset Q, \int x^\alpha g(x)dx = 0, \ |\alpha| \le k - 1. \tag{89}$$

By the duality of L_q and $L_{q'}$

$$E_k(f_\varepsilon; Q; L_q) = \sup\left\{\int_Q f_\varepsilon g dx\, ; g \in \mathcal{P}_{k-1}^{\perp}(Q)\right\}. \tag{90}$$

On the other hand,

$$\left|\int_Q f_\varepsilon g dx\right| \le \int_{\mathbb{R}^d} \left|\int_{Q-\varepsilon y} f(x)g(x + \varepsilon y)dx\right| \varphi(y)dy \tag{91}$$

and the function $x \mapsto g(x + \varepsilon y), x \in Q$, clearly, belongs to the set $\mathcal{P}_{k-1}^{\perp}(Q - \varepsilon y)$.Therefore for every polynomial m of degree $k - 1$

$$\int_{Q-\varepsilon y} f(x)g(x + \varepsilon y)dx = \int_{Q-\varepsilon y} f(x)(g - m)(x + \varepsilon y)dx.$$

Combining this with (91) and using the Hölder inequality we obtain

$$\left|\int_Q f_\varepsilon g dx\right| \le \int_{\mathbb{R}^d} \varphi(y)\|f - m\|_{L_q(Q-\varepsilon y)}\|g\|_{L_{q'}(\mathbb{R}^d)}dy.$$

Taking here infimum over all polynomials m and then supremum over all $g \in \mathcal{P}_{k-1}^{\perp}(Q)$ we get by (90)

$$E_k(f_\varepsilon; Q; L_q) \le \int_{\mathbb{R}^d} E_k(f_\varepsilon; Q - \varepsilon y; L_q)\varphi(y)dy.$$

The proof is complete.

To this end, first we estimate $\mathrm{var}_p^k(f; \Delta; L_q)$, see (81), for the disjoint family of cubes of Q. Due to (88) the Minkowski inequality gives for such Δ

$$\mathrm{var}_p^k(f_\varepsilon; \Delta; L_q) \le \int_{\mathbb{R}^d} \mathrm{var}_p^k(f; \Delta - \varepsilon y; L_q)\varphi(y)dy.$$

Since $\Delta - \varepsilon y := \{\hat{Q} - \varepsilon y; \hat{Q} \in \Delta\}$ is the disjoint family of cubes containing for small ε in Q^d, the right-hand side is bounded by $\mathrm{var}_p^k(f; L_q)$.Taking then supremum over all such Δ, we obtain the inequality

$$\mathrm{var}_p^k(f_\varepsilon; Q; L_q) \le \mathrm{var}_p^k(f; L_q)$$

Unfortunately, the corresponding extension theorem is unknown though it exists for some spaces V_{pq}^k, e.g., for $s(V_{pq}^k) = k$. The special case of the last assertion for the space $BV\ (Q^d)$ and even for more general class of domains is presented, e.g., in [157].

This remark leads to the following:

Conjecture. For every $f \in V_{pq}^k$ there is a sequence $\{f_j\} \subset C^\infty(\mathbb{R}^d)$ such that

$$\|f - f_j\|_{L_q} \to 0 \text{ as } j \to \infty$$

and, moreover,

$$\lim_{i\to\infty} \mathrm{var}_p^k(f_j; L_q) \le \mathrm{var}_p^k(f; L_q) \tag{92}$$

Theorem (5.2.11)[142]: Let $Q \subset Q^*$ be dyadic subcubes of Q^d. Then it is true that

$$E_k(f; Q^* \backslash Q; L_q) \le c(k,d)\mathrm{var}_p^k(f; Q^* \backslash Q; L_q). \tag{93}$$

Lemma (5.2.12)[142]: Let $S_1, S_2 \subset \mathbb{R}^d$ be subsets of finite measure such that for $\varepsilon > 0$

$$|S_1 \cap S_2| \ge \varepsilon \cdot \min_{i=0,1}\{|S_i|\}. \tag{94}$$

Then the following is true:

$$E_k(f; S_1 \cup S_2; L_q) \le c\varepsilon^{-k+1} \sum_{i=0}^{1} E_k(f; S_j; L_q). \tag{95}$$

For the proof see, e.g., [143].

Lemma (5.2.13)[142]: Let $\{S_j\}_{1\le j\le N}$ be a family of subsets in $\mathbb{R}^d$ of finite measure such that for some $\varepsilon > 0$

$$|S_j \cap S_{j+1}| \ge \varepsilon \min\{|S_j|, |S_{j+1}|\}, \quad 1 \le j < N. \tag{96}$$

Then it is true that

$$E_k\left(f; \bigcup_{j=1}^{N} S_j; L_q\right) \le c \sum_{j=1}^{N} E_k(f; S_j; L_q). \tag{97}$$

where $c = (c(k,d)\varepsilon^{-k+1})^{N-1}$.

Proof (induction on *N*). For $N = 2$ the result follows from (95). Now assume that (97) holds for $N \ge 2$ and prove it for $N + 1$.

Setting $S^M := \bigcup_{j=1}^{M} S_j$ we get from (94)

$$|S^N \cap S_{N+1}| \ge \varepsilon |S_N \cap S_{N+1}| \ge \varepsilon \min\{|S_N|, |S_{N+1}|\} = \varepsilon \min\{|S^N|, |S_{N+1}|\}.$$

Further, Lemma (5.2.12) implies

$$E_k\left(f;\bigcup_{j=1}^{N+1} S_j\,;L_q\right)\le c(k,d)\varepsilon^{-k+1}\left(E_k\big(f;S^N;L_q\big)+E_k\big(f;S_{N+1};L_q\big)\right).$$

while the induction hypothesis gives

$$E_k\big(f;S^N;L_q\big)\le (c(k,d)\varepsilon^{-k+1})^{N-1}\sum_{j=1}^{N} E_k\big(f;S_j;L_q\big).$$

Combining these we get the result for $N+\ 1$.

Theorem (5.2.14)[142]: There exists a cover $\mathcal{K}$ of $Q^*\backslash Q$ by cubes such that the following holds:

For every overlapping pair [2] $K_1,K_2\in\mathcal{K}$

$$|K_1\cap K_2|\ge\frac{1}{2}\min\{|K_1|,|K_2|\},\tag{98}$$

and, moreover,

$$\text{card}\,\mathcal{K}\le 4(2^d-1).\tag{99}$$

Now we complete the proof of Theorem (5.2.11).

By Lemma (5.2.13) and Theorem (5.2.14) we have

$$E_k\big(f;Q^*\backslash Q;L_q\big)\le (c(k,d)2^{k-1})^{4(2^d-1)}\sum_{K\in\mathcal{K}} E_k\big(f;K;L_q\big).$$

Moreover, by the definition of (k,p)-variation, see (78),

$$E_k\big(f;K;L_q\big)\le \text{var}_p^k\big(f;Q^*\backslash Q;L_q\big).$$

for every $K\in\mathcal{K}$.

Together with the previous inequality this gets the required result

$$E_k\big(f;Q^*\backslash Q;L_q\big)\le c(k,d)\ \text{var}_p^k\big(f;Q^*\backslash Q;L_q\big).$$

We begin with part (a) of this result and then derive from (a) part (b).

Let $f\in V_{pq}^k(:=\ V_{pq}^k(Q^d))$ where

$$1\ \le p<q<\ \infty, d\ \ge\ 2 \text{ and } 0<s:=\ s\big(V_{pq}^k\big)\le k.\tag{100}$$

Without loss of generality we assume that

$$|f|_{V_{pq}^k}=1.\tag{101}$$

Under these assumptions given $N\in\mathbb{N}$ we prove existence of a cover Δ_N of Q^d by at most N dyadic cubes and a piecewise polynomial $g_N\in\mathcal{P}_{k-1}(\Delta_N)$ such that

$$\|f-g_N\|_q\le c(k,d)N^{-s/d}.\tag{102}$$

First, let $C^\infty(\mathbb{R}^d)\cap V_{pq}^k$. The proof of (102) for this case begins with the construction of the cover Δ_N. This will be obtained by the algorithm presented now.

The important ingredient of the algorithm is a weight W defined on the σ-algebra $A(\mathcal{D})$ generated by dyadic cubes of Q^d. This by definition is a function $W:\ A(\mathcal{D})\ \to\ \mathbb{R}_+$ satisfying the conditions.

(Subadditivity) For a disjoint family $\{S_i\}\ \subset A(\mathcal{D})$

$$\sum W(S_i)\le W\left(\bigcup S_i\right).\tag{103}$$

(Absolute continuity)

$$\lim_{|S|\to 0} W(S)=0.\tag{104}$$

We normalize W by

$$W(Q^q)=\ 1.\tag{105}$$

To prove Theorem (5.2.22)(a) for $f \in V^k_{pq} \cap C^\infty$ we define a weight W by

$$W(S) := \operatorname{var}^k_p(f; S; L_q)^p, S \in A(\mathcal{D}). \tag{106}$$

Due to Propositions (5.2.7), (5.2.8)and (101) satisfies the required properties (103)- (105).

In the construction of the algorithm we essentially exploit the canonical graph structure of the set $\mathcal{D}$ regarding as the vertex set while the edge set consists of pairs $\{Q', Q\} \subset \mathcal{D}$ such that $Q' \subset Q$ and $|Q'| = 2^{-d}|Q|$. In this case, we use the notation $Q' \to Q$ and call Q' the son of Q and Q the father of Q'.

The set of all 2^d sons of Q is denoted by $\mathcal{D}_1(Q)$. This clearly is the uniform partition of Q into 2^d congruent subcubes.

Further, a path in the graph $\mathcal{D}$ is a sequence

$$P := \{Q_1 \to Q_2 \to \cdots \to Q_n\}. \tag{107}$$

The vertices (cubes) Q_1, Q_n are called the tail and the head of P, respectively. Moreover, we use the notations

$$P := [Q_1, Q_n], Q_1 =: T_P =: P^{-1}, Q_n =: H_P =: P^+. \tag{108}$$

It is readily seen that the following is true.

Proposition (5.2.15)[142]: If $Q' \subset Q$ are dyadic cubes of $\mathcal{D}$, there exists a unique path joining Q' and Q.

In terms of Graph Theory, $\mathcal{D}$ is a rooted tree with the root Q^d.

More generally, the set $\mathcal{D}(Q)$ of all dyadic subcubes of $Q \in \mathcal{D}$ is a rooted tree with the root Q.

For $N \in \mathbb{N}$ and W given by (106) the subset of "bad" cubes of $\mathcal{D}$ is defined by

$$G_N := \{Q \in \mathcal{D};\ W(Q) \geq N^{-1}\}; \tag{109}$$

clearly, $Q^d \in G_N$, see (105), and GN is finite, see (104).

The algorithm gives the following partition of G_N into the set of (basic) paths, see Proposition II. 1 of Appendix II for the proof.

Proposition (5.2.16)[142]: There exists partition $\mathcal{B}_N$ of the set $G_N \backslash \{Q^d\}$ into N paths such that

$$W(H_B \backslash T_B) < N^{-1},\ \ B \in \mathcal{B}_N, \tag{110}$$

and, moreover,

$$\operatorname{card} \mathcal{B}_N \leq 3N + 1. \tag{111}$$

Now we decompose the remaining part of $\mathcal{D}$

$$G^c_N := \mathcal{D} \backslash G_N. \tag{112}$$

To this end we define the boundary of G_N denoted by ∂G_N that consists of all maximal cubes of G^c_N with respect to the set–inclusion order.

In other words, every $Q' \in \mathcal{D}$ containing $Q \in \partial G_N (\subset G^c_N)$ as a proper subset belongs to G_N. In particular, if Q^+ is the father of $Q \in \partial G_N$, then

$$W(Q) < N^{-1} \text{ and } W(Q^+) \geq N^{-1}. \tag{113}$$

Proposition (5.2.17)[142]: (a) The family $\{\mathcal{D}(Q); Q \in \partial G_N\}$ is disjoint and

$$G^c_N = \bigcup_{Q \in \partial G_N} \mathcal{D}(Q)\ , \tag{114}$$

i.e, the family is a partition of G^c_N.

(b) The following is true

$$\operatorname{card}(\partial G_N) \leq 2^d N. \tag{115}$$

Proof: (a) Maximal cubes are pairwise disjoint. Hence, ∂G_N is a disjoint family.

Further, cubes $Q \in \partial G_N$ are roots of the trees $\mathcal{D}(Q)$ from (114). Since the roots are disjoint, the corresponding trees are as well, i.e., $\{\mathcal{D}(Q); Q \in \partial G_N\}$ is a disjoint family.

To prove that the family is a partition of G_N^c we check that every $Q' \in G_N^c$ belongs to some $\mathcal{D}(Q)$ where $Q \in \partial G_N$.

Let $Q' =: Q_1 \to Q_2 \to \cdots \to Q_d$ be the path joining Q' and Q^d, and $Q_i, i \geq 2$, be the smallest cube of the path belonging to G_N. Then its son Q_{i-1} belongs to G_N^c, i.e., Q_{i-1} is maximal, and $Q' \in \mathcal{D}(Q_{i-1})$ as required.

(b) Let Q^+ be the father of $Q \in \partial G_N$ and $(\partial G_N)^+ := \{Q^+; Q \in \partial G_N\}$. Since $Q^+$is unique, the set $(\partial G_N)^+$is disjoint.

Further, every father has 2^d sons and therefore

$$\operatorname{card}(\partial G_N) \leq 2^d \operatorname{card}(\partial G_N)^+. \tag{116}$$

Finally, (113), subadditivity of W and (105) imply

$$N^{-1} \operatorname{card}\,(\partial G_N)^+ < \sum_{Q \in (\partial G_N)^+} W(Q) \leq W(Q^d) = 1. \tag{117}$$

This and (116) give (115).

Finally, the required cover Δ_N is given by

$$\Delta_N := \{Q^d\} \bigcup \left(\bigcup_{B \in \mathcal{B}_N} \{T_B, H_B\} \bigcup \mathcal{D}_1(T_B) \right). \tag{118}$$

Due to (111)

$$\operatorname{card} \Delta_N \leq 1 + 2(3N+1) + 2^d(3N+1) =: c(d)N. \tag{119}$$

We define the required $g_N \in \mathcal{P}_{k-1}(\Delta_N)$ using to this end polynomials of best approximation determined by

$$\|f - m_s\|_{L_q(S)} = E_k(f; S; L_q). \tag{120}$$

Further, we use for brevity the following notations

$$M_Q := \sum_{Q' \in \mathcal{D}_1(Q)} m_{Q'} \cdot 1_{Q'} - m_Q \cdot 1_Q \quad , Q \in \mathcal{D}; \tag{121}$$

and, moreover,

$$B^+ := H_B, \quad B^0 := H_B \backslash T_B, \quad B^- := T_B. \tag{122}$$

Using this we write

$$g_N := m_{Q^d} + \sum_{B \in \mathcal{B}_N} [(m_{B^+} - m_{B^0}) \cdot 1_{B^+} + (m_{B^0} - m_{B^-}) \cdot 1_{B^-} + M_{B^-}]. \tag{123}$$

This clearly is a piecewise polynomial of degree $k-1$ over Δ_N, see (118).

Let us note that for B being a singleton $B^{\pm} = \{B\}, B^0 = \emptyset$, i.e., the corresponding terms in (123) and (118) equal M_B and $\{\{B\}, \mathcal{D}_1(\{B\})\}$, respectively.

Theorem (5.2.22)(a) will be derived from the next key result.

To introduce the family Δ_N we use the algorithm for the weight W given by (106).

Since W satisfies the assumptions of Proposition (5.2.16), see (103)–(105), it determines the finite set $G_N \subset \mathcal{D}$, and the algorithm gives the partition $\mathcal{B}_N$ of $G_N \backslash \{Q^d\}$ into the basic paths which in turn determines the required cover Δ_N, see (118) and (119).

To estimate $f - g_N$ we need a suitable presentation of this difference; the next lemmas are used for its derivation.

Lemma (5.2.18)[142]: Let $f \in L_q(Q) \cap C(\mathbb{R}^d)$, $1 \leq q \leq \infty, Q \in \mathcal{D}$. Then the following holds

$$f = m_Q + \sum_{Q' \in \mathcal{D}(Q)} M_{Q'} \tag{124}$$

with convergence in $L_q(Q)$.

Proof: Let $\mathcal{D}_j(Q), j \in \mathbb{Z}$, be the partition of Q into 2^{jd} congruent (dyadic) cubes, e.g., $\mathcal{D}_0(Q) = \{Q\}$ and $\mathcal{D}_1(Q)$ is the set of sons for Q. Then $P_j \in \mathcal{P}_{k-1}\left(\mathcal{D}_j(Q)\right)$ is defined by

$$P_j := \sum_{Q' \in \mathcal{D}_j(Q)} m_{Q'} \cdot 1_{Q'}. \tag{125}$$

We show that

$$f - m_Q = \sum_{j \geq 0} (P_{j+1} - P_j) \ \left(\text{convergence in } L_q(Q)\right). \tag{126}$$

Let s_n be the n-th partial sum of the series (126). Then

$$f - m_Q - s_n = f - P_{n-1} = \sum_{Q' \in \mathcal{D}_n(Q)} \left(f - m_{Q'}\right) \cdot 1_{Q'}.$$

This and (120) imply that

$$\left\|f - m_Q - s_n\right\|_q = \left\{\sum_{Q' \in \mathcal{D}_n(Q)} \left\|f - m_{Q'}\right\|_{L_q(Q')}^q\right\}^{1/q} = \left\{\sum_{Q' \in \mathcal{D}_n(Q)} E_k(f; Q'; L_q)\right\}^{1/q}.$$

By Theorem 4 of [149] the right–hand side is bounded by $c(k,d)\omega_k\left(f; \frac{|Q|^{1/d}}{2^n}; L_q(Q)\right)$. Since this bound tends to 0 as $n \to \infty$ for $q < \infty$ and for $q = \infty$ and $f \in C(\mathbb{R}^d)$, (126) is proved. Using now notations (121) and (125) we obtain

$$P_{j+1} - P_j = \sum_{Q' \in \mathcal{D}_j(Q)} M_{Q'}.$$

Summing over $j \geq 0$ and using (126) we then have

$$f - m_Q = \sum_{j \geq 0} \sum_{Q' \in \mathcal{D}_j(Q)} M_{Q'} = \sum_{Q' \in \mathcal{D}(Q)} M_{Q'}.$$

The proof is complete.

Now we apply (124) for $Q = Q^d$ and present $\mathcal{D} = \mathcal{D}(Q^d)$ as follows:

$$\mathcal{D} = \left(\sum_{B \in \mathcal{B}_N} \sum_{Q \in B} Q\right) \bigcup \left(\sum_{Q \in \partial G_N} \mathcal{D}(Q)\right),$$

see Proposition (5.2.16) and (114). This then implies the identity

$$f - m_{Q^d} = \sum_{B \in \mathcal{B}_N} \sum_{Q \in B} M_Q + \sum_{Q \in \partial G_N} \sum_{Q' \in \mathcal{D}(Q)} M_{Q'}.$$

Rewriting the second sum here by (124) we have

$$f - m_{Q^d} = \sum_{B \in \mathcal{B}_N} \sum_{Q \in B} M_Q + \sum_{Q \in \partial G_N} \left(f - m_Q\right) \cdot 1_Q.$$

Subtracting from here equality (123) for g_N we obtain the required presentation

$$f - g_N = \sum_{B \in \mathcal{B}_N} S_B + \sum_{Q \in \partial G_N} \left(f - m_Q\right) \cdot 1_Q\,; \qquad (127)$$

here we set

$$S_B = \sum_{Q \in B \setminus \{B-\}} M_Q - [(m_{B^+} - m_{B^0}) \cdot 1_{B^+} + (m_{B^0} - m_{B^-}) \cdot 1_{B^-}]. \qquad (128)$$

The next result gives the basic presentation of S_B.

Lemma (5.2.19)[142]: The following is true

$$S_B = \sum_{Q \in B \setminus \{B-\}} \sum_{Q' \in \mathcal{D}_1(Q) \setminus B} \left(m_{B^0} - m_{Q'}\right) \cdot 1_{Q'}. \qquad (129)$$

Proof: We begin with the identity

$$\sum_{Q \in B \setminus \{B-\}} M_Q = \sum_{Q \in B \setminus \{B-\}} \sum_{Q' \in \mathcal{D}_1(Q) \setminus B} \left[\left(m_{Q'} - m_{B^+}\right) \cdot 1_{Q'} + (m_{B^-} - m_{B^+}) \cdot 1_{B^-}\right]. \qquad (130)$$

proved by induction on card B.

Let $B := [Q_1, Q_n] = \{Q_1 \to Q_2 \to \cdots \to Q_n\}$, i.e., $B^- := Q_1, B^+ := Q_n$. Since $\mathcal{D}_1(Q) \setminus B$ for $Q \in B \setminus \{B^-\}$ consists of all sons of Q excluding the son belonging to B,

$$\mathcal{D}_1(Q_i) \setminus B = \mathcal{D}_1(Q_i)\{Q_{i-1}\}, i \geq 2.$$

Denoting the right–hand side by $\mathcal{D}^*(Q_i)$we then rewrite (130) as follows.

$$\sum_{i=2}^{n} M_{Q_i} = \sum_{i=2}^{n} \sum_{Q \in \mathcal{D}_1^*(Q_i)} \left[\left(m_Q - m_{Q_n}\right) \cdot 1_Q + \left(m_{Q_1} - m_{Q_n}\right) \cdot 1_{Q_1}\right] \qquad (131)$$

For $n = 2$ the right–hand side of (131) equals

$$\sum_{Q \in \mathcal{D}_1^*(Q_2)} \left[\left(m_Q - m_{Q_2}\right) \cdot 1_Q\right] + \left(m_{Q_1} - m_{Q_2}\right) \cdot 1_{Q_1}$$

$$:= \sum_{Q \in \mathcal{D}_1(Q_2)} m_Q \cdot 1_Q - m_{Q_2} \left(\sum_{Q \in \mathcal{D}_1^*(Q_2)} 1_Q + 1_{Q_1} \right).$$

Since $\mathcal{D}_1^*(Q_2)$is a partition of $Q_2 \setminus Q_1$, the sum in the parentheses equals $1_{Q_2 \setminus Q_1} + 1_{Q_1} = 1_{Q_2}$. Hence, the right–hand side here equals M_{Q_2}, see (121), as required.

Now let (130) hold for all paths of cardinality $n \geq 2$. To prove it for n + 1 we write (131) for the n-term path $\{Q_2 \to \ldots \to Q_{n+1}\}$and add to it (131) for $n = 2$ written equivalently as follows:

$m_{Q_2} =$

$$\sum_{Q \in \mathcal{D}_1^*(Q_2)} \left(m_Q - m_{Q_{n+1}}\right) \cdot 1_Q + \left(m_{Q_{n+1}} - m_{Q_2}\right) \cdot 1_{Q_2} + \left(m_{Q_1} - m_{Q_{n+1}}\right) \cdot 1_{Q_1}$$

Together with the equality

$$\sum_{i=3}^{n+1} M_{Q_i} = \sum_{i=3}^{n+1} \sum_{Q\in\mathcal{D}_1^*(Q_i)} \left(m_Q - m_{Q_{n+1}}\right)\cdot 1_Q + \left(m_{Q_2} - m_{Q_{n+1}}\right)\cdot 1_{Q_2} \quad (132)$$

this gives

$$\sum_{i=2}^{n+1} M_{Q_i} = \sum_{i=2}^{n+1} \sum_{Q\in\mathcal{D}_1^*(Q_i)} \left(m_Q - m_{Q_{n+1}}\right)\cdot 1_Q + R \quad (133)$$

where we set

$$R \coloneqq \left(m_{Q_{n+1}} - m_{Q_2}\right)\cdot 1_{Q_2} + \left(m_{Q_1} - m_{Q_{n+1}}\right)\cdot 1_{Q_1} + \left(m_{Q_2} - m_{Q_{n+1}}\right)\cdot 1_{Q_2}$$
$$= \left(m_{Q_1} - m_{Q_{n+1}}\right)\cdot 1_{Q_1}.$$

Hence, (133) proves the required equality (130) for $n + 1$.

Now we transform (131) by adding and subtracting $m_{B^0}(\coloneqq m_{Q_n\setminus Q_1})$.

This gives

$$\sum_{i=2}^{n+1} M_{Q_i} = \sum_{i=2}^{n} \sum_{Q\in\mathcal{D}_1^*(Q_i)} \left(m_Q - m_{B^0}\right)\cdot 1_Q + \left(m_{B^0} - m_{Q_n}\right)\sum_{i=2}^{n} \sum_{Q\in\mathcal{D}_1^*(Q_i)} 1_Q + \left(m_{Q_1} - m_{B^0}\right)$$
$$\cdot 1_{Q_1} + \left(m_{B^0} - m_{Q_n}\right)\cdot 1_{Q_1}.$$

Since the second sum here equals n $\sum_{i=2}^{n} 1_{Q_i\setminus Q_{i-1}} = 1_{Q_n\setminus Q_1}$ and, in the chosen notations, see (128),

$$S_B \coloneqq \sum_{i=2}^{n} M_{Q_i} - \left(m_{Q_n} - m_{B^0}\right)\cdot 1_{Q_n} - \left(m_{B^0} - m_{Q_1}\right)\cdot 1_{Q_1} \quad (134)$$

these two equalities give

$$S_B \coloneqq \sum_{i=2}^{n} \left[\sum_{Q\in\mathcal{D}_1^*(Q_i)} \left(m_Q - m_{B^0}\right)\cdot 1_Q\right] + R$$

where the remainder R equals

$$R \coloneqq \left[\left(m_{B^0} - m_{Q_n}\right)\cdot 1_{Q_n\setminus Q_1} + \left(m_{B^0} - m_{Q_n}\right)\cdot 1_{Q_1} + \left(m_{Q_1} - m_{B^0}\right)\cdot 1_{Q_1}\right]$$
$$- \left[\left(m_{Q_n} - m_{B^0}\right)\cdot 1_{Q_n} + \left(m_{B^0} - m_{Q_1}\right)\cdot 1_{Q_1}\right]. \quad (135)$$

Since the square parentheses here annihilate, $R = 0$.The identity (129) is proved.

Proposition (5.2.20)[142]: Let $f \in V_{pq}^k \cap C^\infty(\mathbb{R}^d)$ where $d, p, q, s = s(V_{pq}^k)$ satisfy (100) and (101). Given $N \in \mathbb{N}$ there exist a cover $\Delta_N \subset \mathcal{D}$ of Q^d and a piecewise polynomial $g_N \in \mathcal{P}_{k-1}(\Delta_N)$such that

$$\|f - g_N\|_q \le c(k,d)N^{-s/d} \quad (136)$$

and, moreover,

$$\operatorname{card} \Delta_N \le c(d)N. \quad (137)$$

Proof. We should prove that for $f \in V_{pq}^k \cap C^\infty(\mathbb{R}^d)$

$$\|f - g_N\|_q \le c(k,d)N^{-s/d}. \quad (138)$$

Due to the presentation (127)

$$\|f - g_N\|_q \le \left\|\sum_{B\in\mathcal{B}_N} S_B\right\|_q + \left\|\sum_{Q\in\partial G_N} \left(f - m_Q\right)\cdot 1_Q\right\|_q \quad (139)$$

and it remains to estimate each term of the sum.

Lemma (5.2.21)[142]: (a) Supports of the functions S_B, $B \in \mathcal{B}_N$, are disjoint.
(b) It is true that

$$\|S_B\|_q \le c(k,d)\operatorname{var}_p^k(f; B^+ \backslash B^-; L_q). \tag{140}$$

Proof. (a) Since $supp\, S_B = B^+ \backslash B^-$, see Lemma (5.2.19), the supports of S_B and $S_{\tilde{B}}$ are disjoint if their heads are. Otherwise, one of these (dyadic) cubes, say, $\tilde{B}^+$, embeds into the other. Then $\tilde{B}^+$ embeds into the tail B^- of the path B. Hence, $supp\, S_{\tilde{B}}$ does not intersect $supp\, S_B = B^+ \backslash B^-$.

(b) By identity (129)

$$S_B = \sum_{Q \in B^*} \sum_{Q' \in \mathcal{D}_1^*(Q)} \left(m_{B^0} - f + f - m_{Q'}\right) \cdot 1_{Q'}$$

where for brevity we set $B^* := B \backslash \{B^-\}$.
Further, we have

$$S_B = (m_{B^0} - f) \sum_{Q \in B^*} \sum_{Q' \in \mathcal{D}_1^*(Q)} 1_Q + \sum_{Q \in B^*} \sum_{Q' \in \mathcal{D}_1^*(Q)} (f - m_{Q'}) 1_{Q'}.$$

Since the family $\bigcup_{Q \in B^*} \mathcal{D}_1^*(Q)$ is a partition of $B^+ \backslash B^-$, the sum of indicators here equals $1_{B^+ \backslash B^-}$ and the equality implies

$$\|S_B\|_q \le \|f - m_{B^0}\|_{L_q(B^+ \backslash B^-)} + \left(\sum_{Q \in B^*} \sum_{Q' \in \mathcal{D}_1^*(Q)} \|f - m_{Q'}\|^q\right)^{1/q}$$

$$= E_k\left(f; B^+ \backslash B^-; L_q\right) + \left(\sum_{Q \in B^*} \sum_{Q' \in \mathcal{D}_1^*(Q)} E_k\left(f; Q'; L_q\right)\right)^{1/q}.$$

By the Jenssen inequality the second term is bounded by

$$\left(\sum_{Q \in B^*} \sum_{Q' \in \mathcal{D}_1^*(Q)} E_k\left(f; Q'; L_q\right)^p\right)^{1/p}.$$

Since the family $\bigcup_{Q \in B^*} \mathcal{D}_1^*(Q)$ is a partition of $B^+ \backslash B^-$, this sum is bounded by $\operatorname{var}_p^k(f; B^+ \backslash B^-; L_q)$, see the definition of (k,p)–variation in (78).

Moreover, by Theorem (5.2.11)

$$\|f - m_{B^0}\|_{L_q(B^0)} := E_k(f; B^+ \backslash B^-) \le c(k,d) \cdot \operatorname{var}_p^k\left(f; B^+ \backslash B^-; L_q\right).$$

Combining this with the previous inequality we obtain (140). Now we use Lemma (5.2.21) to estimate the first term in (139). We have

$$\left\|\sum_{B \in \mathcal{B}_N} S_B\right\|_q \le \left\{\sum_{B \in \mathcal{B}_N} \|S_B\|_q^q\right\}^{1/q} \le c(k,d) \left\{\sum_{B \in \mathcal{B}_N} \operatorname{var}_p^k\left(f; B^+ \backslash B^-; L_q\right)^q\right\}^{1/q}.$$

Moreover, by the definition of the weight W, see (106), and the inequality (110) of Proposition (5.2.16)

$$\operatorname{var}_p^k\left(f; B^+ \backslash B^-; L_q\right) := W(B^+ \backslash B^-)^{1/p} \le N^{-1/p}.$$

Combining with the previous inequality and using (111) we finally have the required estimate

$$\left\|\sum_{B\in\mathcal{B}_N} S_B\right\|_q \le c(k,d)\left(N^{-q/p}\operatorname{card}\mathcal{B}_N\right)^{1/q} \le c(k,d)\left(N^{-q/p}(3N+1)\right)^{1/q}$$
$$\le c_1(k,d)N^{-1/p+1/q} := c_1(k,d)N^{-s/d}.$$
It remains to obtain the similar bound for the sum over boundary ∂G_N in (139). Due to Proposition (5.2.17) and (113) ∂G_N is disjoint and, moreover,
$$\operatorname{var}_p^k(f;Q;L_q)^p =: W(Q) < N^{-1}$$
for every $Q \in \partial G_N$.

This immediately implies
$$\left\|\sum_{Q\in\partial G_N}(f-m_Q)\cdot 1_Q\right\|_q = \left\{\sum\|f-m_Q\|_{L_q(Q)}^q\right\}^{1/q} := \left\{\sum_{Q\in\partial G_N} E_k(f;Q;L_q)^q\right\}^{1/q}$$
$$\le \left\{\sum_{Q\in\partial G_N}\operatorname{var}_p^k(f;Q;L_q)^{q/p}\right\}^{1/p} \le N^{-1/p}(\operatorname{card} G_N)^{1/q}.$$
Since $\operatorname{card} G_N \le 2^d N$, see (115), this finally gives
$$\left\|\sum_{Q\in\partial G_N}(f-m_Q)\cdot 1_Q\right\|_q \le 2^{d/q}N^{-s/d}$$
as required.

Proposition (5.2.20) is proved.

Theorem (5.2.22)[142]: (a) Let $f \in V_{pq}^k(Q^d)$ where the smoothness $s := s(V_{pq}^k)$, see (72), and d,p,q be such that
$$d \ge 2, 0 < s \le k \ and \ 1 \le p < q < \infty. \tag{141}$$
Given $N \in \mathbb{N}$ there exist a cover $\Delta_N \subset \mathcal{D}(Q^d)$ of Q^d with $\operatorname{card}\Delta_N \le N$ and $g_N \in \mathcal{P}_{k-1}(\Delta_N)$ such that
$$\|f-g_N\|_q \le c(d)N^{-s/d}|f|_{V_{pq}^k}. \tag{142}$$
The same is true for $q = \infty$, i.e., for $s/d = 1/p$, if f is uniformly continuous on Q^d.

(b) The cover Δ_N can be replaced by a partition of Q^d into at most N dyadic d-rings.

Proof: (a). We derive the result from Theorem (5.2.9) and Proposition (5.2.20).

Let $Q := [1-\delta,\delta), \delta > 0$, and $f \in V_{pq}^k$ if $q < \infty$ and $f \in V_{pq}^k \cap C(\mathbb{R}^d)$ if $q = \infty$.Given $\varepsilon > 0$ Theorem (5.2.23) then yields a function $f_\varepsilon \in C^\infty(\mathbb{R}^d)$ such that
$$\|f-f_\varepsilon\|_{L_q(Q)} \le \varepsilon \tag{143}$$
and, moreover,
$$\operatorname{var}_p^k(f_\varepsilon;Q;L_q) \le |f|_{V_{pq}^k}. \tag{144}$$
Since Proposition (5.2.20) is homothety–invariant, it remains true for Q substituted for Q^d. Hence, given $N \in \mathbb{N}$ there exist a cover e $\tilde{\Delta}_N \subset \mathcal{D}(Q)$ of Q and a piecewise polynomiale $\tilde{g}_N \in \mathcal{P}_{k-1}(\tilde{\Delta}_N)$ such that
$$\|f_\varepsilon-\tilde{g}_N\|_{L_q(Q)} \le c(k,d)N^{-s/d}\operatorname{var}_p^k(f;Q;L_q) \tag{145}$$
and, moreover,
$$\operatorname{card}\tilde{\Delta}_N \le c(d)N. \tag{146}$$
Now let h be a homothety mapping Qonto Q^d, i.e.,

$$h(x) := \frac{x-\delta e}{1-2\delta}, \qquad x \in \mathbb{R}^d,$$

where $e := (1,\ldots,1)$.

Then $\Delta_N := h(\tilde{\Delta}_N) \subset \mathcal{D}(:= \mathcal{D}(Q^d))$ is a cover of Q^d satisfying

$$\text{card}\, \Delta_N = \text{card}\, \tilde{\Delta}_N \le c(d)N; \tag{147}$$

moreover, $g_N := \tilde{g}_N \circ h^{-1}$ is a piecewise polynomial from $\mathcal{P}_{k-1}(\Delta_N)$.

We will show that for $f \in V_{pq}^k$ with $q < \infty$ and for $f \in V_{p\infty}^k \cap C(\mathbb{R}^d)$

$$\|f - g_N\|_q \le c(k,d)N^{-s/d}|f|_{V_{pq}^k}; \tag{148}$$

this clearly implies Theorem (5.2.23)(a) for $N \ge c(d)$, see (147).

Let $h^*g := g \circ h^{-1}, g \in L_q(Q^d)$. Then $h^* : L_q(Q^d) \to L_q(Q)$ and $\|h^*\| = (1-2\delta)^{d/q}$.

Further, we write

$$\|f - g_N\|_q \le \|(f\circ h - f_\varepsilon)\circ h^{-1}\|_q + \|(f_\varepsilon - \tilde{g}_N)\circ h^{-1}\|_q$$
$$\le (1-2\delta)^{d/q}\Big(\|f - f_\varepsilon\|_{L_q(Q)} + \|f - f\circ h\|_{L_q(Q)} + \|f_\varepsilon - \tilde{g}_N\|_{L_q(Q)}\Big).$$

By (145) and (144) the third term in the parentheses is bounded by $c(k,d)N^{-s/d}|f|_{V_{pq}^k}$ while the first tends to 0 as $\varepsilon \to 0$, see (143), and the second does as $\delta \to 0$ for $q < \infty$, and also for $q = \infty$, if f is uniformly continuous on Q.

This proves (148) and +Theorem (5.2.23)(a) for $N \ge c(d)$.

To obtain the result for $1 \le N < c(d)$ we simply set $\Delta_N := \{Q^d\}$and $g_N := m_{Q^d}$.Then

$$\|f - g_N\|_q = E_k(f; Q^d; L_q) < c(d)^{s/d}N^{-s/d}|f|_{V_{pq}^k}$$

and, moreover, card $\Delta_N = 1 \le N$.

This gives Theorem (5.2.22)(a) for all $N \ge 1$.

(b). We establish the analog of Theorem (5.2.22)(a) with a partition of Q^d into d-rings of cardinality at most $c(d)N$.

To this end we write the piecewise polynomial g_N of Theorem (5.2.22)(a), see (123), in the form

$$g_N := m_{Q^d} + \sum_{Q\in\Delta_N} P_Q \cdot 1_Q \tag{149}$$

where $P_Q \in \mathcal{P}_{k-1}$ and

$$\Delta_N := \bigcup_{B\in\mathcal{B}_N} (\{H_B, T_B\} \bigcup \mathcal{D}_1(B^-)),$$

see (118).

First, we assume that Δ_N covers Q^d. If Δ_N is not a partition (otherwise, the result is clear), it contains at least one tower

$$T := \{Q_1 \underset{\neq}{\subset} \cdots \underset{\neq}{\subset} Q_n\} \subset \Delta_N.$$

This means that for every $0 \le i \le n$ there is no $Q \in \Delta_N$ such that $Q_i \underset{\neq}{\subset} Q \underset{\neq}{\subset} Q_{i+1}$; here $Q_0 := \emptyset, Q_{n+1} := Q^d$; hence, the bottom $Q_1 \neq \emptyset$and the top Q_n are, respectively, minimal and maximal cubes of T closest to Q^d.

According to this definition $(\Delta_N \backslash T) \cup \{Q_n\}$ still covers Q^d. Moreover, the tops of different towers do not intersect.

These, in particular, imply that if $T_j, 1 \le j \le m$, are all towers of Δ_N and $Q(T_j)$ are their tops, then

$$\left(\Delta_N \setminus \bigcup_{j=1}^{m} T_j\right) \bigcup \left(\bigcup_{j=1}^{m} Q(T_j)\right)$$

is a partition of Q^d. Hence, it suffices to subdivide each $Q(T_j)$ into a set of d-rings whose cardinality equals card T_j. We do this for $m = 1$ and then repeat the procedure for the remaining towers.

Now let $T := \{Q_1 \underset{\neq}{\subset} \dots \underset{\neq}{\subset} Q_n\}$ be the single tower of Δ_N. Setting

$$R_i := Q_i \setminus Q_{i-1}, \ \ 1 \le i \le n,$$

where $R_1 = Q_1$ as $Q_0 := \emptyset$, we obtain the partition $\mathcal{R}_n := \{R_i\}_{1 \le i \le n}$ of $Q_n := Q(T)$ into d-rings.

Further, we define the family of polynomials $\{P_{R_i}\} \subset \mathcal{P}_{k-1}$ given by$P_{R_i} := \left(\sum_{j=i}^{n} P_{Q_j}\right) \cdot 1_{R_j}, 1 \le i \le n$. These definitions imply the identity

$$\sum_{i=1}^{n} P_{Q_i} \cdot 1_{Q_i} = \sum_{R \in \mathcal{R}_n} P_R \cdot 1_R. \tag{150}$$

Moreover, the T is single in Δ_N, hence, $\mathcal{R}_n \cup (\Delta_N \setminus T)$ is a partition of Q^d into $\le n + (N - n) = N$ d-rings while the piecewise polynomial

$$\tilde{g}_N := m_{Q^d} + \sum_{R \in \mathcal{R}_n} P_R \cdot 1_R + \sum_{Q \in \Delta_N \setminus T} P_Q \cdot 1_Q$$

belongs to $\mathcal{P}_{k-1}(\mathcal{R}_n \cup (\Delta_N \setminus T))$ and equals g_N by (150) and (149).

This gives the result for Δ_N being a cover of Q^d.

Now suppose that Δ_N is not a cover of Q^d. Then $\mathcal{D}_1(Q^d) \cap G_N^c \neq \emptyset$, sinceotherwise $\mathcal{D}_1(Q^d) \subset G_N$, i.e., every son of Q^d is the head of a basic path. By thedefinition of Δ_N, see (118), this implies that $\mathcal{D}_1(Q^d) \subset \Delta_N$, i.e., Δ_N is a cover of Q^d, a contradiction.

Further, the set of heads $\mathcal{D}_1(Q^d) \cap G_N$is contained in $\Delta_N \subset \mathcal{D} \setminus \{Q^d\}$, and, moreover, it is nonempty as for otherwise $\Delta_N = \{Q^d\}$.

Hence, the set

$$\tilde{\Delta}_N := \Delta_N \bigcup (\mathcal{D}_1(Q^d) \bigcap G_N^c)$$

is a cover of Q^d and its cardinality is bounded by

$$N + \operatorname{card} \mathcal{D}_1(Q^d) - 1 = N + 2^d - 1 \le 2^d N$$

To complete the proof it suffices to modify the g_N to obtain e $\tilde{g}_N \in \mathcal{P}_{k-1}\left(\Delta_N \cup (\mathcal{D}_1(Q^d) \cap G_N^c)\right)$ such that

$$\|f - \tilde{g}_N\|_q \le c(k,d) N^{-s/d} |f|_{V_{pq}^k}. \tag{151}$$

We define $\tilde{g}_N$ by

$$\tilde{g}_N := g_N + \sum_{Q \in \mathcal{D}_1(Q^d) \cap G_N^c} (m_Q - m_{Q^d}) \cdot 1_Q$$

and then prove (151).

Substituting here g_N by the right–hand side of (149) and using the notations

$$S := \bigcup_{Q \in \Delta_N} Q, \qquad \tilde{\Delta} := \mathcal{D}_1(Q^d) \bigcap G_N^c$$

we have

$$\tilde{g}_N := g_N \cdot 1_S + \left(\sum_{Q\in\tilde{\Delta}} m_Q \cdot 1_Q\right) \cdot 1_{Q^d \setminus S}.$$

This, in turn, implies

$$\|f - \tilde{g}_N\|_q \leq \|f - g_N\|_q + \left\|\sum_{Q\in\tilde{\Delta}} (f - m_Q) \cdot 1_Q\right\|_q.$$

The first summand is clearly bounded by the right–hand side of (151).
Moreover, the second one equals

$$\left(\sum_{Q\in\tilde{\Delta}} E_k(f; Q; L_q)^q\right)^{1/q} \leq \left\{\sum_{Q\in\tilde{\Delta}} E_k(f; Q; L_q)^p\right\}^{1/p} \leq \left\{\sum_{Q\in\tilde{\Delta}} \mathrm{var}_p^k(f; Q; L_q)^p\right\}^{1/p}$$

$$=: \left\{\sum_{Q\in\tilde{\Delta}} W(Q)\right\}^{1/p}.$$

Since $\tilde{\Delta} \subset G_N^c$, every $W(Q) < N^{-1}$.

Theorem (5.2.23)[142]: (a) Given $f \in W_p^k([0,1)^d), k \in \mathbb{N}$ and $1 \leq p < q^* < \infty$ such that

$$\frac{k}{d} = \frac{1}{p} - \frac{1}{q^*}, d \geq 2,$$

there exist a cover Δ_N of $[0,1)^d$ by at most N dyadic subcubes and a family of polynomials $\{P_Q\}_{Q\in\Delta_N} \subset \mathcal{P}_{K-1}$ (of degree $k-1$) such that

$$\left\|f - \sum_{Q\in\Delta_N} P_Q . 1_Q\right\|_{q^*} \leq c(k,d) N^{-k/d} \sup_{|\alpha|=k} \|D^\alpha f\|_p. \tag{152}$$

(b) For $p := 1$, hence, $q^* = \frac{d}{d-k}$ the previous holds for $f \in L_1$, whose derivatives of order k are bounded Radon measures.

The associated seminorm of the latter function space denoted by $BV^k([0,1)^d)$ is given by

$$|f|_{BV^k} := \sup_{|\alpha|=k} \operatorname*{var}_{[0,1)^d} (D^\alpha f). \tag{153}$$

Proof. We obtain this result from Theorem (5.2.22) with $s(V_{pq}^k) = k$ and $q < \infty$. It asserts in this case that under the assumptions

$$d \geq 2, \quad 1 \leq p < q < \infty \text{ and } \frac{k}{d} = \frac{1}{p} - \frac{1}{q} \tag{154}$$

there exist a cover $\Delta_N \subset \mathcal{D}$ of Q^d of at most N cubes and $g_N \in \mathcal{P}_{k-1}(\Delta_N)$ such that

$$\|f - g_N\|_q \leq c(k,d) N^{-k/d} |f|_{V_{pq}^k}. \tag{155}$$

It remains to replace here $|f|_{V_{pq}^k}$ by the Sobolev seminorm $|f|_{W_p^k(Q^d)}$ if $p > 1$ and by the $BV^k(Q^d)$ seminorm if $p = 1$. This substitution is justified by the two–sided inequality

$$|f|_{V_{pq}^k} \approx \begin{cases} |f|_{W_p^k} & \text{if } p > 1, \\ |f|_{BV^k} & \text{if } p = 1, \end{cases} \tag{156}$$

where constants are independent of f, see Theorems 4 and 12 from [149]. The proof is complete.

Theorem (5.2.24)[142]: Let $f \in \dot{B}_p^\lambda$ and d, p, q, λ be such that

$$d \geq 2, 1 \leq p < q < \infty \quad \text{and} \quad \frac{\lambda}{d} = \frac{1}{p} - \frac{1}{q}.$$

Given $N \in \mathbb{N}$ there exist a cover $\Delta_N \subset \mathcal{D}$ of Q^d by at most N cubes and $g_N \in \mathcal{P}_{k-1}(\Delta_N)$ such that

$$\|f - g_N\|_q \leq c(k, d) N^{-\lambda/d} |f|_{B_p^\lambda}.$$

The second result concerns approximation in the uniform norm ($q = \infty$).

The named object is a set-function defined by (70) with Q^d substituted for a measurable set $S \subset \mathbb{R}^d$ of nonempty interior.

In fact, we replace osc_p^k by local best approximation, a set-function given for $f \in L_q^{loc}(\mathbb{R}^d)$ and $S \subset \mathbb{R}^d$by

$$E_k(f; S; L_q) := \inf_{m \in \mathcal{P}_{k-1}} \|f - m\|_{L_q(S)}.$$

Proof. We should prove the analog of the previous result for the homogeneous Besov $\dot{B}_p^\lambda(Q), \lambda > 0, Q \subset \mathbb{R}^d$, whose associated seminorm is given by

$$|f|_{B_p^\lambda(Q)} := \left\{ \int_0^{|Q|^{1/d}} \left(\frac{\omega_k\left(f; t; L_p(Q)\right)}{t^\lambda} \right)^\lambda \frac{dt}{t} \right\}^{1/p} \tag{157}$$

where $k = k(\lambda) := \min\{n \in \mathbb{N}; n > \lambda\}$.

We derive this from Theorem (5.2.22) with $s(V_{pq}^k) = \lambda, k = k(\lambda)$ and $q < \infty$. Hence,in this case,

$$1 \leq p < q < \infty, \quad \frac{\lambda}{d} = \frac{1}{p} - \frac{1}{q}, \tag{158}$$

and Theorem (5.2.22) gives under these assumptions the inequality

$$\|f - g_N\|_q \leq c(k, d) N^{-\lambda/d} |f|_{V_{pq}^k}. \tag{159}$$

with the corresponding $g_N \in \mathcal{P}_{k(\lambda)-1}(\Delta_N)$ and Δ_N.

It remains to replace here $|f|_{V_{pq}^k}$ by $|f|_{B_p^\lambda(Q^d)}$.

To this end we use the classical embedding theorem that under the assumptions (158) gives the inequality

$$E_k(f; Q; L_q) \leq c(d, \lambda, q) |f|_{B_p^\lambda(Q)} \tag{160}$$

see Remark (5.2.26) below for details.

Now let $\Delta := \{Q\}$ be a disjoint family of cubes from Q^d. Then (160) implies

$$\left(\sum_{Q \in \Delta} E_k(f; Q; L_q)^p \right)^{1/p} \leq c(d, \lambda, q) \left(\sum_{Q \in \Delta} \left(|f|_{B_p^\lambda(Q)} \right)^p \right)^{1/p}$$

Due to Lemma 2 from [27] the sum in the right–hand side is bounded by $c(d, q)|f|_{B_p^\lambda(Q^d)}$.

Taking supremum over Δ we then obtain the required inequality

$$|f|_{V_{pq}^k} \leq c(k, \lambda, q) |f|_{B_p^\lambda(Q^d)} \tag{161}$$

and prove Theorem (5.2.24).

Theorem (5.2.25)[142]: Let $f \in \dot{B}_p^{\lambda_1}$ and

$$d \geq 2, 1 \leq p < \infty \text{ and } \frac{\lambda}{d} = \frac{1}{p}.$$

Given $N \in \mathbb{N}$ there exist $\Delta_N \subset \mathcal{D}$ of satisfying the condition of Theorem (5.2.24) and $g_N \in \mathcal{P}_{k-1}(\Delta_N)$ such that

$$\|f - g_N\|_\infty \leq c(\lambda, d) N^{-\lambda/d} |f|_{B_p^{\lambda_1}}.$$

Proof. Now we deal with the homogeneous space $\dot{B}_p^{\lambda 1}(Q)$ whose associated seminorm is given by

$$|f|_{B_p^{\lambda 1}(Q)} := \int_0^{|Q|^{1/d}} \frac{\omega_k\left(f; t; L_p(Q)\right)}{t^{\lambda+1}} dt \qquad (162)$$

where $k = k(\lambda)$.

We prove, under the conditions

$$1 \leq p < q = \infty, \qquad d \geq 2 \text{ and } \frac{\lambda}{d} = \frac{1}{p}, \qquad (163)$$

existence of the corresponding Δ_N and $g_N \in \mathcal{P}_k(\Delta_N)$ such that the next inequality is true:

$$\|f - g_N\|_\infty \leq c(d, \lambda, p) N^{-\lambda/d} |f|_{B_p^{\lambda 1}(Q^d)}; \qquad (164)$$

here $k = k(\lambda)$.

Due to (163) $\lambda = \frac{d}{p} \leq d$ and therefore $k(\lambda) \leq d + 1$. Since norms $\|f\|_{B_p^{\lambda 1}(Q)} := \|f\|_{L_p(Q)} + |f|_{B_p^{\lambda 1}(Q)}$ with different $k \geq k(\lambda)$ are equivalent, it suffices to prove (164) for $k := d + 1$ instead of $k(\lambda)$.

We derive (164) from Theorem (5.2.22)(a) with $s\left(V_{pq}^k\right) = \lambda$ and $q = \infty$. This requires the embedding

$$\dot{B}_p^{\lambda 1}(Q^d) \subset V_{p\infty}^k(Q^d) \bigcap C(\mathbb{R}^d), \qquad (165)$$

because Theorem (5.2.22) with $q = \infty$ holds only for $f \in V_{p\infty}^k \cap C(\mathbb{R}^d)$.But $C(\mathbb{R}^d)$ in(165) can be removed as condition (163) implies that $B_p^{\lambda 1}(Q^d) \subset C(\mathbb{R}^d)|_{Q^d}$, see,e.g., [32]. By a reason explained later we begin with the case

$$\dot{B}_p^{\lambda 1}(\mathbb{R}^d) \subset V_{p\infty}^k(\mathbb{R}^d), \quad k = d + 1, \lambda = \frac{d}{p}. \qquad (166)$$

This will be proved for $p = 1$ and ∞ while the general case will be then derived from those by the method of real interpolation.

If $p = 1$, then (163) implies $\lambda = d$ and $k(\lambda) = d + 1$; moreover, by definition $\dot{B}_1^{\lambda 1} = \dot{B}_1^{\lambda}$. In this case (160) is still true, i.e., we have

$$E_{d+1}(f; Q; L_\infty) \leq c(k, d) |f|_{B_1^d(Q)}, \qquad (167)$$

see Remark (5.2.26) below.

Using the argument used in the proof of (161) we obtain from (167) the required inequality

$$|f|_{V_{1\infty}^{d+1}(\mathbb{R}^d)} \leq c(d) |f|_{B_1^d(\mathbb{R}^d)}$$

This proves (166) for $p = 1$.

Now let $p = \infty$, hence, $\lambda = \frac{d}{p} = 0$. The arising space $\dot{B}^{01}_{\infty}(\mathbb{R}^d)$ is defined by the seminorm

$$|f|_{B^{01}_{\infty}(\mathbb{R}^d)} := \sum_{j\in\mathbb{Z}} \|f * \varphi_j\|_{L_\infty(\mathbb{R}^d)}$$

where $\{\varphi_j\}$ is a sequence of test functions, satisfying, in particular, the condition

$$f = \sum_j f * \varphi_j$$

with convergence in the distributional sense, see, e.g., [32].

This implies

$$|f|_{V^{d+1}_{\infty\infty}(\mathbb{R}^d)} := \sup_{Q\subset\mathbb{R}^d} E_d(f;Q;L_\infty) \le \|f\|_{L_\infty(\mathbb{R}^d)} \le \sum_j \|f * \varphi_j\|_{L_\infty(\mathbb{R}^d)} = |f|_{B^{01}_{\infty}(\mathbb{R}^d)}$$

Hence, we prove (166) for $p = \infty$ as well.

Interpolating the embeddings obtained we then have

$$(\dot{B}^{01}_{\infty}, \dot{B}^{d1}_{1})_{\theta p} \subset \left(V^{d+1}_{\infty\infty}, V^{d+1}_{1\infty}\right)_{\theta p}; \tag{168}$$

hereafter $\mathbb{R}^d$ is omitted for brevity.

Taking $\theta := \frac{\lambda}{d} = \frac{1}{p}$ we obtain for the left–hand side the embedding

$$\dot{B}^{\lambda 1}_{p} \subset \left(\dot{B}^{01}_{\infty}, \dot{B}^{d1}_{1}\right)_{\theta p} \tag{169}$$

see [155].

Now we show that the right–hand side is contained in $V^{d+1}_{p\infty}(\mathbb{R}^d)$ with $p := \frac{\lambda}{d}$.

Let $\mathcal{E} : L_\infty(\mathbb{R}^d) \to l_\infty(\Delta)$ be a map given by

$$\mathcal{E}: f \mapsto (E_d(f;Q;L_\infty))_{Q\in\Delta};$$

here Δ is a disjoint family of cubes $Q \subset \mathbb{R}^d$.

By definition

$$\|\mathcal{E}(f)\|_{l_p(\Delta)} := \left(\sum_{Q\in\Delta} E_d(f;Q;L_\infty)^p\right)^{1/p} \le |f|_{V^{d+1}_{p\infty}(\mathbb{R}^d)},$$

i.e., $\mathcal{E}$ maps $V^{d+1}_{p\infty}(\mathbb{R}^d)$into $l_p(\Delta)$ and $\|\mathcal{E}\| \le 1, 1 \le p \le \infty$.

Interpolating this sublinear operator by the real method we obtain

$$\|\mathcal{E}(f)\|_{(l_\infty(\Delta), l_1(\Delta))_{\theta p}} \le |f|_{(V^{d+1}_{\infty\infty}, V^{d+1}_{1\infty})_{\theta p}},$$

see, e.g., [146] for validity of the interpolation result for sublinear operators. Moreover, $\left(l_\infty(\Delta), l_1(\Delta)\right)_{\theta p}$with $\theta = \frac{1}{p}$ equals $l_p(\Delta)$, see, e.g., [32]. Together with the previous this implies

$$\left(\sum_{Q\in\Delta} E_d(f;Q;L_\infty)^p\right)^{1/p} \le |f|_{(V^{d+1}_{\infty\infty}, V^{d+1}_{1\infty})_{\theta p}}$$

where $= \frac{1}{p} = \frac{\lambda}{d}$.

Taking here supremum over all Δ we obtain the embedding

$$\left(V^{d+1}_{\infty\infty}, V^{d+1}_{1\infty}\right)_{\theta p} \subset V^{d+1}_{p\infty},$$

implying the required embedding (166).

To derive from (166) the similar embedding for Q^d we use a bounded linear extension operator

$$Ext : \dot{B}_p^{\lambda 1}(Q^d) \to \dot{B}_p^{\lambda 1}(\mathbb{R}^d),$$

with $\|Ext\| \le c(\lambda, d)$, see, e.g., [143], and the restriction operator

$$Res : V_{p\infty}^{d+1}(\mathbb{R}^d) \to V_{p\infty}^{d+1}(Q^d).$$

Denoting the embedding operator in (166) by U and composing it with the now introduced ones we obtain the operator $U_{Q^d} := Ext\ o\ U\ o\ Res$ that embeds $\dot{B}_p^{\lambda 1}(Q^d)$ into $V_{p\infty}^{d+1}(Q^d)$with the embedding constant $\|Ext\| \le c(d, \lambda)$.

This proves the required inequality (165) and, therefore, Theorem (5.2.25).

Remark (5.2.26)[142]: We prove inequalities (160) and (167).

Let $f \in L_p(Q), 1 \le p < q \le \infty$, and $m_Q \in \mathcal{P}_{k-1}$ be the best approximation of f in $L_p(Q)$. Setting for brevity

$$\omega(t) := \omega_k\left(f; t; L_p(Q)\right), t > 0,$$

we estimate the nonincreasing rearrangement of $f - m_Q$ as follows

$$\left(f - m_Q\right)^*(t) \le c(k, d) \int_{t/2}^{|Q|} \frac{\omega\left(u^{1/d}\right)}{u^{1+1/p}} du, \quad t \le |Q|, \tag{170}$$

see [27].

TakingL_q–norm and applying the Hardy inequality we have

$$\left\|f - m_Q\right\|_{L_q(Q)} = \left\|\left(f - m_Q\right)^*\right\|_{L_q(0,|Q|)}$$

$$\le c(k, d)\left\|\mathcal{H}_{1/q}\right\| \left(\int_0^{|Q|} \left(\frac{\omega\left(u^{1/d}\right)}{u^{1/p-1/q}}\right)^q du\right)^{1/q} \tag{171}$$

where $\mathcal{H}_\mu, \mu > 0$, is the Hardy operator given by

$$\mathcal{H}_\mu g(t) := t^\mu \int_t^{|Q|} \frac{g(u)}{u^\mu} \frac{du}{u}.$$

Since $\left\|\mathcal{H}_\mu\right\| < \infty$ for $\mu > 0$, inequality (171) is true for $1/q > 0$, i.e., for $q < \infty$.

Since $\frac{1}{p} - \frac{1}{q} = \frac{\lambda}{d}$, the integral in (171) is bounded by

$$d^{1/q} \left(\int_0^{|Q|^{1/d}} \left(\frac{\omega(t)}{t^\lambda}\right)^q \frac{dt}{t}\right)^{1/q} \le c(k, \lambda) d^{1/q} \left(\int_0^{|Q|^{1/d}} \left(\frac{\omega(t)}{t}\right)^p \frac{dt}{t}\right)^{1/p} = c(k, d, \lambda)|f|_{B_p^\lambda(Q)}.$$

Hence, for $q < \infty$

$$\left\|f - m_Q\right\|_{L_q(Q)} \le c(k, d, \lambda)|f|_{B_p^\lambda(Q)}$$

which implies (160) as the left–hand side is clearly bigger than $E_k(f; Q; L_q)$.

For $q = \infty$ we pass in (170) to the limit as $t \to 0^+$ to obtain

$$\left\|f - m_Q\right\|_{L_\infty(Q)} = \lim_{t \to 0}\left(f - m_Q\right)^*(t) \le c(k, d, \lambda) \int_0^{|Q|} \frac{\omega\left(u^{1/d}\right)}{u^{1/p}} \frac{du}{u} = d \cdot c(k, d)|f|_{B_p^{\lambda 1}(Q)}$$

Hence, (167) follows.

First, let Q be a dyadic subcube of Q^* such that

$$\text{dist}(Q, \mathbb{R}^d \setminus Q^*) > 0. \qquad (172)$$

The general result will be reduced to this case.

Theorem (5.2.27). Let (172) hold. There exists a cover $\mathcal{K}$ of $Q^* \setminus Q$ by cubes [6] such that for every overlapping [7] pair $\{\mathcal{K}_1, \mathcal{K}_2\} \subset \mathcal{K}$

$$|\mathcal{K}_1 \cap \mathcal{K}_2| \geq \frac{1}{2} \min_{i=1,2} |\mathcal{K}_i| \qquad (173)$$

and, moreover,

$$\text{card}\, \mathcal{K} = 4(2^d - 1). \qquad (174)$$

Proof. Without loss of generality we assume that $Q^* = Q^d := [0,1)^d$. By (172) the dyadic cube Q is contained in one of sons of Q^d, say, in $[1/2e, e) := \prod_{i=1}^d [1/2, 1)$, $e := (1,1,\ldots,1)$. Denoting $Q := \prod_{i=1}^d [a_i, b_i)$ we, in particular, have

$$0 < 1 - a_i \leq 1/2, \quad 1 \leq i \leq d. \qquad (175)$$

Now let π denote a partition of Q^d by hyperplanes passing through the vertex $a \in Q$ and parallel to the coordinate hyperplanes. It consists of 2^d parallelotopes every of which consists a single vertex $\varepsilon \in \{0,1\}^d$ of Q^d. We enumerate elements of π by these vertices, so that $\pi := \{\Pi_\varepsilon\}_{\varepsilon \in \{0,1\}^d}$ and ε is contained in the closure of $Q^d \cap \Pi_\varepsilon$. Then Π_ε and $\Pi_{\varepsilon'}$ have a (unique) common face whenever $\varepsilon, \varepsilon'$ differ by a single coordinate. Moreover, the edge $[\varepsilon, \varepsilon')$ of Q^d is orthogonal to this face and intersects Π_ε and $\Pi_{\varepsilon'}$.

Let $G(\pi)$ denote a graph with the vertex set $\pi = \{\Pi_\varepsilon\}$and the edges consisting of pairs $\{\Pi_\varepsilon, \Pi_{\varepsilon'}\}$ with a common face. The bijection $\varphi : \Pi_\varepsilon \leftrightarrow \varepsilon$ is an isomorphism of $G(\pi)$ onto the hypercube graph Γ_d whose vertices and edges are those of the cube Q^d.

In fact, $\varepsilon = \varphi(\Pi_\varepsilon)$ and $\varepsilon' = \varphi(\Pi_{\varepsilon'})$ are joined by an edge in Γ_d whenever ε differs from ε' by a single coordinate, i.e., whenever Π_ε and $\Pi_{\varepsilon'}$ have a common face and therefore are joined by an edge in $G(\pi)$.

Further, the graph Γ_d has a Hamiltonian cycle, i.e., a cycle that visits each vertex of Γ_d exactly once, see, e.g., [153]. Therefore, $G(\pi)$ also has such a cycle denoted by $\mathcal{C}(\pi)$.

Now we apply this construction to the parallelotope $\Pi_e := \prod_{i=1}^d [a_i, 1)$ containing $Q = [a, b) := \prod_{i=1}^d [a_i, b_i)$ and the vertex b substituting for that of a. This gives a partition $\hat{\pi}$ of Π_e into 2^d parallelotopes one of which is Q. Then we enumerate them by the vertex set V of Π_e such thatb $\hat{\pi} = \{\Pi_v\}_{v \in V}$ and v belong to the closure of $\Pi_v \cap \Pi_e$, e.g., $\Pi_a = Q$.

Using the partition $\hat{\pi}$ we, as above, define the graph $G(\hat{\pi})$ isomorphic to Γ_d and denote by $\mathcal{C}(\hat{\pi})$the corresponding Hamiltonian cycle. Hence, $\Pi_v, \Pi_{v'}$ are neighbours in $\mathcal{C}(\hat{\pi})$ if they have a common face orthogonal to $[v, v']$.

Now we define a new graph G with the vertex set

$$V(G) := (\pi \setminus \{\Pi_e\}) \bigcup (\hat{\pi} \setminus \{\Pi_a\})$$

where $\Pi_a = Q$, and with the edge set $E(G)$ of two parts.

The first consists of edges from $G(\pi)$ and $G(\hat{\pi})$ such that both of their endpoints belong to either $\pi \setminus \{\Pi_e\}$or $\hat{\pi} \setminus \{\Pi_a\}$

The second part is as follows.

Let $\Pi_\varepsilon, \Pi_{\varepsilon'}$ from $\mathcal{C}(\pi)$ have common faces with $\Pi_e (\in \mathcal{C}(\hat{\pi}))$. Since $\Pi_e := [a, e)$, the vertex $a \in Q$ belongs to Π_ε and to $\Pi_{\varepsilon'}$. Therefore there exist parallelotopes Π_v and $\Pi_{v'}$ from $\hat{\pi}$ each having one of faces common with that of Q and another containing in Π_ε and $\Pi_{\varepsilon'}$, respectively.

Then the pairs $\{\Pi_\varepsilon, \Pi_v\}, \{\Pi_{\varepsilon'}, \Pi_{v'}\}$ from $V(G)$ form the remaining part of edges from $E(G)$.

It is now the matter of definition to check that

$$\mathcal{C} = \mathcal{C}_1 \bigcup \mathcal{C}_2 := (\mathcal{C}(\pi)\backslash\{\Pi_e\}) \bigcup (\mathcal{C}(\hat{\pi})\backslash\{Q\})$$

is a Hamiltonian cycle in G.

Now we construct the desired cover $\mathcal{K}$ of $Q^d \setminus Q$ beginning first with extension of each parallelotope of $\mathcal{C}_i, i = 1, 2$, to a cube contained in $Q^d \setminus Q$.

We begin with the set

$$\mathcal{C}_1 := \{\Pi_\varepsilon;\ \varepsilon \in \{0,1\}^d \setminus \{e\}\}$$

containing $2^d - 1$ elements.

Let $\Pi_\varepsilon := \prod_{i=1}^d [a_i^\varepsilon, b_i^\varepsilon)$ and l^ε be the maximal edgelength of Π_ε. Since by the definition of Π_ε every edge $[a_i^\varepsilon, b_i^\varepsilon)$ equals either $A_i := [0, a_i)$or $B_i := [a_i, 1)$ and $|A_i| \geq |B_i|$, see $(I.4)$, the maximal edge of Π_ε, say, $[a_{i_0}^\varepsilon, b_{i_0}^\varepsilon)$, has the form

$$\left[a_{i_0}^\varepsilon, b_{i_0}^\varepsilon\right) = A_{i_0} = \left[0, a_{i_0}\right). \tag{176}$$

Now we extend Π_ε to a cube replacing every edge $[a_i^\varepsilon, b_i^\varepsilon) = A_i$ by $\left[\hat{a}_i^\varepsilon, \hat{b}_i^\varepsilon\right) := \left[0, a_{i_0}\right)$ and every edge equal to B_i by $\left[\hat{a}_i^\varepsilon, \hat{b}_i^\varepsilon\right) := [1 - a_{i_0}, 1)$.

In this way, we obtain the cube

$$Q_\varepsilon := \prod_{i=1}^d \left[\hat{a}_i^\varepsilon, \hat{b}_i^\varepsilon\right) \subset Q^d$$

of edgelength a_{i_0} that contains Π_ε and, moreover, is contained in $Q^d \setminus \Pi_\varepsilon$.

In fact, the projections of Q_ε and Π_e on the x_{i_0} – axis are $\left[\hat{a}_i^\varepsilon, \hat{b}_i^\varepsilon\right) = \left[0, a_{i_0}\right)$, see (176), and $\left[a_{i_0}, 1\right)$ respectively, that do not intersect.

Thus, we have

$$\bigcup \mathcal{C}_1 = \bigcup_{\varepsilon \neq e} Q_\varepsilon, Q_\varepsilon \supset \Pi_\varepsilon, \ \ \varepsilon \neq e,$$

where $\bigcup \mathcal{C}_1 := \bigcup\{\Pi; \Pi \in \mathcal{C}_1\}$.

Further, we cover $\bigcup \mathcal{C}_2$ similarly. By definition

$$\mathcal{C}_2 = \left\{\Pi_v := \prod_{i=1}^d [a_i^v, b_i^v)\,; v \in V\backslash\{a\}\right\}$$

where $[a_i^v, b_i^v)$ equals either $A_i := [b_i, 1)$ or $B_i := [a_i, b_i)$.

Let us show that $|A_i| \geq |B_i|$. In fact, Q is a dyadic cube, say, $Q := 2^{-n}(\alpha + Q^d)$, $\alpha \in \mathbb{Z}_+^d$, and therefore $|B_i| = 2^{-n}$ while $|A_i| = 1 - b_i = 2^{-n}(2^n - \alpha_i - 1) \geq 2^{-n}$ as $b_i < 1$.

Then the maximal edge of Π_v, say, $\left[a_{i_0}^v, b_{i_0}^v\right) = B_i$ has the form

$$\left[a_{i_0}^v, b_{i_0}^v\right) = A_{i_0} = \left[b_{i_0}, 1\right)(177).$$

Now we extend Π_v replacing every $[a_i^v, b_i^v) = A_i$ by $\left[\hat{a}_i^v, \hat{b}_i^v\right) = [1 - l^v, 1)$ and every $[a_i^v, b_i^v) = B_i$ by $[b_i, b_i - l^v)$; here $l^v = 1 - b_{i_0}$ is the maximal edgelength of Π_v.

In this way, we obtain the cube

$$Q_v := \prod_{i=1}^{d} \left[\hat{a}_i^v, \hat{b}_i^v\right) \subset Q^d$$

of volume $(l^v)^d$ that contains Π_v and, moreover, is contained in $Q^d \setminus Q$.

In fact, the embedding $\Pi_v \subset Q_v$ follows from the inequality $a_i \geq \hat{a}_i^v := b_i - l^v$ equivalent to

$$|B_i| = b_i - a_i \leq |A_i| \leq l^v$$

Further, $Q_v \cap Q = \emptyset$, as the projections on the x_{i_0} – axis of these cubes $\left[a_{i_0}^v, b_{i_0}^v\right) = \left[b_{i_0}, 1\right)$ and $[a_{i_0}, b_{i_0})$, respectively, do not intersect.

Thus, we have

$$\bigcup \mathcal{C}_2 \subset \prod_{v \neq a} Q_v \subset Q \setminus Q^d \text{ and } \Pi_v \subset Q_v.$$

This gives the family $\mathcal{F} := \{Q_\varepsilon\} \cup \{Q_v\}$ of $2(2^d - 1)$ cubes that cover the d-ring $Q^d \setminus Q$ such that Q_ε, Q_v are uniquely defined by the corresponding $\Pi_\varepsilon \supset Q_\varepsilon$, $\Pi_v \supset Q_v$ from the Hamiltonian cycle $\mathcal{C}$.

Further, we enumerate the cycle $\mathcal{C}$ by integers to obtain

$$\mathcal{C} = \{\Pi_i;\ 1 \leq i \leq 2 \cdot 2^d - 1\}$$

where $\Pi_i := \Pi_1$ for $i = 2 \cdot 2^d - 1$, such that Π_i, Π_{i+1} are neighbours in $\mathcal{C}$. Hence, they adjoint to some edge of Q^d denoted by $[v_i, v_{i+1})$ such that a small shift along this edge of the smaller parallelotope remains in $\Pi_i \cup \Pi_{i+1} \subset Q^d \setminus Q$.

Now let $\{Q_i;\ 1 \leq i \leq 2 \cdot 2^d - 1\}$ where $Q_i := Q_1$ for $i = 2 \cdot 2^d - 1$ and the numeration of the family $\mathcal{F}$ is induced by that of $\mathcal{C}$.

Then by the definition of cubes from $\mathcal{F}$ the following is true.

(a) $\cup_i Q_i$ covers $Q^d \setminus Q$;

(b) cubes $Q_i \supset \Pi_i, Q_{i+1} \supset \Pi_{i+1}$ adjoint to the edge $[v_i, v_{i+1})$and the shift along this edge of the smaller one, say Q_i, by its length remains in $Q_{i+1} \subset Q^d \setminus Q$.

Let then $Q_{i+1/2}$denote the image of Q_i under such a shift by the one–half of its length. Then the cover $\mathcal{K} := \{Q_i, Q_{i+1/2}\}$ of $Q^d \setminus Q$ consists of $4(2^d - 1)$ cubes satisfying the inequality

$$|Q_j \cap Q_{i+1/2}| \geq 1/2 \min\{|Q_j|, |Q_{i+1/2}|\}$$

for $j = i, i + 1$.

Hence, Theorem (5.2.27) is proved for Q contained in the interior of Q^d, see (172).

We describe the algorithm giving as output the cover Δ_N in Theorem (5.2.22). In what follows, we freely use terms and definitions, e.g., weight, dyadic tree $\mathcal{D} := \mathcal{D}(Q^d)$, paths etc. Proofs of some statements below will be left to the reader (all of them are presented in details in [24]).

Let $W : A(\mathcal{D}) \to \mathbb{R}_+$ be a subadditive absolutely continuous weight normed by the condition

$$W(Q^d) = 1. \tag{178}$$

Then the set

$$G_N := \{Q \in \mathcal{D}; W(Q) \geq N^{-1}\}, N \in \mathbb{N}, \tag{179}$$

is a finite rooted subtree of $\mathcal{D}$ with the root Q^d. Hence, every path connecting $Q \in G_N$ and Q^d is unique and belongs toG_N.

Further, let $G_N^{\min}$ be the set of minimal elements of G_N with respect to the set–inclusion order.

Hence, every $Q \in G_N$ contains properly some minimal cube and a son Q' of such a cube satisfies

$$W(Q') < N^{-1}.$$

In particular, $G_N^{\min}$ is disjoint and as every disjoint subset of G_N has at most N elements.

Somehow enumerating $G_N^{\min}$, say,

$$G_{\min}^N := \{Q_i\}_{1\le i\le m_N}$$

where

$$m_N := card\ G_N^{\min} \le N, \tag{180}$$

we then denote by L_i a (unique) path in G_N joining Q_i and Q^d.

By the definition of $G_N^{\min}$

$$G_N = \bigcup_{i=1}^{m_N} L_i. \tag{181}$$

We divide each L_i into more small paths

$$P_i := L_i \backslash \bigcup_{j=0}^{i-1} L_j , 1 \le i \le m_N$$

where $L_0 := \{Q^d\}$.

Lemma (5.2.28)[142]: ([24]) (a) Family $\{P_i\}_{1\le i\le m_N}$ is a partition of $G_N \backslash\{Q^d\}$.

(b) Every Pi is of the form

$$P_i := [Q_i, Q_i^c) := [Q_i, Q_i^c]\backslash\{Q_i^c\} \tag{182}$$

where Q_i^c is the tail of a path $L_i \cap L_j$ with $j < i$.

The set

$$\mathcal{C}_N := \{Q^d\} \bigcup \{Q_i^c\}_{1\le i\le m_N}$$

contains at most $m_N + 1$ elements called contact cubes.

Now we refine G_N subdividing each P_i by contact cubes from $P_i \cap \mathcal{C}_N$. In this way, we define a set of subpaths $[Q', Q'')$ where Q' is either a minimal cube or a contact cube, and Q'' is a contact cube.

Denoting the set of these subpaths by $\mathcal{P}_N$ we obtain from (180)

$$card\ \mathcal{P}_N \le 2m_N + (m_N + 1) = 3m_N + 1. \tag{183}$$

Finally, we divide each path $P \in \mathcal{P}_N$ in the required basic paths. To this end, we use an auxiliary weight defined on paths $P = [T_P , H_P]$ of D by

$$\widetilde{W}(P) := W(H_P \backslash T_P). \tag{184}$$

Now we define for each $P \in \mathcal{P}_N$ a family of vertices (cubes) $\{Q_i(P) \in P; 1 \le i \le i_p\}$ using induction on i.

We begin with $Q_1(P) := T_P$ and then having $Q_i(P)$ define $Q_{i+1}(P)$ as a vertex in the half–open from the left path

$$(Q_i(P), H_P] := [Q_i(P), H_P] \backslash \{Q_i(P)\}$$

satisfying the conditions

$$\widetilde{W}([Q_i(P), Q_{i+1}(P)]) \ge N^{-1},$$
$$\widetilde{W}([Q_i(P), Q_{i+1}(P))) < N^{-1}.$$

Then we define the i-th basic path $B_i(P)$ by setting

$$B_i(P) := \left[Q_i(P), Q_{i+1}(P)\right). \quad (185)$$

The vertex $Q_{i+1}(P)$ may be undetermined, if

$$\widetilde{W}([Q_i(P), H_P]) < N^{-1}$$

In this case, we complete induction setting $i_P := i$ and defining $B_i(P)$ to be equal to $[Q_i(P), H_P]$. However, to preserve formula (185) for this case, we define $Q_{i+1}(P)$ as the father of H_P . Denoting it, say, $H_P^+(\in P)$ we define $B_i(P)$for this case by $(II.8)$ with $Q_{i+1}(P) \coloneqq H_P^+$ and $\coloneqq i_P$.

Hence, the induction has been completed with $Q_{i+1}(P) = H_P$ or $Q_{i+1}(P) = H_P^+$ for $= i_P$. In this way, we obtain a partition of P by subpaths $B_i(P) \coloneqq \left[Q_i(P), Q_{i+1}(P)\right), 1 \leq i \leq i_P$. Let us single out that if $Q_{i+1} = H_P^+$, then $B_i(P)$ may be a singleton $\{H_p\}$.

By definition these subpaths satisfy

$$\widetilde{W}([Q_i(P), Q_{i+1}(P)]) < N^{-1},$$
$$\widetilde{W}([Q_i(P_i), Q_{i+1}(P))) \geq N^{-1} \quad (186)$$

for $1 \leq i \leq i_P - \varepsilon_P$ where $\varepsilon_P := 0$ if $Q_{i_P+1} = H_P^+$ and $\varepsilon_P := 1$ otherwise; in the first case, only the first of inequalities (186) holds.

Collecting all the basic paths we obtain the refinement of $\mathcal{P}_N$ given by

$$\mathcal{B}_N := \{B_i(P);\ 1 \leq i \leq i_P, P \in \mathcal{P}_N\}. \quad (187)$$

The next result (Proposition (5.2.16)) gives the output of the algorithm.

Proposition (5.2.29)[142]: (a) $\mathcal{B}_N$ is a partition of $G_N \setminus \{Q^d\}$.

(b) For every $B = [T_P, H_P] \in \mathcal{B}_N$

$$W(H_P \setminus T_P) := \widetilde{W}(B) < N^{-1}. \quad (188)$$

(c) The following is true

$$\text{card } \mathcal{B}_N \leq 3N + 1. \quad (189)$$

Proof. (a) $\mathcal{B}_N$ is a refinement of the partition $\mathcal{P}_N$, hence, it is also a partition.

(b) is given by the first inequality in (186) and the definition of $B_i(P)$.

(c) Let $\{P_i\}$ be a strictly monotone sequence of subpaths in a path P, i.e.. the head of P_i is a proper subset of the tail of P_{i+1}. Then by the definition of $\widetilde{W}$, see (184),

$$\sum_i \widetilde{W}(P_i) \leq W(H_P \setminus T_P).$$

Now let $B_i(P) := \left[Q_i(P), Q_{i+1}(P)\right), 1 \leq i \leq i_P$, be the partition of $P \in \mathcal{P}_N$ into the basic paths. By the second inequality (186)

$$(i_P - \varepsilon_P)N^{-1} \leq \sum_{i=1}^{i_P - \varepsilon_P} \widetilde{W}([Q_i, Q_{i+1}]). \quad (190)$$

Since the sequence $\{[Q_i, Q_{i+1}]\}_{1 \leq i \leq i_P - \varepsilon_P}$ has multiplicity 2, it can be divided into two strictly monotone subsequences. Hence, the right–hand side of (190) is bounded by $2W(H_P \setminus T_P)$. This implies

$$\text{card } \mathcal{B}_N = \sum_{P \in \mathcal{P}_N} i_P \leq 2N \sum_{P \in \mathcal{P}_N} W(H_P) + \sum_{P \in \mathcal{P}_N} \varepsilon_P$$

Since the set $\{H_P\}_{P \in \mathcal{P}_N}$ is disjoint, $\sum_{P \in \mathcal{P}_N} W(H_P) \leq W(Q^d) = 1$.

Further, $\varepsilon_P = 1$ if and only if the endpoint of $B_i(P)$ with $i = i_P$ is H_P^+ . By the definition of $\mathcal{P}_N$ every head H_P of $P \in \mathcal{P}_N$ is a contact cube. Hence,

$$\sum_{P \in \mathcal{P}_N} \varepsilon_P \leq \text{card}\, \mathcal{C}_N \leq m_N + 1 \leq N + 1,$$

see (180).
Combining this with the previous estimates we finally get

$$\text{card}\, \mathcal{B}_N \leq 2N + N + 1 = 3N + 1.$$

Chapter 6
Maximal Function Characterization for Hardy Spaces and of $H^1_{\Delta_N}(\mathbb{R}^n)$ and $\mathrm{BMO}_{\Delta_N}(\mathbb{R}^n)$

We show the characterizations of Hardy spaces associated to self-adjoint operator, via atomic decomposition or the nontangential maximal functions. The proof is based on a modification of a technique due to A. Calderón [163]. While for the space $\mathrm{BMO}_{\Delta_N}(\mathbb{R}^n)$ (which contains the classical $\mathrm{BMO}(\mathbb{R}^n)$) we show that it can be characterized in terms of the action of the Riesz transforms associated to the Neumann Laplacian on $L^\infty(\mathbb{R}^n)$ functions and in terms of the behavior of the commutator with the Riesz transforms. The results obtained extend many of the fundamental results known for $H^1(\mathbb{R}^n)$ and $\mathrm{BMO}(\mathbb{R}^n)$.

Section (6.1): Associated to Nonnegative Self-Adjoint Operators Satisfying Gaussian Estimates

We development of Hardy spaces on Euclidean spaces $\mathbb{R}^n$ in the 1960s played an important role in modern harmonic analysis and applications in partial differential equations. Let us recall the definition of the Hardy spaces (see [165], [167], [74], [176], [177], [19], [84]). Consider the Laplace operator $\Delta = -\sum_{i=1}^n \partial_{x_i}^2$ the Euclidean spaces $\mathbb{R}^n$. For $0 < p < \infty$, the Hardy space $H^p(\mathbb{R}^n)$is defined as the space of tempered distribution $f \in L^1(\mathbb{R}^n)$ for which the area integral function of f satisfying

$$Sf(x) := \left(\int_0^\infty \int_{|y-x|<t} \left|t^2 \Delta e^{-t^2\Delta} f(y)\right|^2 \frac{dydt}{t^{n+1}}\right)^{1/2} \tag{1}$$

belongs to $L^p(\mathbb{R}^n)$. If this is the case, define

$$\|f\|_{H^p(\mathbb{R}^n)} := \|Sf\|_{L^p(\mathbb{R}^n)} \tag{2}$$

When $p > 1$, $H^p(\mathbb{R}^n) = L^p(\mathbb{R}^n)$. For $p \le 1$, the space $H^p(\mathbb{R}^n)$ involves many different characterizations. For example, if $f \in L^1(\mathbb{R}^n)$, then

$$\begin{aligned} f \in H^p(\mathbb{R}^n) &\overset{(i)}{\Leftrightarrow} \sup_{t>0}\left|e^{-t^2\Delta} f(x)\right| \in L^p(\mathbb{R}^n) \\ &\overset{(ii)}{\Leftrightarrow} \sup_{|y-x|<t}\left|e^{-t^2\Delta} f(y)\right| \in L^p(\mathbb{R}^n) \\ &\overset{(iii)}{\Longleftrightarrow} f \ has\ a\ (p,q)\ atomic\ decomposition \\ & f = \sum_{j=0}^\infty \lambda_j a_j \ with \ \sum_{j=0}^\infty |\lambda_j|^p < \infty \end{aligned} \tag{3}$$

Recall that a function a supported in ball B of $\mathbb{R}^n$ is called a (p,q) $-$atom, $0 < p \le 1 \le q \le \infty$, $p < q$, if $\|a\|_{L^q(B)} \le |B|^{\frac{1}{q}-\frac{1}{p}}$, and $\int_B x^\alpha a(x)dx = 0$, where α is a multi-index of order $|\alpha| \le \left[n(\frac{1}{p}-1)\right]$, the integer part of $n(\frac{1}{p}-1)$(see [165], [176], [19]).

The theory of classical Hardy spaces has been very successful and fruitful in the past decades. However, there are important situations in which the standard theory of Hardy spaces is not applicable, including certain problems in the theory of partial differential equation which involve generalizing the Laplacian. There is a need to consider Hardy spaces that are adapted to a linear operator L, similarly to the way that the standard theory of Hardy spaces isadapted to the Laplacian. See [58], [160], [161], [168], [71], [169], [171], [172], [173], [174], [53].

We assume that L is a densely-defined operator on $L^2(\mathbb{R}^n)$and satisfies the following properties:

(H1) L is a non-negative self-adjoint operator on $L^2(\mathbb{R}^n)$;

(H2)The kernel of e^{-tL}, denoted by $p_t(x,y)$, is a measurable function on $\mathbb{R}^n \times \mathbb{R}^n$ and satisfies a Gaussian upper bound, that is

$$(GE) \quad |p_t(x,y)| \leq Ct^{-n/2} \exp\left(-\frac{|x-y|^2}{ct}\right)$$

for all $t > 0$, and $x, y \in \mathbb{R}^n$, where C and c are positive constants.

Given a function $f \in L^2(\mathbb{R}^n)$, consider the following area function $S_L f$associated to the heat semigroup generated by L

$$S_L f(x) := \left(\int_0^\infty \int_{|y-x|<t} \left|t^2 \mathrm{L} e^{-t^2 \mathrm{L}} f(y)\right|^2 \frac{dydt}{t^{n+1}}\right)^{1/2}, x \in \mathbb{R}^n \qquad (4)$$

Under the assumptions (H1) and (H2) of an operator L, it is known that the null space $N(L) = \{0\}$(see [171]) and the function S_L is bounded on $L^p(\mathbb{R}^n)$, $1 < p < \infty$ and

$$\|S_L f\|_{L^p(\mathbb{R}^n)} \simeq \|f\|_{L^p(\mathbb{R}^n)}$$

See [159], [58].

Definition (6.1.1)[158]: Suppose that an operator L satisfies(H1)–(H2). Given $0 < p \leq 1$. The Hardy space $H^p_{L,s}(\mathbb{R}^n)$ is defined as the completion of $\{f \in L^2(\mathbb{R}^n) : \|SLf\|_{L^p(\mathbb{R}^n)} < \infty\}$ with norm

$$\|f\|_{H^p_{L,s}(\mathbb{R}^n)} := \|SLf\|_{L^p(\mathbb{R}^n)}$$

To describe an atomic character of the Hardy spaces, let us recall the notion of a (p, q, M)-atom associated to an operator L [160], [168], [171].

Definition (6.1.2)[158]: Given $0 < p \leq 1 \leq q \leq \infty$, $p < q$ and $M \in \mathbb{N}$, a function $a \in L^2(\mathbb{R}^n)$is called a (p, q, M) $-$atom associated to the operator L if there exist a function $b \in D(L^M)$and a ball $B \subset \mathbb{R}^n$ such that

(i) $a = L^M b$;

(ii) $supp L^k b \subset B, k = 0,1,\ldots, M$;

(iii) $\left\|(r_B^2 L)^k b\right\|_{L^q(\mathbb{R}^n)} \leq r_B^{2M} |B|^{1/q-1/p}, k = 0, 1, \ldots, M.$

The atomic Hardy space $H^p_{L,\text{at,q,M}}(\mathbb{R}^n)$ is defined as follows.

Definition (6.1.3)[158]: We will say that $f = \sum \lambda_j a_j$ is an atomic (p, q, M)-representation (of f) if $\{\lambda_j\}_{j=0}^\infty \in \ell^p$, each a_j is a (p, q, M) $-$atom, and the sum converges in $L^2(\mathbb{R}^n)$. Set

$$H^p_{L,\text{at,q,M}}(\mathbb{R}^n) := \{f : f \textit{ has atomic } (p, q. M) - \textit{representation}\}$$

with the norm $\|f\|_{H^p_{L,\text{at,q,M}}}(\mathbb{R}^n)$given by

$$\inf\left\{\left(\sum_{j=0}^\infty |\lambda_j|^p\right)^{1/p} : f = \sum_{j=0}^\infty \lambda_j a_j \textit{ is an atomic } (p, q. M) - \textit{representation}\right\}$$

The space $H^p_{L,\text{at,q,M}}(\mathbb{R}^n)$ is then defined as the completion of $H^p_{L,\text{at,q,M}}(\mathbb{R}^n)$ with respect to this norm.

Obviously, $H^p_{L,\text{at},q_2,\text{M}}(\mathbb{R}^n) \subseteq H^p_{L,\text{at},q_1,\text{M}}(\mathbb{R}^n)$ when $1 \leq q_1 \leq q_2 \leq \infty$. Under the assumption that an operator L satisfies conditions(H1)–(H2), S.Hofmann, $G. Lu$, D.Mitrea,

M.Mitrea and obtained a $(1,2,M)$ −atomic decomposition of the Hardy space $H^1_{L,\mathrm{S}}(\mathbb{R}^n)$, and showed that for every integer $M \geq 1$, the spaces $H^1_{L,\mathrm{S}}(\mathbb{R}^n)$ and $H^1_{L,\mathrm{at},2,\mathrm{M}}(\mathbb{R}^n)$ coincide (see [171]). In particular,

$$\|f\|_{H^1_{L,\mathrm{S}}}(\mathbb{R}^n) \approx \|f\|_{H^1_{L,\mathrm{at},2,\mathrm{M}}}(\mathbb{R}^n)$$

A proof of an equivalence between the spaces $H^p_{L,\mathrm{S}}(\mathbb{R}^n)$ and $H^p_{L,\mathrm{at},2,\mathrm{M}}(\mathbb{R}^n)$for $p < 1$was shown by Duong and Li in [168], and by Jiang and Yang in [175].

Given a function $f \in L^2(\mathbb{R}^n)$, consider the non-tangential maximal function associated to the heat semigroup generated by the operator L,

$$f^*_L(x) \coloneqq \sup_{|y-x|<t} \left|e^{-t^2L}f(y)\right|$$

We may define the spaces $H^p_{L,\max}(\mathbb{R}^n)$, $0 < p \leq 1$as the completion of $\{f \in L^2(\mathbb{R}^n) : \|f^*_L\|_{L^p(\mathbb{R}^n)} < \infty\}$with respect to L^p-norm of the non-tangential maximal function; i.e.,

$$\|f\|_{H^p_{L,\max}(\mathbb{R}^n)} \coloneqq \|f^*_L\|_{L^p(\mathbb{R}^n)}$$

It can be verified (see [171], [168]) that for all $q > p$ with $1 \leq q \leq \infty$ and every number $M > \frac{n}{2}(\frac{1}{p} - 1)$, any (p,q,M)-atom ais in $H^p_{L,\max}(\mathbb{R}^n)$ and so the following continuous inclusion holds:

$$H^p_{L,\mathrm{at,q,M}}(\mathbb{R}^n) \subseteq H^p_{L,\max}(\mathbb{R}^n) \quad (5)$$

A natural question is to show the following continuous inclusion: $H^p_{L,\max}(\mathbb{R}^n) \subseteq H^p_{L,\mathrm{at,q,M}}(\mathbb{R}^n)$. It is known that the inclusion $H^p_{L,\max}(\mathbb{R}^n) \subseteq H^p_{L,\mathrm{at,q,M}}(\mathbb{R}^n)$. holds for certain operators including Schrödinger operators with nonnegative potentials via particular PDE technique (see [169], [74], [171], [172]). However, this question is still open assuming merely that an operator L satisfies(H1)–(H2). We give an affirmative answer to this question to get an atomic decomposition directly from $H^p_{L,\max}(\mathbb{R}^n)$.

We should mention that using the theory of tent spaces, a $(p,2,M)$-atomic decomposition of the Hardy space $H^p_{L,\mathrm{S}}(\mathbb{R}^n)$in terms of area functions was given in [168], [171]. We shall use a different argument to build a (p,∞,M)-atomic decomposition of the Hardy spaces $H^p_{L,\max}(\mathbb{R}^n)$ in terms of maximal functions. Our proof is based on a modification of a technique due to A.

Calderón [163], where a decomposition of the function $F(x,t) = f * \varphi_t(x)$ associated with the distribution f was given, and convolution operation of the function F played an important role in the proof. In our setting, there is, however, no analogue of convolution operation of the function $t^2Le^{-t^2L}f(x)$, we have to modify Calderón's construction and the geometry in conducting the analysis (see Fig.1). On the other hand, we do not assume that the heat kernel $p_t(x,y)$ satisfies the standard regularity condition, thus standard techniques of Calderón–Zygmund theory ([164], [19]) are not applicable. The lacking of smoothness of the kernel will be overcome in Proposition (6.1.7) below by using some estimates on heat kernel bounds, finite propagation speed of solutions to the wave equations and spectral theory of non-negative self-adjoint operators.

Throughout, the letter "c" and "C" will denote (possibly different) constants that are independent of the essential variables.

Recall that, if L is a nonnegative, self-adjoint operator on $L^2(\mathbb{R}^n)$, and $E_L(\lambda)$ denotes a spectral decomposition associated with L, then for every bounded B orel function $F: [0, \infty) \to \mathbb{C}$, one defines the operator $F(L): L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ by the formula

$$F(L) := \int_0^\infty F(\lambda) dE_L(\lambda) \tag{6}$$

In particular, the operator $cos(t\sqrt{L})$is then well-defined on $L^2(\mathbb{R}^n)$. Moreover, it follows from [166]that the integral kernel $K_{cos(t\sqrt{L})}$ of $cos(t\sqrt{L})$ satisfies

$$\operatorname{supp} K_{cos(t\sqrt{L})} \subseteq \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : |x - y| \le t\} \tag{7}$$

the Fourier inversion formula, whenever F is an even bounded Borel function with the Fourier transform of F, $\hat{F} \in L^1(\mathbb{R})$, we can write $F(\sqrt{L})$in terms of $cos(t\sqrt{L})$. Concretely, by recalling (6) we have

$$F(\sqrt{L}) = (2\pi)^{-1} \int_{-\infty}^{\infty} \hat{F}(t) \cos(t\sqrt{L})\, dt$$

which, when combined with (7), gives

$$K_{F(\sqrt{L})}(x, y) = (2\pi)^{-1} \int_{|t| \ge |x-y|} \hat{F}(t) K_{\cos(t\sqrt{L})}(x, y) dt \tag{8}$$

Lemma (6.1.4)[158]: Let $\varphi \in C_0^\infty(\mathbb{R})$ be even, $supp\varphi \subset (-1, 1)$. Let Φ denote the Fourier transform of φ. Then for every $\kappa = 0, 1, 2, \dots$, and for every $t > 0$, the kernel $K_{(t^2L)^\kappa \Phi(t\sqrt{L})}(x, y)$ of the operator $(t^2L)^\kappa \Phi(t\sqrt{L})$ which was defined by the spectral theory, satisfies

$$\operatorname{supp} K_{(t^2L)^\kappa \Phi(t\sqrt{L})} \subseteq \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n : |x - y| \le t\} \tag{9}$$

and

$$\left|K_{(t^2L)^\kappa \Phi(t\sqrt{L})}\right| \le Ct^{-n} \tag{10}$$

for all $x, y \in \mathbb{R}^n$.

Proof: For the proof, we refer it to [170]and [171].

Lemma (6.1.5)[158]: Assume that an operator L satisfies(H1)–(H2). Let $R > 0$, $s > 0$. Then for any $\epsilon > 0$, there exists a constant $C = C(s, \epsilon)$ such that

$$\int_{\mathbb{R}^n} \left|K_{F(\sqrt{L})}(x, y)\right|^2 (1 + R|x - y|)^s dy \le C_\epsilon \mathbb{R}^n \|FR\,\cdot\|^2_{C^{\frac{s}{2}+\epsilon}(\mathbb{R})}$$

for all $F \in C^{\frac{s}{2}+\epsilon}(\mathbb{R})$ with $supp\ F \subseteq [0, R]$, where C_ϵ is a constant independent of $F$and R.

Proof: For the proof, see Lemma 7.18, [80]. See also [97].

Next we show the following result.

Lemma (6.1.6)[158]: Assume that an operator L satisfies(H1)–(H2). Let $\psi_i \in S(\mathbb{R})$ be even functions, $\psi_i(0) = 0$, $i = 1, 2$. Then for every $\eta > 0$, there exists a positive constant $C = C(n, \eta, \psi_1, \psi_2)$ such that the kernel $K_{\psi_1(s\sqrt{L})\psi_2(t\sqrt{L})}(x, y) of\ \psi_1(s\sqrt{L})\psi_2(t\sqrt{L})$ satisfies

$$\left\|K_{\psi_1(s\sqrt{L})\psi_2(t\sqrt{L})}(x, y)\right\| \le C \left(\frac{min(s, t)}{max(s, t)}\right) \frac{max(s, t)^\eta}{(max(s, t) + |x - y|)^{n+\eta}} \tag{11}$$

for all $t > 0$ and $x, y \in \mathbb{R}^n$.

Proof: By symmetry, it suffices to show that if $s \le t$, then

$$\left|K_{\psi_1(s\sqrt{L})\psi_2(t\sqrt{L})}(x, y)\right| \le C\left(\frac{s}{t}\right) \frac{t^\eta}{(t + |x - y|)^{n+\eta}} \tag{12}$$

To do this, we fix $s, t > 0$ and let $\Psi(tx) = \frac{t}{s}\psi_1(sx)\psi_2(tx)$, and so $\psi_1(s\sqrt{L})\psi_2(t\sqrt{L}) = \frac{s}{t}\Psi(t\sqrt{L})$. Let us show that

$$\left|K_{\psi(t\sqrt{L})}(x,y)\right| \le Ct^{-n}, \qquad x,y \in \mathbb{R}^n \tag{13}$$

Indeed, for any $\kappa \in N$, we have the relationship

$$(I + t^2L)^{-\kappa} = \frac{1}{(\kappa-1)!}\int_0^\infty e^{-ut^2L}e^{-u}u^{k-1}du \tag{14}$$

and so when $\kappa > n/4$,

$$\|(I + t^2L)^{-\kappa}\|_{L^2(\mathbb{R}^n)\to L^\infty(\mathbb{R}^n)} \le \frac{1}{(\kappa-1)!}\int_0^\infty \left\|e^{-ut^2L}\right\|_{L^2(\mathbb{R}^n)\to L^\infty(\mathbb{R}^n)} e^{-u}u^{k-1}du \le Ct^{-n/2}$$

Now $\|(I + t^2L)^{-\kappa}\|_{L^1(\mathbb{R}^n)\to L^2(\mathbb{R}^n)} = \|(I + t^2L)^{-\kappa}\|_{L^2(\mathbb{R}^n)\to L^\infty(\mathbb{R}^n)} \le Ct^{-n/2}$, and so

$$\left\|\psi(t\sqrt{L})\right\|_{L^1(\mathbb{R}^n)\to L^\infty(\mathbb{R}^n)}$$
$$\le \left\|(I + t^2L)^{2\kappa}\psi(t\sqrt{L})\right\|_{L^2(\mathbb{R}^n)\to L^2(\mathbb{R}^n)}\|(I + t^2L)^{-\kappa}\|^2_{L^2(\mathbb{R}^n)\to L^\infty(\mathbb{R}^n)}$$

Since $\psi_1 \in S(R)$ and $\psi_1(0) = 0$, we have that $(s\lambda)^{-1}\psi_1(s\lambda) = \int_0^1 \psi_1'(s\lambda y)dy \in L^\infty(\mathbb{R})$, and then the L^2 operator norm of the last term is equal to the $L^\infty(\mathbb{R})$ norm of the function

$$(1 + t^2|\lambda|)^{2m}\Psi\left(t\sqrt{|\lambda|}\right) = \left[\frac{\psi_1(s\sqrt{|\lambda|})}{s\sqrt{|\lambda|}}\right]\left[(1 + t^2|\lambda|)^{2m}\left(t\sqrt{|\lambda|}\right)\psi_2(t\sqrt{|\lambda|})\right]$$

which is uniformly bounded in $t > 0$. This implies that (13) holds.

Next, we write $F(t\lambda) = \Psi(t\lambda)(1 + t^2\lambda^2)^m$, where $m > n/2$. Then we have $\Psi(t\sqrt{L}) = F(t\sqrt{L})(1 + t^2L)^{-m}$. From (14), it can be verified that for $m > n/2$, there exist some positive constants C and csuch that for every $t > 0$, the kernel $K_{(1+t^2L)^{-m}}(x,y)$ of the operator $(1 + t^2L)^{-m}$ satisfies

$$\left|K_{(1+t^2)^{-m}}(x,y)\right| \le \frac{C}{t^n}\exp\left(-\frac{|x-y|}{ct}\right)$$

which, in combination with $(1 + \frac{|x-y|}{t}) \le (1 + \frac{|x-z|}{t})(1 + \frac{|y-z|}{t})$, shows

$$\left|\left(1 + \frac{|x-y|}{t}\right)^{n+\eta} K_{\Psi(t\sqrt{L})}(x,y)\right|$$
$$= \left(1 + \frac{|x-y|}{t}\right)^{n+\eta}\left|\int_{\mathbb{R}^n} K_{F(t\sqrt{L})}(x,z)K_{(1+t^2L)^{-m}}(z,y)\,dz\right|$$
$$\le Ct^n\int_{\mathbb{R}^n}\left|K_{F(t\sqrt{L})}(x,z)\right|\left(1 + \frac{|x-y|}{t}\right)^{n+\eta}dz$$

By symmetry, estimate (12) will be proved if we show that

$$\int_{\mathbb{R}^n}\left|K_{F(t\sqrt{L})}(x,z)\right|\left(1 + \frac{|x-z|}{t}\right)^{n+\eta}dz \le C \tag{15}$$

Let $\varphi \in C_c^\infty(0,\infty)$ be a non-negative function satisfying supp $\varphi \subseteq [\frac{1}{4}, 1]$and let $\varphi_0 = 1 - \sum_{\ell=1}^\infty \varphi(2^{-\ell}\lambda)$. So,

$$\varphi_0(\lambda) + \sum_{\ell=1}^\infty \varphi(2^{-\ell}\lambda) = 1, \qquad \forall\lambda > 0$$

Let $F^0(t\lambda)$ denote the function $\varphi_0(t\lambda)F(t\lambda)$ and for $\ell \geq 1$, $F^\ell(t\lambda) := \varphi(2^{-\ell}t\lambda)F(t\lambda)$. From (13), the proof of (15) reduces to estimate the following:

$$\int_{\mathbb{R}^n} \left|K_{F(t\sqrt{L})}(x,z)\right| \left(1+\frac{|x-z|}{t}\right)^{n+\eta} dz$$

$$\leq C + \int_{\mathbb{R}^n} \left|K_{F^0(t\sqrt{L})}(x,z)\right|^2 \left(1+\frac{|x-z|}{t}\right)^{n+\eta} dz$$

$$+\sum_{\ell=1}^{\infty} \int_{|x-z|\geq t} \left|K_{F^\ell(t\sqrt{L})}(x,z)\right| \left(\frac{|x-z|}{t}\right)^{n+\eta} dz =: C + \sum_{\ell=0}^{\infty} I_\ell \qquad (16)$$

By Lemma(6.1.5)

$$I_0 \leq Ct^{n/2}\left(\int_{\mathbb{R}^n} \left|K_{F^0 t\sqrt{L}}(x,z)\right|^2 \left(1+\frac{|x-z|}{t}\right)^{3n+2\eta+1} dz\right)^{1/2} \leq C\left\|\delta_{1/t}F^0(t\,\cdot)\right\|_{C^{\frac{3}{2}n+2\eta+1}}$$

Since $\psi_1 \in S(\mathbb{R})$ and $\psi_1(0) = 0$, we have that $(s\lambda)^{-1}\psi_1(s\lambda) = \int_0^1 \psi_1'(s\lambda y)dy \in S(R)$. Then we have

$$I_0 \leq C\|\varphi_0(\lambda)\Psi(\lambda)(1+\lambda^2)^m\|_{C^{\frac{3}{2}n+2\eta+1}}$$

$$= \left\|\varphi_0(\lambda)\int_0^1 \varphi_1'(s\,\lambda_y/t)dy\,[\lambda\varphi_2(\lambda)(1+\lambda^2)^m]\right\|_{C^{\frac{3}{2}n+2\eta+1}} \leq C \qquad (17)$$

For the term I_ℓ, we use Lemma(6.1.5) again to obtain

$$I_\ell \leq Ct^{n/2}\left(\int_{\mathbb{R}^n} \left|K_{F^\ell t\sqrt{L}}(x,z)\right|^2 \left(\frac{|x-z|}{t}\right)^{3n+2\eta+1} dz\right)^{1/2}$$

$$\leq Ct^{n/2}2^{-\ell(3n+2\eta+1/2)}\left(\int_{\mathbb{R}^n} \left|K_{F^\ell t\sqrt{L}}(x,z)\right|^2 \left(1+\frac{2^\ell|x-z|}{t}\right)^{3n+2\eta+1} dz\right)^{1/2}$$

$$\leq C2^{-\ell(3n+2\eta+1/2)}2^{\ell n/2}\left\|\delta_{2^\ell/t}F^\ell(t\,\cdot)\right\|_{C^{\frac{3}{2}n+2\eta+1}}$$

It can be verified that for $\psi_i \in S(R)$, $i = 1, 2$,

$$\left\|\varphi\delta_{2^\ell}F\right\|_{C^{\frac{3}{2}n+2\eta+1}} = \left\|\varphi(\lambda)\int_0^1 \varphi_1'(2^\ell s\,\lambda_y/t)dy\,[2^\ell\lambda\varphi_2(2^\ell\lambda)(1+2^{2l}\lambda^2)^m]\right\|_{C^{\frac{3}{2}n+2\eta+1}}$$

$$\leq C2^{\ell\left(\frac{3}{2}n+2\eta+1\right)}2^{-2\ell}$$

which gives

$$\sum_{\ell=1}^{\infty} I_\ell \leq C\sum_{\ell=1}^{\infty} C2^{-\ell(3n+2\eta+1)/2}2^{\ell n/2}2^{\ell\left(\frac{3}{2}n+2\eta+1\right)}2^{-2\ell}$$

$$\leq C\sum_{\ell=1}^{\infty} 2^{-n\ell} \leq C \qquad (18)$$

Putting (17)and (18)into (16), estimate (15)follows readily. The proof of Lemma(6.1.6)is complete.

We devoted to the proof of Theorem(6.1.8), which give a (p,∞,M)-atomic representation for the Hardy spaces $H^p_{L,max}(\mathbb{R}^n)$. To do it, we begin with the following proposition.

Proposition (6.1.7)[158]: Let $0 < p \leq 1$. Let L be a non-negative self-adjoint operator on $L^2(\mathbb{R}^n)$ satisfying Gaussian estimate (GE). Let $\varphi_i \in S(\mathbb{R})$ be even functions with $\varphi_i(0) =$

1 and $\alpha_i > 0$, $i = 1, 2$. Then there exists a constant $C = C(n, \varphi_1, \varphi_2, \alpha_1, \alpha_2)$ such that for every $f \in L^2(\mathbb{R}^n)$, the functions $\varphi^*_{i,L,\alpha} f = sup|y - x| < \alpha_t |\varphi_i(t\sqrt{L})f(y)|$, $i = 1,2$, satisfy

$$\left\|\varphi^*_{1,L,\alpha_1} f\right\|_{L^p(\mathbb{R}^n)} \leq C \left\|\varphi^*_{2,L,\alpha_2} f\right\|_{L^p(\mathbb{R}^n)} \tag{19}$$

As a consequence, for any even function $\varphi \in S(\mathbb{R})$ with $\varphi(0) = 1$ and $\alpha > 0$,

$$C^{-1}\|f^*_L\|_{L^p(\mathbb{R}^n)} \leq \left\|\varphi^*_{L,\alpha} f\right\|_{L^p(\mathbb{R}^n)} \leq \|f^*_L\|_{L^p(\mathbb{R}^n)}$$

Proof: Recall that for any $0 < \alpha_2 \leq \alpha_1$,

$$\left\|\varphi^*_{1,L,\alpha_1} f\right\|_{L^p(\mathbb{R}^n)} \leq C \left(1 + \frac{\alpha_1}{\alpha_2}\right)^{n/p} \left\|\varphi^*_{L,\alpha_2} f\right\|_{L^p(\mathbb{R}^n)}$$

for any $\varphi \in S(\mathbb{R})$ ([164]). Now, we let $\psi(x) := \phi_1(x) - \phi_2(x)$, and then the proof of (19)reduces to show that

$$\left\|\varphi^*_{L,\alpha_1} f\right\|_{L^p(\mathbb{R}^n)} \leq C \left\|\varphi^*_{2,L,1} f\right\|_{L^p(\mathbb{R}^n)} \tag{20}$$

Let us show (20). Let $\Psi(x) = x^{2\kappa}\Phi(x)$ where $\Phi(x)$ is the function as in Lemma(6.1.4)and $2\kappa > (n+1)/p$. By the spectral theory ([53]), we have

$$f = C_{\Psi,\varphi_2} \int_0^\infty \Psi(s\sqrt{L})\varphi_2(s\sqrt{L}) f \frac{ds}{s}$$

Therefore,

$$\psi(t\sqrt{L})f(x) = C \int_0^\infty \left(\psi(t\sqrt{L})\Psi(s\sqrt{L})\right) \varphi_2(s\sqrt{L}) f(x) \frac{ds}{s}$$

Let us denote the kernel of $\psi(t\sqrt{L})\Psi(s\sqrt{L})$ by $K_{\psi(t\sqrt{L})\Psi(s\sqrt{L})}(x, y)$. Then for $\lambda \in (\frac{n}{p}, 2\kappa)$,

$$\sup_{|\omega|<t} \left|\psi(t\sqrt{L}) f(x - \omega)\right| = C \sup_{|\omega|<t} \left| \iint_{\mathbb{R}^{n+1}_+} K_{\psi(t\sqrt{L})\Psi(s\sqrt{L})}(x - \omega, z) \varphi_2(s\sqrt{L}) f(z) \frac{dzds}{s} \right|$$

$$\leq C \sup_{|\omega|<t} \int_{\mathbb{R}^{n+1}_+} \left|K_{\psi(t\sqrt{L})\Psi(s\sqrt{L})}(x - \omega, z)\right| \left(1 + \frac{|x - z|}{s}\right)^\lambda$$

$$\times \left|\varphi_2(s\sqrt{L}) f(z)\right| \left(1 + \frac{|x - z|}{s}\right)^{-\lambda} \frac{dzds}{s}$$

$$\leq \sup_{s,z} \left|\varphi_2(s\sqrt{L}) f(z)\right| \left(1 + \frac{|x - z|}{s}\right)^{-\lambda}$$

$$\times \sup_{|\omega|<t} \int_{\mathbb{R}^{n+1}_+} \left|K_{\psi(t\sqrt{L})\Psi(s\sqrt{L})}(x - \omega, z)\right| \left(1 + \frac{|x - z|}{s}\right)^\lambda \frac{dzds}{s} \tag{21}$$

Next we will prove that

$$\sup_{|\omega|<t} \int_{\mathbb{R}^{n+1}_+} \left|K_{\psi(t\sqrt{L})\Psi(s\sqrt{L})}(x - \omega, z)\right| \left(1 + \frac{|x - z|}{s}\right)^\lambda \frac{dzds}{s} \leq C \tag{22}$$

Once estimate (22)is shown, (20)follows. Indeed, it follows from (21), (22)and the condition $\lambda \in (\frac{n}{p}, 2\kappa)$ that

$$\|\varphi_{L,\alpha_1}^* f\|_{L^p(\mathbb{R}^n)} = \left\| \sup_{|\omega|<t} |\psi(t\sqrt{L})f(x-\omega)| \right\|_{L_x^p(\mathbb{R}^n)}$$

$$\leq C \left\| \sup_{z,s} |\varphi_2(s\sqrt{L})f(z)| \left(1+\frac{|x-z|}{s}\right)^{-\lambda} \right\|_{L_x^p(\mathbb{R}^n)}$$

$$\leq C \left\| \sup_{|y-x|<t} |\varphi_2(t\sqrt{L})f(y)| \right\|_{L_x^p(\mathbb{R}^n)} = C\|\varphi_{2,L,1}^* f\|_{L^p(\mathbb{R}^n)}$$

where we used Theorem 2.4 of [164]in the second inequality.
Let us prove (22). Note that $|w| < t$. We write

$$\psi(t\sqrt{L})\Psi(s\sqrt{L}) = \begin{cases} \left(\frac{s}{t}\right)^{2k} \left[\psi(t\sqrt{L})(t\sqrt{L})^{2k}\phi(s\sqrt{L})\right] & if\ s \leq t \\ \left(\frac{t}{s}\right)^{2} \left[(t\sqrt{L})^{-2}\psi(t\sqrt{L})(s\sqrt{L})^{2k+2}\phi(s\sqrt{L})\right], & if\ s > t \end{cases}$$

We then apply Lemma(6.1.6)to obtain that for $\eta \in (\lambda, 2\kappa)$,

$$\left|K_{\psi(t\sqrt{L})\Psi(s\sqrt{L})}(x-\omega,z)\right| \leq Cmin\left(\left(\frac{s}{t}\right)^{2k}, \left(\frac{t}{s}\right)^{2}\right) \frac{\max(s,t)^\eta}{(\max(s,t)+|x-\omega-z|)^{n+\eta}}$$

This, together with the fact that

$$\int_{|u|\geq s} \frac{\max(s,t)^\eta}{(\max(s,t)+|u-\omega|)^{n+\eta}} \left(1+\frac{|u|}{s}\right)^{\lambda} du \leq \int_{|u|\geq s} \frac{\max(s,t)^\eta}{(\max(s,t)+|u-\omega|)^{n+\eta}} du$$

Shows

$$\int_{\mathbb{R}^n} \left|K_{\psi(t\sqrt{L})\Psi(s\sqrt{L})}(x-\omega,z)\right| \left(1+\frac{|x-z|}{s}\right)^{\lambda} dz$$

$$\leq Cmin\left(\left(\frac{s}{t}\right)^{2k}, \left(\frac{t}{s}\right)^{2}\right) \left[1 + \int_{|u|\geq s} \frac{\max(s,t)^\eta}{(\max(s,t)+|u-\omega|)^{n+\eta}} \left(1+\frac{|u|}{s}\right)^{\lambda} du)\right] \tag{23}$$

To estimate the integrals over $|u| \geq s$, we note that if $s \geq t$, then we use the fact that $\eta > \lambda$ and $s + |u - w| \geq t + |u - w| \geq |w| + |u - w| \geq |u|$ to obtain

$$\int_{|u|\geq s} \frac{s^\eta}{(t+|u-\omega|)^{n+\eta}} \left(1+\frac{|u|}{s}\right)^{\lambda} du \leq 2^\lambda \int_{|u|\geq s} \frac{s^\eta}{(t+|u-\omega|)^{n+\eta}} \frac{|u|^\lambda}{s^\lambda} du$$

$$\leq \int_{|u|\geq s} \frac{s^\eta}{t+|u|^{n+\eta}} \frac{|u|^\lambda}{s^\lambda} du \leq C \tag{24}$$

If $s < t$, then it follows from the fact that $t + |u - w| \geq |w| + |u - w| \geq |u|$ and $\eta > \lambda$,

$$\int_{|u|\geq s} \frac{t^\eta}{(t+|u-\omega|)^{n+\eta}} \left(1+\frac{|u|}{s}\right)^{\lambda} du \leq 2^\lambda \int_{|u|\geq s} \frac{t^\eta}{(t+|u-\omega|)^{n+\eta}} \frac{|u|^\lambda}{s^\lambda} du$$

$$\leq \int_{|u|\geq s} \frac{t^\eta}{t+|u|^{n+\eta}} \frac{|u|^\lambda}{s^\lambda} du \leq C\left(\frac{t}{s}\right)^{\eta} \tag{25}$$

Putting estimates (24)and (25)into (23), we have obtained that for any $|w| < t$,

$$\int_{\mathbb{R}^n} \left|K_{\psi(t\sqrt{L})\Psi(s\sqrt{L})}(x-\omega,z)\right| \left(1+\frac{|x-z|}{s}\right)^{\lambda} dz$$
$$\leq Cmin\left(\left(\frac{s}{t}\right)^{2k},\left(\frac{s}{t}\right)^{2}\right)\left[1+\max(1.\left(\frac{t}{s}\right)^{\eta})\right] \leq Cmin\left(\left(\frac{s}{t}\right)^{2k-\eta},\left(\frac{t}{s}\right)^{2}\right)$$

Observe that $\eta < 2\kappa$. It follows

$$\sup_{|\omega|<t} \int_{\mathbb{R}_+^{n+1}} \left|K_{\psi(t\sqrt{L})\Psi(s\sqrt{L})}(x-\omega,z)\right| \left(1+\frac{|x-z|}{s}\right)^{\lambda} \frac{dzds}{s}$$
$$\leq C\int_0^{\infty} \min((\frac{s}{t})^{2k-n},(\frac{t}{s})^2)\frac{ds}{s} \leq C$$

which shows estimate (22), and the proof of Proposition (6.1.7) is end.

Theorem (6.1.8)[158]: Suppose that an operator L satisfies(H1)and(H2). Fix $0 < p \leq 1$. For all $q > p$ with $1 \leq q \leq \infty$ and for all integers $M > \frac{n}{2}(\frac{1}{p}-1)$, we have that $H^p_{L,\max}(\mathbb{R}^n) \subseteq H^p_{L,\text{at,q,M}}(\mathbb{R}^n)$, and hence by (5),

$$H^p_{L,\max}(\mathbb{R}^n) \simeq H^p_{L,\text{at,q,M}}(\mathbb{R}^n)$$

Proof. It suffices to show that for $f \in H^p_{L,max}(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$, fhas a (p,∞,M)atomic representation. We start with a suitable version of the Calderón repro-ducing formula. Let Φ be a function defined in Lemma(6.1.4), and set $\Psi(x) := x^{2M}\Phi(x), x \in \mathbb{R}$. By the spectral theory ([91]), for every $f \in L^2(\mathbb{R}^n)$ one can write

$$f = c_\Psi \int_0^{\infty} \Psi(t\sqrt{L})t^2Le^{-t^2L}\frac{dt}{t} = \lim_{\epsilon\to 0}\int_{\epsilon}^{1/\epsilon} \Psi(t\sqrt{L})t^2Le^{-t^2L}\frac{dt}{t} \tag{26}$$

with the integral converging in $L^2(\mathbb{R}^n)$.

Set

$$\eta(x) = c_\Psi \int_1^{\infty} t^2x^2\Psi(tx)e^{-t^2x^2}\frac{dt}{t} = c_\Psi\int_x^{\infty} y\Psi(y)e^{-y^2}dy, x \neq 0$$

with $\eta(0) = 1$. It follows that $\eta \in S(R)$is an even function, and

$$\eta(ax) - \eta(bx) = c_\Psi \int_a^b t^2x^2\Psi(tx)e^{-t^2x^2}f\frac{dt}{t}$$

By the spectral theory ([91]) again, one has

$$c_\Psi \int_a^b \Psi(t\sqrt{L})t^2Le^{-t^2L}f\frac{dt}{t} = \eta(a\sqrt{L})f(x) - \eta(b\sqrt{L})f(x) \tag{27}$$

Define,

$$M_Lf(x) := \sup_{|x-y|<5\sqrt{nt}} (\left|t^2Le^{-t^2L}f(y)\right| + \left|\eta(t\sqrt{L})f(y)\right|)$$

By Proposition(6.1.7), it follows that

$$\|M_Lf\|_{L^p(\mathbb{R}^n)} \leq C\|f\|_{H^p_{L,max}(\mathbb{R}^n)},\ \ 0 < p \leq 1$$

Recall that $\mathbb{R}_+^{n+1}$denotes the upper half-space in $\mathbb{R}^{n+1}$. If O is an open subset of $\mathbb{R}^n$, then the "tent" over O, denoted by $\hat{O}$, is given as $\hat{O}:= \{(x,t) \in \mathbb{R}^n + 1+ : B(x,4\sqrt{nt}) \subset O\}$. For $i \in \mathbb{Z}$, we define the family of sets $O_i := \{x \in \mathbb{R}^n : M_Lf(x) > 2^i\}$. Now let $\{Q_{ij}\}_j$ be a Whitney decomposition of O_i such that $O_i = \cup_j Q_{ij}$ and let $\hat{O}_i$ be a tent region. Set $\bar{e} = (1,\cdots,1) \in \mathbb{R}^n$. For every i,j, we define

$$\tilde{Q}_{ij} := \{(y,t) \in \mathbb{R}_+^{n+1} : y + 3t\bar{e} \in Q_{ij}\} \tag{28}$$

It can be verified that $\hat{O}_i \subset \cup_j \tilde{Q}_{ij}$. Indeed, for each $(y^0, t^0) \in \hat{O}_i$, we have that $B(y^0, 4\sqrt{nt^0}) \subset O_i$. Let $\tilde{y}^0 := y^0 + 3\bar{e}t^0$. Observe that $\tilde{y}^0 \in B(y^0, 4\sqrt{nt^0})$and then $\tilde{y}^0 \in O_i$. Then there exists some $Q_{ij_0} \subset O_i$ such that $\tilde{y}^0 \in Q_{ij_0}$, hence $(y^0, t^0) \in \tilde{Q}_{ij_0}$ and $\hat{O}_i \subset \bigcup_j \tilde{Q}_{ij}$. Note that $\tilde{Q}_{ij} \cap \tilde{Q}_{ij'} = \emptyset$ when $j \neq j'$. We obtain an decomposition for $\mathbb{R}^{n+1}_+$ as follows:

$$\mathbb{R}^{n+1}_+ = \cup_i \widehat{O_\iota} = \cup_i \widehat{O_\iota}/\widehat{O_{\iota+1}} = \cup_i \cup_j T_{ij}$$

where

$$T_{ij} := \tilde{Q}_{ij} \cap \widehat{O_\iota}/\widehat{O_{\iota+1}}$$

Using the formula (26), one can write

$$f = \sum_{ij} c_\Psi \Psi(t\sqrt{L})(\mathcal{X}T_{ij}(y,t)t^2 L e^{-t^2 L} f)\frac{dt}{t}$$

with the sum converging in $L^2(\mathbb{R}^n)$, where $\lambda_{ij} := 2^i |Q_{ij}|^{1/p}$, $a_{ij} := L^M b_{ij}$, and

$$b_{ij} := (\lambda_{ij})^{-1} c_\Psi \int_0^\infty t^{2M} \phi(t\sqrt{L})(\mathcal{X}T_{ij}(y,t)t^2 L e^{-t^2 L} f)\frac{dt}{t}$$

Let us show that the sum (29) converges in $L^2(\mathbb{R}^n)$. Indeed, since for each $f \in L^2(\mathbb{R}^n)$,

$$\left(\int_{\mathbb{R}^{n+1}_+} \left|t^2 L e^{-t^2\sqrt{L}} f(y)\right|^2 \frac{dydt}{t}\right)^{1/2} \leq C\|f\|_{L^2(\mathbb{R}^n)}$$

we use (29)to obtain

$$\left\|\sum_{|i|>N_1, |j|>N_2} \lambda_{ij} a_{ij}\right\|_{L^2(\mathbb{R}^n)}$$

$$= c_\Psi \left\|\sum_{|i|>N_1, |j|>N_2} \int_{\mathbb{R}^{n+1}} K_{(t^2 L)^M \Psi(t\sqrt{L})}(x,y)\mathcal{X}T_{ij}(y,t)t^2 L e^{-t^2 L} f(y)dy\frac{dt}{t}\right\|_{L^2(\mathbb{R}^n)}$$

$$\leq \sup_{\|g\|_2 \leq 1} \sum_{|i|>N_1, |j|>N_2} \int_{t_{ij}} \left|(t^2 L)^M \phi(t\sqrt{L})(x,y)g(y)t^2 L e^{-t^2\sqrt{L}} f(y)\right| \frac{dydt}{t}$$

$$\leq C\left(\sum_{|i|>N_1, |j|>N_2} \int_{t_{ij}} \left|t^2 L e^{-t^2\sqrt{L}} f(y)\right|^2 \frac{dydt}{t}\right)^{1/2} \longrightarrow 0$$

as $N_1 \to \infty$, $N_2 \to \infty$.

Next, we will show that, up to a normalization by a multiplicative constant, the a_{ij} are (p, ∞, M) −atoms. Once the claim is established, we shall have

$$\sum_{i,j} |\lambda_{ij}|^p = \sum_{i,j} 2^{ip}|Q_{ij}| \leq C\sum_i 2^{ip}|O_i| \leq C\|f\|^p_{H^p_{L,max}(\mathbb{R}^n)}$$

as desired.

We prove that for every i, j, the function $C^{-1}a_{ij}$ is a (p, ∞, M)-atom as-sociated with the cube $30Q_{ij}$ for some constant C. Observe that if $(y, t) \in T_{ij}$, then $B(y, 4\sqrt{nt}) \in O_i$. Denote by $\tilde{y} := y + 3t\bar{e}$, and so $\tilde{y} \in Q_{ij}$ and $B(\tilde{y}, \sqrt{nt}) \in O_i$. The fact that Q_{ij} is the Whitney cube of Oiimplies that $5Q_{ij} \cap O_i^c \neq \emptyset$. Denote the side length of Q_{ij} by $\ell(Q_{ij})$. It then follows that $t \leq 3\ell(Q_{ij})$. Since $y + 3\bar{e}t \in Q_{ij}$, we have that $y \in 20Q_{ij}$. From Lemma(6.1.4), the integral kernel $K_{(t^2 L)^k \Phi(t\sqrt{L})}$ of the operator $(t^2 L)^k \Phi(t\sqrt{L})$ satisfies

$$suppK_{(t^2L)^K\Psi(t\sqrt{L})} \subseteq \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n : |x-y| \le t\}$$

This concludes that for every $k = 0,1,\cdots,M$

$$supp(L^K b_{ij}) \subseteq 30Q_{ij}$$

It remains to show that $\|(\ell(Q_{ij})^2L)^k b_{ij}\|_{L^\infty(\mathbb{R}^n)} \le C(\ell(Q_{ij}))^{2M}|Q_{ij}|^{-1/p}$, $k = 0,1,\cdots,M$. When $K = 0,1,\cdots,M-1$, it reduces to show

$$\left|\int_0^\infty \int_{\mathbb{R}^n} K_{t^{2M}L^K\Psi(t\sqrt{L})}(x,y)\mathcal{X}T_{ij}(y,t)t^2Le^{-t^2L}f(y)dy\frac{dt}{t}\right| \le C2^i\ell(Q_{ij})^{2(M-K)}. \quad (30)$$

Indeed, if $\chi T_{ij}(y,t) = 1$, then $(y,t) \in (\widehat{O_{i+1}})^c$, and so $B(y,4\sqrt{n}t) \cap (O_{i+1})^c \ne \emptyset$. Let $\bar{x} \in B(y,4\sqrt{n}t) \cap (O_{i+1})^c$. We have that $|t^2Le^{-t^2L}f(y)| \le M_Lf(\bar{x}) \le 2^{i+1}$. By Lemma(6.1.4),

$$\left|\int_0^\infty \int_{\mathbb{R}^n} K_{t^{2M}L^K\Psi(t\sqrt{L})}(x,y)\mathcal{X}T_{ij}(y,t)t^2Le^{-t^2L}f(y)dy\frac{dt}{t}\right|$$

$$\le C2^i\left|\int_0^{\ell(Q_{ij})} t^{2(M-k)}\int_{\mathbb{R}^n}\left|K_{(t^2L)^K\Psi(t\sqrt{L})}(x,y)\right|dy\frac{dt}{t}\right|$$

$$\le C2^i\int_0^{\ell(Q_{ij})} t^{2(M-k)}\frac{dt}{t} \le C2^i\ell(Q_{ij})^{2(M-K)}$$

since $K = 0,1,\cdots,M-1$.

Now we consider the case $k = M$. The proof is based on a modification of a technique due to A. Calderón [163]. In this case, we need to prove that for every i,j,

$$\left|\int_0^\infty \int_{\mathbb{R}^n} K_{\Psi(t\sqrt{L})}(x,y)\mathcal{X}T_{ij}(y,t)t^2Le^{-t^2L}f(y)dy\frac{dt}{t}\right| < C2^i \quad (31)$$

To show (31), we fix x and let $d(x,Q_{ij}) < 30\sqrt{n}\ell(Q_{ij})$. We claim the following result:

(P1)The properties of the set defining $\chi T_{ij}(y,t)$ imply that there exist intervals $(0,b_0),(a_1,b_1),\cdots,(a_N,\infty), 0 < b_0 \le a_1 < b_1 \le\cdots\le a_N, 1 \le N \le 2n+2$ such that for $l = 0,1,\cdots,N-1$, there hold $a_{l+1} \le 3^{2n+2}b_l$ and

(a)$K_{\Psi(t\sqrt{L})}(x,y)\chi T_{ij}(y,t) = 0$ for $t > a_N$;

(b)either $K_{\Psi(t\sqrt{L})}(x,y)\chi T_{ij}(y,t) = 0$ or $K_{\Psi(t\sqrt{L})}(x,y)\chi T_{ij}(y,t) = K_{\Psi(t\sqrt{L})}(x,y)$ for all $t \in (a_l,b_l)$;

(c)either $K_{\Psi(t\sqrt{L})}(x,y)\chi T_{ij}(y,t) = 0$ or $K_{\Psi(t\sqrt{L})}(x,y)\chi T_{ij}(y,t) = K_{\Psi(t\sqrt{L})}(x,y)$ for all $t \in (0,b_0)$.

Assuming this claim (P1)for the moment, we observe that for $d(x,Q_{ij}) < 30\sqrt{n}\ell(Q_{ij})$, one can write

$$\int_0^\infty \int_{\mathbb{R}^n} K_{\Psi(t\sqrt{L})}(x,y)\mathcal{X}T_{ij}(y,t)t^2Le^{-t^2L}f(y)dy\frac{dt}{t}$$

$$= \left\{\int_0^{b_0} + \sum_{l=1}^{N-1}\int_{a_l}^{b_l}\right\}\int_{\mathbb{R}^n} K_{\Psi(t\sqrt{L})}(x,y)\mathcal{X}T_{ij}(y,t)t^2Le^{-t^2L}f(y)dy\frac{dt}{t}$$

$$+ \sum_{l=0}^{N-1}\int_{b_l}^{a_{l+1}}\int_{\mathbb{R}^n} K_{\Psi(t\sqrt{L})}(x,y)\mathcal{X}T_{ij}(y,t)t^2Le^{-t^2L}f(y)dy\frac{dt}{t}$$

$$= I_1(x) + I_2(x) \quad (32)$$

To estimate $I_1(x)$, we note that if $a_l \le a < b \le b_l$ or $0 \le a < b \le b_0$, then one has either

$$\int_a^b \int_{\mathbb{R}^n} K_{\Psi(t\sqrt{L})}(x,y)\mathcal{X}T_{ij}(y,t)t^2Le^{-t^2L}f(y)dy\frac{dt}{t}$$

or by (27),

$$\int_a^b \int_{\mathbb{R}^n} K_{\Psi(t\sqrt{L})}(x,y)\mathcal{X}T_{ij}(y,t)t^2Le^{-t^2L}f(y)dy\frac{dt}{t} = \int_a^b \Psi(t\sqrt{L})t^2Le^{-t^2L}f(x)\frac{dt}{t}$$
$$= \eta(a\sqrt{L})f(x) - \eta(b\sqrt{L})f(x)$$

Observe that for each $a \le t \le b$, if $|x - y| < t$, then $\chi T_{ij}(y,t) = 1$. This tells us that $(y,t) \in (\widehat{O_{i+1}})^c$, hence $B(y, 4\sqrt{n}t) \cap (O_{i+1})^c \ne \emptyset$. Assume that $\bar{x} \in B(y, 4\sqrt{n}t) \cap (O_{i+1})^c$. From this, we have that $|x - \bar{x}| \le |x - y| + |y - \bar{x}| < 5\sqrt{n}t$ and $M_L f(\bar{x}) \le 2^{i+1}$. It implies that $|\eta(t\sqrt{L})f(x)| \le M_L f(\bar{x}) \le C2^{i+1}$for every $a \le t \le b$. Therefore, $|\eta(a\sqrt{L})f(x)| \le C2^{i+1}$ and $|\eta(b\sqrt{L})f(x)| \le C2^{i+1}$, and so $|I_1(x)| \le C2^{i+1}$.

Consider $I_2(x)$. If $\chi T_{ij}(y,t) = 1$, then $(y,t) \in (\hat{O}_{i+1})^c$. Thus $B(y, 4\sqrt{n}t) \cap (O_{i+1})^c = \emptyset$. Assume that $\bar{x} \in B(y, 4\sqrt{n}t) \cap (O_{i+1})^c$. We have that $|t^2Le^{-t^2L}f(y)| \le M_L f(\bar{x}) \le 2^{i+1}$. This, together with $a_{l+1} \le cb_l$, implies that

$$\left|\int_{b_l}^{a_{l+1}} \int_{\mathbb{R}^n} K_{\Psi(t\sqrt{L})}(x,y)\mathcal{X}T_{ij}(y,t)t^2Le^{-t^2L}f(y)dy\frac{dt}{t}\right|$$
$$\le 2^{i+1}\left|\int_{b_l}^{cb_l} \int_{\mathbb{R}^n} \left|K_{\Psi(t\sqrt{L})}(x,y)\right| dy\frac{dt}{t}\right| \le C2^{i+1}\int_{b_l}^{cb_l} \frac{1}{t}dt$$
$$\le C2^{i+1} \qquad (33)$$

which yields that $|I_2(x)| \le C2^{i+1}$.

Combining (32)and (33), we obtain (31). It follows that $\|a_{ij}\|_{L^\infty(\mathbb{R}^n)} \le C|Q_{ij}|^{-1/p}$. Up to a normalization by a multiplicative constant, the a_{ij} are (p, ∞, M)-atoms.

It remains to prove the claim(P1). Note that $\mathcal{X}T_{ij}(y,t) = \chi_{\widehat{O_l}}(y,t) \cdot \chi_{(\widehat{O_{l+1}})^c}(y,t) \cdot \chi_{\widetilde{Q_{ij}}}(y,t)$; Assume that $Q_{ij} = \{(y_1, \cdots, y_n) : c_l \le y_l \le d_l, l = 1, \cdots, n\}$. Then

$$\chi_{\check{Q}_{ij}}(y,t) = \prod_{l=1}^{n} \chi_{\{c_l \le y_l + 3t \le d_l\}}(y,t) = \prod_{l=1}^{n} \chi_{\{y_l+3t \ge c_l\}}(y,t)\chi_{\{y_l+3t \le d_l\}}(y,t)$$

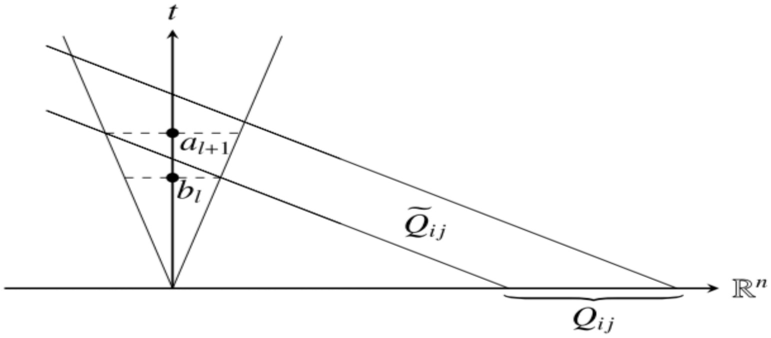


Fig. (1)[158]: The case of $x_l < c_l$.

Let $\chi_l(y,t)$ be one of the characteristic functions $\chi_{\{y_l+3t \ge c_l\}}(y,t)$, $\chi_{\{y_l+3t \le d_l\}}(y,t)$, $\chi_{\widehat{O_l}}(y,t)$ and $\chi_{(\widehat{O_{l+1}})^c}(y,t)$. We will show the following property:

(P2)There exist numbers b_l and $a_{l+1}$with $0 < b_l \le a_{l+1}$ and $a_{l+1} \le 3b_l$ such that for every x, either $K_{\Psi(t\sqrt{L})}(x,y)\chi_l(y,t) = 0$ or $K_{\Psi(t\sqrt{L})}(x,y)\chi_l(y,t) = K_{\Psi(t\sqrt{L})}(x,y)$ for all tin each of the intervals complementary to (b_l, a_{l+1}). And for at least one of $\chi_l(y,t), K_{\Psi(t\sqrt{L})}(x,y)\chi_l(y,t) = 0$ for $t > a_{l+1}$.

Then the same holds for $\chi T_{ij}(y,t) = \prod_{l=1}^{2n+2} \chi_l(y,t) K_{\Psi(t\sqrt{L})}(x,y)$ in each of the intervals complementary to the union of the intervals (b_l, a_{l+1}), which is what was asserted in the claim. Thus we merely have to prove(P2). To do this, we consider four cases.

Case 1. $\chi_l(y,t) = \chi\{y_l + 3t \ge c_l\}(y,t)$.

In this case, since $suppK_{\Psi(t\sqrt{L})}(x,y) \subseteq \{y: |x-y| \le t\}$, we have that $\mathrm{supp}K_{\Psi(t\sqrt{L})}(x,y) \subseteq \{y: x_l - t \le y_l \le x_l + t\}$. If $x_l \ge c_l$, then $y_l + 3t \ge x_l + 2t \ge c_l$ for any $t > 0$. This yields

$$K_{\Psi(t\sqrt{L})}(x,y)\chi_l(y,t) = K_{\Psi(t\sqrt{L})}(x,y), \qquad t > 0.$$

If $x_l < c_l$, then we choose $b_l = \frac{c_l - x_l}{4}$and $a_{l+1} = \frac{c_l - x_l}{2}$ (see Fig.1).

In the case of t <bl, we have $y_l + 3t \le x_l + 4t < c_l$, which implies that $K_{\Psi(t\sqrt{L})}(x,y)\chi T_{ij}(y,t) = 0$. In the case of $t > a_{l+1}$, we have $y_l + 3t \ge x_l + 2t > c_l$. This implies that $K_{\Psi(t\sqrt{L})}(x,y)\chi_l(y,t) = K_{\Psi(t\sqrt{L})}(x,y)$.

Case 2. $\chi_l(y,t) = \chi\{y_l + 3t \le d_l\}(y,t)$.

Since $suppK_{\Psi(t\sqrt{L})}(x,y) \subseteq \{y: |x-y| \le t\}$, we have that $K_{\Psi(t\sqrt{L})}(x,y) \subseteq \{y: x_l - t \le y_l \le x_l + t\}$. When $x_l \ge d_l$, we have that $y_l + 3t \ge x_l + 2t > dl$ for any $t > 0$. This tells us

$$K_{\Psi(t\sqrt{L})}(x,y)\chi_l(y,t) = 0, \quad for\ t > 0$$

When $x_l < d_l$, we choose $b_l = \frac{d_l - x_l}{4}$ and $a_{l+1} = \frac{d_l - x_l}{2}$. If $t < b_l$, then $y_l + 3t \le x_l + 4t < d_l$, which implies that $K_{\Psi(t\sqrt{L})}(x,y)\chi_l(y,t) = K_{\Psi(t\sqrt{L})}(x,y)$. If $t > a_{l+1}$, then $y_l + 3t \ge x_l + 2t > d_l$. From this, we have that $K_{\Psi(t\sqrt{L})}(x,y)\chi_l(y,t) = 0$.

Case 3. $\chi_l(y,t) = \chi + O_i(y,t)$.

In this case, we choose $b_l = \frac{1}{5\sqrt{n}} d(x, O_i^c)$ and $a_{l+1} = \frac{1}{2\sqrt{n}} d(x, O_i^c)$. Let $|x-y| < t$. If $t < \frac{1}{5\sqrt{n}} d(x, O_i^c)$, then $d(y, O_i^c) \ge d(x, O_i^c) - |x - y| > 5\sqrt{n}t - t \ge 4\sqrt{n}t$. This tells us

$$K_{\Psi(t\sqrt{L})}(x,y)\chi_l(y,t) = K_{\Psi(t\sqrt{L})}(x,y)$$

for $t < \frac{1}{5\sqrt{n}} d(x, O_i^c)$. If $t > \frac{1}{2\sqrt{n}} d(x, O_i^c)$, then $d(y, O_i^c) \le d(x, O_i^c) + d(x,y) < (2\sqrt{n} + 1)t < 4\sqrt{n}t$. Hence, if $t > \frac{1}{2\sqrt{n}} d(x, O_i^c)$, then

$$K_{\Psi(t\sqrt{L})}(x,y)\chi_l(y,t) = 0$$

Case 4. $\chi_l(y,t) = \chi(\widehat{O_{l+1}})^c(y,t)$.

In this case, we can choose $b_l = \frac{1}{5\sqrt{n}} d(x, O_{i+1}^c)$ and $a_{l+1} = \frac{1}{2\sqrt{n}} d(x, O_{i+1}^c)$. The proof can be an adaptation of the proof as in Case 3, and we omit the detail here.

This concludes the proof of the property(P2). We have obtained the claim(P1), and then the proof of Theorem(6.1.8)is complete.

As a consequence of Theorem(6.1.8), we have the following equivalent characterization for functions in $H^p_{L,at,q,M}(\mathbb{R}^n)$.

Corollary (6.1.9)[158]: Suppose that an operator L satisfies (H1)–(H2). Fix $0 < p \le 1$. For all $q > p$ with $1 \le q \le \infty$ and for all integers $M > \frac{n}{2}(\frac{1}{p} - 1)$, if $f \in L^2(\mathbb{R}^n)$, then the following conditions on fare equivalent:

(i) $f \in H^p_{L,at,q,M}(\mathbb{R}^n)$;

(ii) Given $\alpha > 0$, $\varphi^*_{L,\alpha} f = \sup_{|y-x|<\alpha t} |\phi(t\sqrt{L})f(y)| \in L^p(\mathbb{R}^n)$ for some even function $\varphi \in S(R)$, $\varphi(0) = 1$;

(iii) $G^*_L(f) = \sup_{\phi \in A} \sup_{|y-x|<t} |\varphi(t\sqrt{L})f(y)| \in L^p(\mathbb{R}^n)$,

$$A = \{\varphi \in L\,(\mathbb{R}): even\ function\ with\ \varphi(0) \neq 0, \int_{\mathbb{R}} (1 + |x|)^N \sum_{k \le N} \left|\frac{d^k}{dx^k}\varphi(x)\right|^2 dx \le 0\}$$

where N is a large number depending only on p and n.

Proof: The proof follows the line of (ii)⇒(i) ⇒(iii)⇒(ii). From Theorem(6.1.8)and Proposition(6.1.7), it tells us an implication (ii) ⇒(i). The proof of (i) ⇒(iii) will be an adaptation of the proof of the earlier known implication of (i) ⇒(ii) (see [171]). Obviously, (iii)⇒(ii). This proves Corollary(6.1.9).

We consider an electromagnetic Laplacian

$$L = (i\nabla - A(x))^2 + V(x), \quad n \ge 3$$

Recall that a measurable function Von $\mathbb{R}^n$ is in the Kato class when

$$\lim_{r \downarrow 0} \sup_x \int_{|x-y|<r} \frac{|V(y)|}{|x-y|^{n-2}} dy = 0$$

while the Kato norm is defined by

$$\|V\|_K = \sup_x \int \frac{|V(y)|}{|x-y|^{n-2}} dy = 0$$

Proposition (6.1.10)[158]: Consider an electromagnetic Laplacian

$$L = (i\nabla - A(x))^2 + V(x), \quad n \ge 3$$

Assume that $A \in L^2_{loc}(\mathbb{R}^n, \mathbb{R}^n)$, and the positive and negative parts $V_\pm$ of V satisfy V_+ is of Kato class, $\|V_-\|_K < c_n = \pi^{n/2}/\Gamma(n/2\ - 1)$. Then for every integer $M > \frac{n}{2}(\frac{1}{p} - 1)$, the spaces $H^p_{L,max}(\mathbb{R}^n)$ and $H^p_{L,\mathrm{at},\infty,\mathrm{M}}(\mathbb{R}^n)$ coincide. In particular,

$$\|f\|_{H^p_{L,max}(\mathbb{R}^n)} \approx \|f\|_{H^p_{L,\mathrm{at},\infty,\mathrm{M}}(\mathbb{R}^n)}$$

Proof: It is known (see [162]) that under assumptions of Proposition(6.1.10), the operator L has a unique nonnegative self-adjoint extension, and e^{-tL} is an integral operator whose kernel satisfies the Gaussian estimate (H2). Now Proposition(6.1.10)is a straightforward consequence of Theorem(6.1.8).

Corollary (6.1.11)[183]: Assume that an operator L satisfies(H1)–(H2). Let $\psi_i \in S(\mathbb{R})$ be even functions, $\psi_i(0) = 0$, $i = 1, 2$. Then for every $\eta > 0$, there exists a positive constant $C = C(n, \eta, \psi_1, \psi_2)$ such that the kernel $K_{\psi_1((1+2\epsilon)\sqrt{L})\psi_2\left((1+\epsilon)\sqrt{L}\right)}(x, y)$ of $\psi_1((1 + 2\epsilon)\sqrt{L})\psi_2((1 + \epsilon)\sqrt{L})$ satisfies

$$\left\|K_{\psi_1((1+2\epsilon)\sqrt{L})\psi_2((1+\epsilon)\sqrt{L})}(x, y)\right\| \le C \left(\frac{\min(1 + 2\epsilon, 1 + \epsilon)}{\max(1 + 2\epsilon, 1 + \epsilon)}\right) \frac{max(1 + 2\epsilon, 1 + \epsilon)^\eta}{(\max(1 + 2\epsilon, 1 + \epsilon) + |x - y|)^{n+\eta}} \qquad (34)$$

for all $\epsilon \ge 0$ and $x, y \in \mathbb{R}^n$.

Proof: By symmetry, it suffices to show that if $\epsilon \geq 0$, then

$$\left|K_{\psi_1((1+\epsilon)\sqrt{L})\psi_2((1+2\epsilon)\sqrt{L})}(x,y)\right| \leq C\left(\frac{1+\epsilon}{1+2\epsilon}\right)\frac{(1+2\epsilon)^{\eta}}{(1+2\epsilon+|x-y|)^{n+\eta}} \tag{35}$$

To do this, we fix $\epsilon > 0$ and let $\Psi((1+\epsilon)x) = \frac{1+\epsilon}{1+2\epsilon}\psi_1((1+2\epsilon)x)\psi_2((1+\epsilon)x)$, and so $\psi_1((1+2\epsilon)\sqrt{L})\psi_2((1+\epsilon)\sqrt{L}) = \frac{1+2\epsilon}{1+\epsilon}\Psi((1+\epsilon)\sqrt{L})$. Let us show that

$$\left|K_{\psi((1+\epsilon)\sqrt{L})}(x,y)\right| \leq C(1+\epsilon)^{-n}, \qquad x,y \in \mathbb{R}^n \tag{36}$$

Indeed, for any $\kappa \in N$, we have the relationship

$$(I+(1+\epsilon)^2L)^{-\kappa} = \frac{1}{(\kappa-1)!}\int_0^{\infty} e^{-u(1+\epsilon)^2L}e^{-u}u^{k-1}du \tag{37}$$

and so when $\kappa > n/4$,

$$\|(I+(1+\epsilon)^2L)^{-\kappa}\|_{L^2(\mathbb{R}^n)\to L^{\infty}(\mathbb{R}^n)} \leq \frac{1}{(\kappa-1)!}\int_0^{\infty}\left\|e^{-u(1+\epsilon)^2L}\right\|_{L^2(\mathbb{R}^n)\to L^{\infty}(\mathbb{R}^n)}e^{-u}u^{k-1}du$$
$$\leq C(1+\epsilon)^{-n/2}$$

Now $\|(I+(1+\epsilon)^2L)^{-\kappa}\|_{L^1(\mathbb{R}^n)\to L^2(\mathbb{R}^n)} = \|(I+(1+\epsilon)^2L)^{-\kappa}\|_{L^2(\mathbb{R}^n)\to L^{\infty}(\mathbb{R}^n)} \leq C(1+\epsilon)^{-n/2}$, and so

$$\left\|\psi((1+\epsilon)\sqrt{L})\right\|_{L^1(\mathbb{R}^n)\to L^{\infty}(\mathbb{R}^n)} \leq \left\|(I+(1+\epsilon)^2L)^{2\kappa}\psi((1+\epsilon)\sqrt{L})\right\|_{L^2(\mathbb{R}^n)\to L^2(\mathbb{R}^n)}\|(I+(1+\epsilon)^2L)^{-\kappa}\|^2_{L^2(\mathbb{R}^n)\to L^{\infty}(\mathbb{R}^n)}$$

Since $\psi_1 \in S(R)$ and $\psi_1(0) = 0$, we have that $((1+2\epsilon)\lambda)^{-1}\psi_1((1+2\epsilon)\lambda) = \int_0^1 \psi_1'((1+2\epsilon)\lambda y)dy \in L^{\infty}(\mathbb{R})$, and then the L^2 operator norm of the last term is equal to the $L^{\infty}(\mathbb{R})$ norm of the function

$$(1+(1+\epsilon)^2|\lambda|)^{2m}\Psi\left((1+\epsilon)\sqrt{|\lambda|}\right) = \left[\frac{\psi_1((1+2\epsilon)\sqrt{|\lambda|})}{(1+2\epsilon)\sqrt{|\lambda|}}\right]\left[(1+(1+\epsilon)^2|\lambda|)^{2m}\left((1+\epsilon)\sqrt{|\lambda|}\right)\psi_2\left((1+\epsilon)\sqrt{|\lambda|}\right)\right]$$

which is uniformly bounded in $\epsilon \geq 0$. This implies that (36) holds.

Next, we write $F((1+\epsilon)\lambda) = \Psi((1+\epsilon)\lambda)(1+(1+\epsilon)^2\lambda^2)^m$, where $m > n/2$. Then we have $\Psi((1+\epsilon)\sqrt{L}) = F((1+\epsilon)\sqrt{L})(1+(1+\epsilon)^2L)^{-m}$. From (37), it can be verified that for $m > n/2$, there exist some positive constants C and csuch that for every $\epsilon \geq 0$, the kernel $K_{(1+(1+\epsilon)^2L)^{-m}}(x,y)$ of the operator $(1+(1+\epsilon)^2L)^{-m}$ satisfies

$$\left|K_{(1+(1+\epsilon)^2)^{-m}}(x,y)\right| \leq \frac{C}{(1+\epsilon)^n}\exp\left(-\frac{|x-y|}{c(1+\epsilon)}\right)$$

which, in combination with $(1+\frac{|x-y|}{1+\epsilon}) \leq (1+\frac{|x-z|}{1+\epsilon})(1+\frac{|y-z|}{1+\epsilon})$, shows

$$\left|\left(1+\frac{|x-y|}{1+\epsilon}\right)^{n+\eta}K_{\Psi((1+\epsilon)\sqrt{L})}(x,y)\right|$$
$$= \left(1+\frac{|x-y|}{1+\epsilon}\right)^{n+\eta}\left|\int_{\mathbb{R}^n}K_{F((1+\epsilon)\sqrt{L})}(x,z)K_{(1+(1+\epsilon)^2L)^{-m}}(z,y)\,dz\right|$$

$$\leq C(1+\epsilon)^n \int_{\mathbb{R}^n} \left|K_{F((1+\epsilon)\sqrt{L})}(x,z)\right| \left(1+\frac{|x-y|}{1+\epsilon}\right)^{n+\eta} dz$$

By symmetry, estimate (35) will be proved if we show that

$$\int_{\mathbb{R}^n} \left|K_{F((1+\epsilon)\sqrt{L})}(x,z)\right| \left(1+\frac{|x-z|}{1+\epsilon}\right)^{n+\eta} dz \leq C \tag{38}$$

Let $\varphi \in C_c^\infty(0,\infty)$ be a non-negative function satisfying supp $\varphi \subseteq [\frac{1}{4}, 1]$and let $\varphi_0 = 1 - \sum_{\ell=1}^{\infty} \varphi(2^{-\ell}\lambda)$. So,

$$\varphi_0(\lambda) + \sum_{\ell=1}^{\infty} \varphi(2^{-\ell}\lambda) = 1, \qquad \forall \lambda > 0$$

Let $F^0((1+\epsilon)\lambda)$ denote the function $\varphi_0((1+\epsilon)\lambda)F((1+\epsilon)\lambda)$ and for $\ell \geq 1$, $F^\ell((1+\epsilon)\lambda) := \varphi(2^{-\ell}(1+\epsilon)\lambda)F((1+\epsilon)\lambda)$. From (36), the proof of (38) reduces to estimate the following:

$$\int_{\mathbb{R}^n} \left|K_{F((1+\epsilon)\sqrt{L})}(x,z)\right| \left(1+\frac{|x-z|}{1+\epsilon}\right)^{n+\eta} dz$$
$$\leq C + \int_{\mathbb{R}^n} \left|K_{F^0((1+\epsilon)\sqrt{L})}(x,z)\right|^2 \left(1+\frac{|x-z|}{(1+\epsilon)}\right)^{n+\eta} dz$$
$$+ \sum_{\ell=1}^{\infty} \int_{|x-z|\geq 1+\epsilon} \left|K_{F^\ell((1+\epsilon)\sqrt{L})}(x,z)\right| \left(\frac{|x-z|}{1+\epsilon}\right)^{n+\eta} dz =: C + \sum_{\ell=0}^{\infty} I_\ell \tag{39}$$

By Lemma (6.1.5)

$$I_0 \leq C(1+\epsilon)^{n/2} \left(\int_{\mathbb{R}^n} \left|K_{F^0(1+\epsilon)\sqrt{L}}(x,z)\right|^2 \left(1+\frac{|x-z|}{1+\epsilon}\right)^{3n+2\eta+1} dz \right)^{1/2}$$
$$\leq C \left\| \delta_{1/(1+\epsilon)} F^0((1+\epsilon)\,\cdot) \right\|_{C^{\frac{3}{2}n+2\eta+1}}$$

Since $\psi_1 \in S(\mathbb{R})$ and $\psi_1(0) = 0$, we have that $((1+2\epsilon)\lambda)^{-1}\psi_1((1+2\epsilon)\lambda) = \int_0^1 \psi_1'((1+2\epsilon)\lambda y)dy \in S(R)$. Then we have

$$I_0 \leq C\|\varphi_0(\lambda)\Psi(\lambda)(1+\lambda^2)^m\|_{C^{\frac{3}{2}n+2\eta+1}}$$
$$= \left\| \varphi_0(\lambda) \int_0^1 \varphi_1'((1+2\epsilon)\,\lambda_y/(1+\epsilon))dy\,[\lambda\varphi_2(\lambda)(1+\lambda^2)^m] \right\|_{C^{\frac{3}{2}n+2\eta+1}}$$
$$\leq C \tag{40}$$

For the term I_ℓ, we use Lemma (6.1.5) again to obtain

$$I_\ell \leq C(1+\epsilon)^{n/2} \left(\int_{\mathbb{R}^n} \left|K_{F^\ell(1+\epsilon)\sqrt{L}}(x,z)\right|^2 \left(\frac{|x-z|}{1+\epsilon}\right)^{3n+2\eta+1} dz \right)^{1/2}$$
$$\leq C(1+\epsilon)^{n/2} 2^{-\ell(3n+2\eta+1/2)} \left(\int_{\mathbb{R}^n} \left|K_{F^\ell(1+\epsilon)\sqrt{L}}(x,z)\right|^2 \left(1+\frac{2^\ell|x-z|}{1+\epsilon}\right)^{3n+2\eta+1} dz \right)^{1/2}$$
$$\leq C 2^{-\ell(3n+2\eta+1/2)} 2^{\ell n/2} \left\| \delta_{2^\ell/(1+\epsilon)} F^\ell((1+\epsilon)\,\cdot) \right\|_{C^{\frac{3}{2}n+2\eta+1}}$$

It can be verified that for $\psi_i \in S(R)$, $i = 1, 2$,

$$\left\| \varphi \delta_{2^\ell} F \right\|_{C^{\frac{3}{2}n+2\eta+1}}$$

$$= \left\| \varphi(\lambda) \int_0^1 \varphi_1'\big(2^\ell(1+2\epsilon)\,\lambda_y/(1+\epsilon)\big)dy\,\big[2^\ell\lambda\varphi_2\big(2^\ell\lambda\big)(1+2^{2l}\lambda^2)^m\big]\right\|_{C^{\frac{3}{2}n+2\eta+1}}$$
$$\leq C2^{\ell\left(\frac{3}{2}n+2\eta+1\right)}2^{-2\ell}$$

which gives

$$\sum_{\ell=1}^{\infty} I_\ell \leq C \sum_{\ell=1}^{\infty} C2^{-\ell(3n+2\eta+1)/2}2^{\ell n/2}2^{\ell\left(\frac{3}{2}n+2\eta+1\right)}2^{-2\ell} \leq C\sum_{\ell=1}^{\infty} 2^{-n\ell} \leq C \qquad (41)$$

Putting (40)and (41)into (39), estimate (38)follows readily. The proof of Corollary (6.1.11) is complete.

Corollary (6.1.12)[183]: Let $0 \leq \epsilon < 1$. Let L be a non-negative self-adjoint operator on $L^2(\mathbb{R}^n)$ satisfying Gaussian estimate (GE). Let $\varphi_i \in S(\mathbb{R})$ be even functions with $\varphi_i(0) = 1$ and $\alpha_i > 0$, $i = 1, 2$. Then there exists a constant $C = C(n, \varphi_1, \varphi_2, \alpha_1, \alpha_2)$ such that for every $f \in L^2(\mathbb{R}^n)$, the functions $\varphi^*_{i,L,\alpha} f = \sup|y - x| < \alpha_{1+\epsilon}|\varphi_i((1 + \epsilon)\sqrt{L})f(y)|$, $i = 1,2$, satisfy

$$\left\|\varphi^*_{1,L,\alpha_1} f\right\|_{L^{1-\epsilon}(\mathbb{R}^n)} \leq C\left\|\varphi^*_{2,L,\alpha_2} f\right\|_{L^{1-\epsilon}(\mathbb{R}^n)} \qquad (42)$$

As a consequence, for any even function $\varphi \in S(\mathbb{R})$ with $\varphi(0) = 1$ and $\alpha > 0$,

$$C^{-1}\|f^*_L\|_{L^{1-\epsilon}(\mathbb{R}^n)} \leq \left\|\varphi^*_{L,\alpha} f\right\|_{L^{1-\epsilon}(\mathbb{R}^n)} \leq \|f^*_L\|_{L^{1-\epsilon}(\mathbb{R}^n)}$$

Proof: Recall that for any $0 < \alpha_2 \leq \alpha_1$,

$$\left\|\varphi^*_{1,L,\alpha_1} f\right\|_{L^{1-\epsilon}(\mathbb{R}^n)} \leq C\left(1 + \frac{\alpha_1}{\alpha_2}\right)^{\frac{n}{1-\epsilon}} \left\|\varphi^*_{L,\alpha_2} f\right\|_{L^{1-\epsilon}(\mathbb{R}^n)}$$

for any $\varphi \in S(\mathbb{R})$ ([164]). Now, we let $\psi(x) := \phi_1(x) - \phi_2(x)$, and then the proof of (42)reduces to show that

$$\left\|\varphi^*_{L,\alpha_1} f\right\|_{L^{1-\epsilon}(\mathbb{R}^n)} \leq C\left\|\varphi^*_{2,L,1} f\right\|_{L^{1-\epsilon}(\mathbb{R}^n)} \qquad (43)$$

Let us show (43). Let $\Psi(x) = x^{2\kappa}\Phi(x)$ where $\Phi(x)$ is the function as in Lemma (6.1.4) and $2\kappa > \frac{n+1}{1-\epsilon}$. By the spectral theory ([53]), we have

$$f = C_{\Psi,\varphi_2} \int_0^\infty \Psi\big((1+2\epsilon)\sqrt{L}\big)\varphi_2((1+2\epsilon)\sqrt{L})f\,\frac{d(1+2\epsilon)}{1+2\epsilon}$$

Therefore,

$$\psi((1+\epsilon)\sqrt{L})f(x) = C\int_0^\infty \Big(\psi((1+\epsilon)\sqrt{L})\Psi\big((1+2\epsilon)\sqrt{L}\big)\Big)\,\varphi_2((1+2\epsilon)\sqrt{L})f(x)\,\frac{d(1+2\epsilon)}{1+2\epsilon}$$

Let us denote the kernel of $\psi((1+\epsilon)\sqrt{L})\Psi((1+2\epsilon)\sqrt{L})$ by $K_{\psi((1+\epsilon)\sqrt{L})\Psi((1+2\epsilon)\sqrt{L})}(x, y)$. Then for $\lambda \in (\frac{n}{1-\epsilon}, 2\kappa)$,

$$\sup_{|\omega|<(1+\epsilon)} \left|\psi\big((1+\epsilon)\sqrt{L}\big)f(x-\omega)\right|$$
$$= C\sup_{|\omega|<1+\epsilon}\left|\iint_{\mathbb{R}^{n+1}_+} K_{\psi((1+\epsilon)\sqrt{L})\Psi((1+2\epsilon)\sqrt{L})}(x-\omega, z)\varphi_2\big((1+2\epsilon)\sqrt{L}\big)f(z)\,\frac{dz\,d(1+2\epsilon)}{1+2\epsilon}\right|$$

$$\leq C \sup_{|\omega|<1+\epsilon} \int_{\mathbb{R}_+^{n+1}} \left|K_{\psi((1+\epsilon)\sqrt{L})\Psi((1+2\epsilon)\sqrt{L})}(x-\omega,z)\right| \left(1+\frac{|x-z|}{1+2\epsilon}\right)^{\lambda}$$

$$\times \left|\varphi_2((1+2\epsilon)\sqrt{L})f(z)\right| \left(1+\frac{|x-z|}{1+2\epsilon}\right)^{-\lambda} \frac{dzd(1+2\epsilon)}{1+2\epsilon}$$

$$\leq \sup_{1+2\epsilon,z} \left|\varphi_2\left((1+2\epsilon)\sqrt{L}\right)f(z)\right| \left(1+\frac{|x-z|}{1+2\epsilon}\right)^{-\lambda}$$

$$\times \sup_{|\omega|<1+\epsilon} \int_{\mathbb{R}_+^{n+1}} \left|K_{\psi\left((1+\epsilon)\sqrt{L}\right)\Psi\left((1+2\epsilon)\sqrt{L}\right)}(x-\omega,z)\right| \left(1 + \frac{|x-z|}{1+2\epsilon}\right)^{\lambda} \frac{dzd(1+2\epsilon)}{1+2\epsilon} \quad (44)$$

Next we will prove that

$$\sup_{|\omega|<1+\epsilon} \int_{\mathbb{R}_+^{n+1}} \left|K_{\psi((1+\epsilon)\sqrt{L})\Psi((1+2\epsilon)\sqrt{L})}(x-\omega,z)\right| \left(1+\frac{|x-z|}{1+2\epsilon}\right)^{\lambda} \frac{dzd(1+2\epsilon)}{1+2\epsilon} \leq C \quad (45)$$

Once estimate (45)is shown, (43)follows. Indeed, it follows from (44), (45)and the condition $\lambda \in (\frac{n}{1-\epsilon}, 2\kappa)$ that

$$\left\|\varphi_{L,\alpha_1}^* f\right\|_{L^{1-\epsilon}(\mathbb{R}^n)} = \left\|\sup_{|\omega|<1+\epsilon} \left|\psi((1+\epsilon)\sqrt{L})f(x-\omega)\right|\right\|_{L_x^{1-\epsilon}(\mathbb{R}^n)}$$

$$\leq C \left\|\sup_{z,1+2\epsilon} \left|\varphi_2((1+2\epsilon)\sqrt{L})f(z)\right| \left(1+\frac{|x-z|}{1+2\epsilon}\right)^{-\lambda}\right\|_{L_x^{1-\epsilon}(\mathbb{R}^n)}$$

$$\leq C \left\|\sup_{|y-x|<1+\epsilon} \left|\varphi_2((1+\epsilon)\sqrt{L})f(y)\right|\right\|_{L_x^{1-\epsilon}(\mathbb{R}^n)} = C\left\|\varphi_{2,L,1}^* f\right\|_{L^{1-\epsilon}(\mathbb{R}^n)}$$

where we used Theorem 2.4 of [164]in the second inequality.
Let us prove (45). Note that $|w| < 1+\epsilon$. We write

$$\psi\big((1+\epsilon)\sqrt{L}\big)\Psi\big((1+2\epsilon)\sqrt{L}\big)$$

$$= \begin{cases} \left(\frac{1+\epsilon}{1+2\epsilon}\right)^{2k} \left[\psi\big((1+2\epsilon)\sqrt{L}\big)\big((1+2\epsilon)\sqrt{L}\big)^{2k}\phi\big((1+\epsilon)\sqrt{L}\big)\right] & \text{if } \epsilon \geq 0 \\ \left(\frac{1+\epsilon}{1+2\epsilon}\right)^{2} \left[\big((1+\epsilon)\sqrt{L}\big)^{-2}\psi\big((1+\epsilon)\sqrt{L}\big)\big((1+2\epsilon)\sqrt{L}\big)^{2k+2}\phi\big((1+2\epsilon)\sqrt{L}\big)\right], & \text{if } \epsilon > 0 \end{cases}$$

We then apply Corollary (6.1.11) to obtain that for $\eta \in (\lambda, 2\kappa)$,

$$\left|K_{\psi((1+\epsilon)\sqrt{L})\Psi((1+2\epsilon)\sqrt{L})}(x-\omega,z)\right|$$

$$\leq Cmin\left(\left(\frac{1+2\epsilon}{1+\epsilon}\right)^{2k}, \left(\frac{1+\epsilon}{1+2\epsilon}\right)^{2}\right) \frac{\max(1+2\epsilon,1+\epsilon)^{\eta}}{(\max(1+2\epsilon,1+\epsilon)+|x-\omega-z|)^{n+\eta}}$$

This, together with the fact that

$$\int_{|u|\geq 1+2\epsilon} \frac{\max(1+2\epsilon,1+\epsilon)^{\eta}}{(\max(1+2\epsilon,1+\epsilon)+|u-\omega|)^{n+\eta}} \left(1+\frac{|u|}{1+2\epsilon}\right)^{\lambda} du$$

$$\leq \int_{|u|\geq 1+2\epsilon} \frac{\max(1+2\epsilon,1+\epsilon)^{\eta}}{(\max(1+2\epsilon,1+\epsilon)+|u-\omega|)^{n+\eta}} du$$

Shows

$$\int_{\mathbb{R}^n} \left|K_{\psi((1+\epsilon)\sqrt{L})\Psi((1+2\epsilon)\sqrt{L})}(x-\omega,z)\right| \left(1+\frac{|x-z|}{1+2\epsilon}\right)^{\lambda} dz$$

$$\leq Cmin\left(\left(\frac{1+2\epsilon}{1+\epsilon}\right)^{2k}, \left(\frac{1+\epsilon}{1+2\epsilon}\right)^{2}\right)\Bigg[1$$

$$+\int_{|u|\geq 1+2\epsilon} \frac{\max(1+2\epsilon,1+\epsilon)^{\eta}}{(\max(1+2\epsilon,1+\epsilon)+|u-\omega|)^{n+\eta}}\Bigg(1$$

$$+\frac{|u|}{1+2\epsilon}\Bigg)^{\lambda} du)\Bigg] \tag{46}$$

To estimate the integrals over $|u| \geq 1+2\epsilon$, we note that if $\epsilon \geq 0$, then we use the fact that $\eta > \lambda$ and $1+2\epsilon+|u-w| \geq 1+\epsilon+|u-w| \geq |w|+|u-w| \geq |u|$ to obtain

$$\int_{|u|\geq 1+2\epsilon} \frac{(1+2\epsilon)^{\eta}}{(1+\epsilon+|u-\omega|)^{n+\eta}}\left(1+\frac{|u|}{1+2\epsilon}\right)^{\lambda} du$$

$$\leq 2^{\lambda}\int_{|u|\geq 1+2\epsilon} \frac{(1+2\epsilon)^{\eta}}{(1+\epsilon+|u-\omega|)^{n+\eta}}\frac{|u|^{\lambda}}{(1+2\epsilon)^{\lambda}} du$$

$$\leq \int_{|u|\geq 1+2\epsilon} \frac{(1+2\epsilon)^{\eta}}{1+\epsilon+|u|^{n+\eta}}\frac{|u|^{\lambda}}{(1+2\epsilon)^{\lambda}} du \leq C \tag{47}$$

If $\epsilon > 0$, then it follows from the fact that $1+2\epsilon+|u-w| \geq |w|+|u-w| \geq |u|$ and $\eta > \lambda$,

$$\int_{|u|\geq 1+\epsilon} \frac{(1+2\epsilon)^{\eta}}{(1+2\epsilon+|u-\omega|)^{n+\eta}}\left(1+\frac{|u|}{1+\epsilon}\right)^{\lambda} du$$

$$\leq 2^{\lambda}\int_{|u|\geq 1+\epsilon} \frac{(1+2\epsilon)^{\eta}}{(1+2\epsilon+|u-\omega|)^{n+\eta}}\frac{|u|^{\lambda}}{(1+\epsilon)^{\lambda}} du$$

$$\leq \int_{|u|\geq 1+\epsilon} \frac{(1+2\epsilon)^{\eta}}{1+2\epsilon+|u|^{n+\eta}}\frac{|u|^{\lambda}}{(1+\epsilon)^{\lambda}} du$$

$$\leq C\left(\frac{1+2\epsilon}{1+\epsilon}\right)^{\eta} \tag{48}$$

Putting estimates (47)and (48)into (46), we have obtained that for any $|w| < 1+2\epsilon$,

$$\int_{\mathbb{R}^n} \left|K_{\psi((1+2\epsilon)\sqrt{L})\Psi((1+\epsilon)\sqrt{L})}(x-\omega,z)\right| \left(1+\frac{|x-z|}{1+\epsilon}\right)^{\lambda} dz$$

$$\leq Cmin\left(\left(\frac{1+\epsilon}{1+2\epsilon}\right)^{2k}, \left(\frac{1+\epsilon}{1+2\epsilon}\right)^{2}\right)\left[1+\max\left(1.\left(\frac{1+2\epsilon}{1+\epsilon}\right)^{\eta}\right)\right]$$

$$\leq Cmin\left(\left(\frac{1+\epsilon}{1+2\epsilon}\right)^{2k-\eta}, \left(\frac{1+2\epsilon}{1+\epsilon}\right)^{2}\right)$$

Observe that $\eta < 2\kappa$. It follows

$$\sup_{|\omega|<1+2\epsilon}\int_{\mathbb{R}^{n+1}_{+}} \left|K_{\psi((1+2\epsilon)\sqrt{L})\Psi((1+\epsilon)\sqrt{L})}(x-\omega,z)\right| \left(1+\frac{|x-z|}{1+\epsilon}\right)^{\lambda} \frac{dzd(1+\epsilon)}{1+\epsilon}$$

$$\leq C\int_0^\infty \min((\frac{1+\epsilon}{1+2\epsilon})^{2k-n},(\frac{1+2\epsilon}{1+\epsilon})^2)\frac{d(1+\epsilon)}{1+\epsilon}\leq C$$

which shows estimate (45), and the proof of Corollary (6.1.12) is end.

Corollary (6.1.13)[183]: Suppose that an operator L satisfies(H1)and(H2). For all $0\leq\epsilon\leq\infty$ and for all integers $M>\frac{n}{2}\left(\frac{\epsilon}{1-\epsilon}\right)$, we have that $H^{1-\epsilon}_{L,\max}(\mathbb{R}^n)\subseteq H^{1-\epsilon}_{L,a(1+\epsilon),1+\epsilon,M}(\mathbb{R}^n)$, and hence by (6),

$$H^{1-\epsilon}_{L,max}(\mathbb{R}^n)\simeq H^{1-\epsilon}_{L,a(1+\epsilon),1+\epsilon,M}(\mathbb{R}^n).$$

Proof. It suffices to show that for $f\in H^{1-\epsilon}_{L,max}(\mathbb{R}^n)\cap L^2(\mathbb{R}^n)$, f has a $(1-\epsilon,\infty,M)$atomic representation. We start with a suitable version of the Calderón repro-ducing formula. Let Φ be a function defined in Lemma (6.1.4), and set $\Psi(x):=x^{2M}\Phi(x),x\in\mathbb{R}$. By the spectral theory ([91]), for every $f\in L^2(\mathbb{R}^n)$ one can write

$$f=c_\Psi\int_0^\infty \Psi((1+2\epsilon)\sqrt{L})(1+2\epsilon)^2Le^{-(1+2\epsilon)^2L}\frac{d(1+2\epsilon)}{1+2\epsilon}$$

$$=\lim_{\epsilon\to 0}\int_\epsilon^{1/\epsilon}\Psi((1+2\epsilon)\sqrt{L})(1+2\epsilon)^2Le^{-(1+2\epsilon)^2L}\frac{d(1+2\epsilon)}{1+2\epsilon}\qquad(49)$$

with the integral converging in $L^2(\mathbb{R}^n)$.

Set

$$\eta(x)=c_\Psi\int_1^\infty(1+2\epsilon)^2x^2\Psi((1+2\epsilon)x)e^{-(1+2\epsilon)^2x^2}\frac{d(1+2\epsilon)}{1+2\epsilon}$$

$$=c_\Psi\int_x^\infty y\Psi(y)e^{-y^2}dy,x\neq 0$$

with $\eta(0)=1$. It follows that $\eta\in S(R)$is an even function, and

$$\eta(ax)-\eta(bx)=c_\Psi\int_a^b(1+2\epsilon)^2x^2\Psi((1+2\epsilon)x)e^{-(1+2\epsilon)^2x^2}f\frac{d(1+2\epsilon)}{1+2\epsilon}$$

By the spectral theory ([91]) again, one has

$$c_\Psi\int_a^b\Psi((1+2\epsilon)\sqrt{L})(1+2\epsilon)^2Le^{-(1+2\epsilon)^2L}f\frac{d(1+2\epsilon)}{1+2\epsilon}$$

$$=\eta\left(a\sqrt{L}\right)f(x)-\eta\left(b\sqrt{L}\right)f(x)\qquad(50)$$

Define,

$$M_Lf(x):=\sup_{|x-y|<5\sqrt{n(1+2\epsilon)}}\left(\left|(1+2\epsilon)^2Le^{-(1+2\epsilon)^2L}f(y)\right|+\left|\eta\left((1+2\epsilon)\sqrt{L}\right)f(y)\right|\right)$$

By Corollary (6.1.12), it follows that

$$\|M_Lf\|_{L^{1-\epsilon}(\mathbb{R}^n)}\leq C\|f\|_{H^{1-\epsilon}_{L,max}(\mathbb{R}^n)},\quad 0\leq\epsilon<1$$

Recall that $\mathbb{R}^{n+1}_+$denotes the upper half-space in $\mathbb{R}^{n+1}$. If O is an open subset of $\mathbb{R}^n$, then the "tent" over O, denoted by $\hat{O}$, is given as $\hat{O}:=\{(x,1+2\epsilon)\in\mathbb{R}^n+1+:B(x,4\sqrt{n(1+2\epsilon)})\subset O\}$. For $i\in\mathbb{Z}$, we define the family of sets $O_i:=\{x\in\mathbb{R}^n:M_Lf(x)>2^i\}$. Now let $\{Q_{ij}\}_j$ be a Whitney decomposition of O_i such that $O_i=\cup_j Q_{ij}$ and let $\hat{O}_i$ be a tent region. Set $\bar{e}=(1,\cdots,1)\in\mathbb{R}^n$. For every i,j, we define

$$\tilde{Q}_{ij}:=\left\{(y,1+2\epsilon)\in\mathbb{R}^{n+1}_+:y+3(1+2\epsilon)\bar{e}\in Q_{ij}\right\}\qquad(51)$$

It can be verified that $\hat{O}_i\subset\cup_j\tilde{Q}_{ij}$. Indeed, for each $(y^0,\,(1+2\epsilon)^0)\in\hat{O}_i$, we have that $B(y^0,4\sqrt{n(1+2\epsilon)^0})\subset O_i$. Let $\tilde{y}^0:=y^0+3\bar{e}(1+2\epsilon)^0$. Observe that $\tilde{y}^0\in$

$B(y^0, 4\sqrt{n(1+2\epsilon)^0})$and then $\tilde{y}^0 \in O_i$. Then there exists some $Q_{ij_0} \subset O_i$ such that $\tilde{y}^0 \in Q_{ij_0}$, hence $(y^0, (1+2\epsilon)^0) \in \tilde{Q}_{ij_0}$ and $\widehat{O}_i \subset \bigcup_j \tilde{Q}_{ij}$. Note that $\tilde{Q}_{ij} \cap \tilde{Q}_{ij'} = \emptyset$ when $j \neq j'$. We obtain an decomposition for $\mathbb{R}_+^{n+1}$ as follows:

$$\mathbb{R}_+^{n+1} = \cup_i \widehat{O}_\iota = \cup_i \widehat{O}_\iota / \widehat{O_{\iota+1}} = \cup_i \cup_j T_{ij}$$

where

$$T_{ij} \coloneqq \tilde{Q}_{ij} \cap \widehat{O}_\iota / \widehat{O_{\iota+1}}$$

Using the formula (49), one can write

$$f = \sum_{ij} c_\Psi \Psi\left((1+2\epsilon)\sqrt{L}\right)\left(\mathcal{X}T_{ij}(y, 1+2\epsilon)(1+2\epsilon)^2 L e^{-(1+2\epsilon)^2 L} f\right)\frac{d(1+2\epsilon)}{1+2\epsilon} \qquad (52)$$

with the sum converging in $L^2(\mathbb{R}^n)$, where $\lambda_{ij} := 2^i |Q_{ij}|^{\frac{1}{1-\epsilon}}$, $a_{ij} := L^M b_{ij}$, and

$$b_{ij} \coloneqq (\lambda_{ij})^{-1} c_\Psi \int_0^\infty (1+2\epsilon)^{2M} \phi\left((1+2\epsilon)\sqrt{L}\right)\left(\mathcal{X}T_{ij}(y, 1 + 2\epsilon)(1+2\epsilon)^2 L e^{-(1+2\epsilon)^2 L} f\right)\frac{d(1+2\epsilon)}{(1+2\epsilon)}$$

Let us show that the sum (52) converges in $L^2(\mathbb{R}^n)$. Indeed, since for each $f \in L^2(\mathbb{R}^n)$,

$$\left(\int_{\mathbb{R}_+^{n+1}} \left|(1+2\epsilon)^2 L e^{-(1+2\epsilon)^2\sqrt{L}} f(y)\right|^2 \frac{dy d(1+2\epsilon)}{1+2\epsilon}\right)^{1/2} \leq C\|f\|_{L^2(\mathbb{R}^n)}$$

we use (52)to obtain

$$\begin{aligned}
&\left\|\sum_{|i|>N_1, |j|>N_2} \lambda_{ij} a_{ij}\right\|_{L^2(\mathbb{R}^n)} \\
&= c_\Psi \left\|\sum_{|i|>N_1, |j|>N_2} \int_{\mathbb{R}^{n+1}} K_{((1+2\epsilon)^2 L)^M \Psi((1+2\epsilon)\sqrt{L})}(x,y) \mathcal{X}T_{ij}(y, 1 + 2\epsilon)(1+2\epsilon)^2 \times L e^{-(1+2\epsilon)^2 L} f(y) dy \frac{d(1+2\epsilon)}{1+2\epsilon}\right\|_{L^2(\mathbb{R}^n)} \\
&\leq \sup_{\|g\|_2 \leq 1} \sum_{|i|>N_1, |j|>N_2} \int_{(1+2\epsilon)_{ij}} \left|((1+2\epsilon)^2 L)^M \phi\left((1+2\epsilon)\sqrt{L}\right)(x,y) g(y) \times (1+2\epsilon)^2 L e^{-(1+2\epsilon)^2\sqrt{L}} f(y)\right| \frac{dy d(1+2\epsilon)}{1+2\epsilon} \\
&\leq C\left(\sum_{|i|>N_1, |j|>N_2} \int_{(1+2\epsilon)_{ij}} \left|(1+2\epsilon)^2 L e^{-(1+2\epsilon)^2\sqrt{L}} f(y)\right|^2 \frac{dy d(1+2\epsilon)}{1+2\epsilon}\right)^{1/2} \longrightarrow 0
\end{aligned}$$

as $N_1 \to \infty$, $N_2 \to \infty$.

Next, we will show that, up to a normalization by a multiplicative constant, the a_{ij} are $(1-\epsilon, \infty, M)$ −atoms. Once the claim is established, we shall have

$$\sum_{i,j} |\lambda_{ij}|^{1-\epsilon} = \sum_{i,j} 2^{i(1-\epsilon)} |Q_{ij}| \leq C \sum_i 2^{i(1-\epsilon)} |O_i| \leq C\|f\|_{H_{L,max}^{1-\epsilon}(\mathbb{R}^n)}^{1-\epsilon}$$

as desired.

Let us now prove that for every i, j, the function $C^{-1}a_{ij}$ is a $(1-\epsilon, \infty, M)$-atom associated with the cube $30Q_{ij}$ for some constant C. Observe that if $(y, 1+2\epsilon) \in T_{ij}$, then $B(y, 4\sqrt{n(1+2\epsilon)}) \in O_i$. Denote by $\tilde{y} := y + 3(1+2\epsilon)\bar{e}$, and so $\tilde{y} \in Q_{ij}$ and $B(\tilde{y}, \sqrt{n(1+2\epsilon)}) \in O_i$. The fact that Q_{ij} is the Whitney cube of O_i implies that $5Q_{ij} \cap O_i^c \neq \emptyset$. Denote the side length of Q_{ij} by $\ell(Q_{ij})$. It then follows that $1+2\epsilon \le 3\ell(Q_{ij})$. Since $y + 3\bar{e}(1+2\epsilon) \in Q_{ij}$, we have that $y \in 20Q_{ij}$. From Lemma (6.1.4), the integral kernel $K_{((1+2\epsilon)^2L)^k\Phi((1+2\epsilon)\sqrt{L})}$ of the operator $((1+2\epsilon)^2L)^k\Phi((1+2\epsilon)\sqrt{L})$ satisfies

$$\operatorname{supp}K_{((1+2\epsilon)^2L)^K\Psi((1+2\epsilon)\sqrt{L})} \subseteq \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n : |x-y| \le 1+2\epsilon\}$$

This concludes that for every $k = 0,1,\cdots, M$

$$\operatorname{supp}(L^K b_{ij}) \subseteq 30Q_{ij}$$

It remains to show that $\left\|(\ell(Q_{ij})^2L)^k b_{ij}\right\|_{L^\infty(\mathbb{R}^n)} \le C(\ell(Q_{ij}))^{2M}|Q_{ij}|^{-\frac{1}{1-\epsilon}}$, $k = 0, 1,\cdots, M$. When $K = 0, 1,\cdots, M-1$, it reduces to show

$$\left|\int_0^\infty \int_{\mathbb{R}^n} K_{(1+2\epsilon)^{2M}L^K\Psi((1+2\epsilon)\sqrt{L})}(x,y)\mathcal{X}T_{ij}(y, 1+2\epsilon)(1+2\epsilon)^2Le^{-(1+2\epsilon)^2L}f(y)dy\frac{d(1+2\epsilon)}{1+2\epsilon}\right| \le C2^i\ell(Q_{ij})^{2(M-K)}. \quad (53)$$

Indeed, if $\chi T_{ij}(y, 1+2\epsilon) = 1$, then $(y, 1+2\epsilon) \in (\widehat{O_{i+1}})^c$, and so $B(y, 4\sqrt{n(1+2\epsilon)}) \cap (O_{i+1})^c \neq \emptyset$. Let $\bar{x} \in B(y, 4\sqrt{n(1+2\epsilon)}) \cap (O_{i+1})^c$. We have that $|(1+2\epsilon)^2Le^{-(1+2\epsilon)^2L}f(y)| \le M_Lf(\bar{x}) \le 2^{i+1}$. By Lemma (6.1.4),

$$\left|\int_0^\infty \int_{\mathbb{R}^n} K_{(1+2\epsilon)^{2M}L^K\Psi((1+2\epsilon)\sqrt{L})}(x,y)\mathcal{X}T_{ij}(y, 1+2\epsilon)(1+2\epsilon)^2Le^{-(1+2\epsilon)^2L}f(y)dy\frac{d(1+2\epsilon)}{1+2\epsilon}\right|$$
$$\le C2^i\left|\int_0^{\ell(Q_{ij})}(1+2\epsilon)^{2(M-k)}\int_{\mathbb{R}^n}\left|K_{((1+2\epsilon)^2L)^K\Psi((1+2\epsilon)\sqrt{L})}(x,y)\right|dy\frac{d(1+2\epsilon)}{1+2\epsilon}\right|$$
$$\le C2^i\int_0^{\ell(Q_{ij})}(1+2\epsilon)^{2(M-k)}\frac{d(1+2\epsilon)}{1+2\epsilon} \le C2^i\ell(Q_{ij})^{2(M-K)}$$

since $K = 0,1,\cdots, M-1$.

Now we consider the case $k = M$. The proof is based on a modification of a technique due to A. Calderón [163]. In this case, we need to prove that for every i, j,

$$\left|\int_0^\infty \int_{\mathbb{R}^n} K_{\Psi((1+2\epsilon)\sqrt{L})}(x,y)\mathcal{X}T_{ij}(y, 1+2\epsilon)(1+2\epsilon)^2Le^{-(1+2\epsilon)^2L}f(y)dy\frac{d(1+2\epsilon)}{1+2\epsilon}\right| < C2^i \quad (54)$$

To show (54), we fix x and let $d(x, Q_{ij}) < 30\sqrt{n}\ell(Q_{ij})$. We claim the following result:

(P1)The properties of the set defining $\chi T_{ij}(y, 1+2\epsilon)$ imply that there exist intervals $(0, b_0), (a_1, b_1),\cdots, (a_N, \infty), 0 < b_0 \le a_1 < b_1 \le\cdots\le a_N, 1 \le N \le 2n+2$ such that for $l = 0, 1,\cdots, N-1$, there hold $a_{l+1} \le 3^{2n+2}b_l$ and

(a)$K_{\Psi((1+2\epsilon)\sqrt{L})}(x,y)\chi T_{ij}(y,1+2\epsilon)=0$ for $1+2\epsilon>a_N$;
(b)either $K_{\Psi((1+2\epsilon)\sqrt{L})}(x,y)\chi T_{ij}(y,1+2\epsilon)=0$ or $K_{\Psi((1+2\epsilon)\sqrt{L})}(x,y)\chi T_{ij}(y,1+2\epsilon)=K_{\Psi((1+2\epsilon)\sqrt{L})}(x,y)$ for all $1+2\epsilon\in(a_l,b_l)$;
(c)either $K_{\Psi((1+2\epsilon)\sqrt{L})}(x,y)\chi T_{ij}(y,1+2\epsilon)=0$ or $K_{\Psi((1+2\epsilon)\sqrt{L})}(x,y)\chi T_{ij}(y,1+2\epsilon)=K_{\Psi((1+2\epsilon)\sqrt{L})}(x,y)$ for all $1+2\epsilon\in(0,b_0)$.
Assuming this claim (P1)for the moment, we observe that for $d(x,Q_{ij})<30\sqrt{n}\ell(Q_{ij})$, one can write

$$\int_0^\infty\int_{\mathbb{R}^n}K_{\Psi((1+2\epsilon)\sqrt{L})}(x,y)\mathcal{X}T_{ij}(y,1+2\epsilon)(1+2\epsilon)^2Le^{-(1+2\epsilon)^2L}f(y)dy\frac{d(1+2\epsilon)}{1+2\epsilon}$$
$$=\left\{\int_0^{b_0}+\sum_{l=1}^{N-1}\int_{a_l}^{b_l}\right\}\int_{\mathbb{R}^n}K_{\Psi((1+2\epsilon)\sqrt{L})}(x,y)\mathcal{X}T_{ij}(y,1+2\epsilon)(1+2\epsilon)^2Le^{-(1+2\epsilon)^2L}f(y)dy\frac{d(1+2\epsilon)}{1+2\epsilon}$$
$$+\sum_{l=0}^{N-1}\int_{b_l}^{a_{l+1}}\int_{\mathbb{R}^n}K_{\Psi((1+2\epsilon)\sqrt{L})}(x,y)\mathcal{X}T_{ij}(y,1+2\epsilon)(1+2\epsilon)^2Le^{-(1+2\epsilon)^2L}f(y)dy\frac{d(1+2\epsilon)}{1+2\epsilon}$$
$$=I_1(x)+I_2(x) \tag{55}$$

To estimate $I_1(x)$, we note that if $a_l\le a<b\le b_l$ or $0\le a<b\le b_0$, then one has either

$$\int_a^b\int_{\mathbb{R}^n}K_{\Psi((1+2\epsilon)\sqrt{L})}(x,y)\mathcal{X}T_{ij}(y,1+2\epsilon)(1+2\epsilon)^2Le^{-(1+2\epsilon)^2L}f(y)dy\frac{d(1+2\epsilon)}{1+2\epsilon}$$

or by (50),

$$\int_a^b\int_{\mathbb{R}^n}K_{\Psi((1+2\epsilon)\sqrt{L})}(x,y)\mathcal{X}T_{ij}(y,1+2\epsilon)(1+2\epsilon)^2Le^{-(1+2\epsilon)^2L}f(y)dy\frac{d(1+2\epsilon)}{1+2\epsilon}$$
$$=\int_a^b\Psi((1+2\epsilon)\sqrt{L})(1+2\epsilon)^2Le^{-(1+2\epsilon)^2L}f(x)\frac{d(1+2\epsilon)}{1+2\epsilon}$$
$$=\eta\big(a\sqrt{L}\big)f(x)-\eta\big(b\sqrt{L}\big)f(x)$$

Observe that for each $a\le 1+2\epsilon\le b$, if $|x-y|<1+2\epsilon$, then $\chi T_{ij}(y,1+2\epsilon)=1$. This tells us that $(y,1+2\epsilon)\in(\widehat{O_{i+1}})^c$, hence $B(y,4\sqrt{n}(1+2\epsilon))\cap(O_{i+1})^c\neq\emptyset$. Assume that $\bar{x}\in B(y,4\sqrt{n}(1+2\epsilon))\cap(O_{i+1})^c$. From this, we have that $|x-\bar{x}|\le|x-y|+|y-\bar{x}|<5\sqrt{n}(1+2\epsilon)$ and $M_Lf(\bar{x})\le 2^{i+1}$. It implies that $|\eta((1+2\epsilon)\sqrt{L})f(x)|\le M_Lf(\bar{x})\le C2^{i+1}$for every $a\le 1+2\epsilon\le b$. Therefore, $|\eta(a\sqrt{L})f(x)|\le C2^{i+1}$ and $|\eta(b\sqrt{L})f(x)|\le C2^{i+1}$, and so $|I_1(x)|\le C2^{i+1}$.

Consider $I_2(x)$. If $\chi T_{ij}(y,1+2\epsilon)=1$, then $(y,1+2\epsilon)\in(\hat{O}_{i+1})^c$. Thus $B(y,4\sqrt{n}(1+2\epsilon))\cap(O_{i+1})^c=\emptyset$. Assume that $\bar{x}\in B(y,4\sqrt{n}(1+2\epsilon))\cap(O_{i+1})^c$. We have that $|(1+2\epsilon)^2Le^{-(1+2\epsilon)^2L}f(y)|\le M_Lf(\bar{x})\le 2^{i+1}$. This, together with $a_{l+1}\le cb_l$, implies that

$$\left|\int_{b_l}^{a_{l+1}}\int_{\mathbb{R}^n} K_{\Psi((1+2\epsilon)\sqrt{L})}(x,y)\mathcal{X}T_{ij}(y,1+2\epsilon)(1+2\epsilon)^2 L e^{-(1+2\epsilon)^2 L} f(y)dy\frac{d(1+2\epsilon)}{1+2\epsilon}\right|$$
$$\leq 2^{i+1}\left|\int_{b_l}^{cb_l}\int_{\mathbb{R}^n}\left|K_{\Psi((1+2\epsilon)\sqrt{L})}(x,y)\right|dy\frac{d(1+2\epsilon)}{1+2\epsilon}\right|$$
$$\leq C2^{i+1}\int_{b_l}^{cb_l}\frac{1}{1+2\epsilon}d(1+2\epsilon)\leq C2^{i+1} \quad (56)$$
which yields that $|I_2(x)|\leq C2^{i+1}$.

Combining (55) and (56), we obtain (54). It follows that $\|a_{ij}\|_{L^\infty(\mathbb{R}^n)}\leq C|Q_{ij}|^{-\frac{1}{1-\epsilon}}$. Up to a normalization by a multiplicative constant, the a_{ij} are $(1-\epsilon,\infty,M)$-atoms.

It remains to prove the claim(P1). Note that $\mathcal{X}T_{ij}(y,1+2\epsilon)=\chi_{\widehat{O_l}}(y,1+2\epsilon)\cdot\chi_{(\widehat{O_{l+1}})^c}(y,1+2\epsilon)\cdot\chi_{\widetilde{Q_{ij}}}(y,1+2\epsilon)$; Assume that $Q_{ij}=\{(y_1,\cdots,y_n): c_l\leq y_l\leq d_l, l=1,\cdots,n\}$. Then

$$\chi_{\check{Q}_{ij}}(y,1+2\epsilon)=\prod_{l=1}^{n}\chi_{\{c_l\leq y_l+3(1+2\epsilon)\leq d_l\}}(y,1+2\epsilon)$$
$$=\prod_{l=1}^{n}\chi_{\{y_l+3(1+2\epsilon)\geq c_l\}}(y,1+2\epsilon)\chi_{\{y_l+3(1+2\epsilon)\leq d_l\}}(y,1+2\epsilon)$$

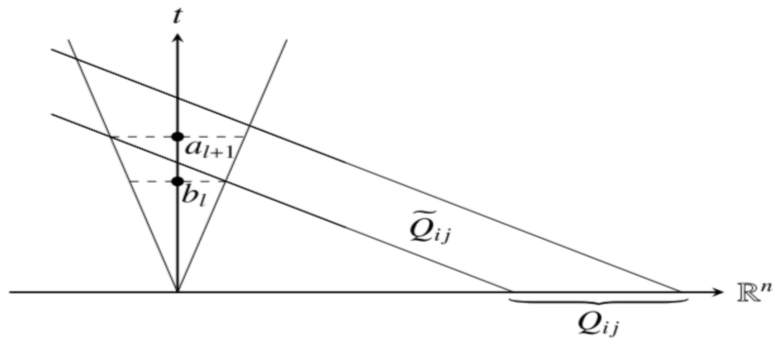


Fig. (1)[183]: The case of $x_l<c_l$.

Let $\chi_l(y,1+2\epsilon)$ be one of the characteristic functions $\chi_{\{y_l+3(1+2\epsilon)\geq c_l\}}(y,1+2\epsilon),\chi_{\{y_l+3(1+2\epsilon)\leq d_l\}}(y,1+2\epsilon),\chi_{\widehat{O_l}}(y,1+2\epsilon)$ *and* $\chi_{(\widehat{O_{l+1}})^c}(y,1+2\epsilon)$. We will show the following property:

(P2)There exist numbers b_l and $a_{l+1}$with $0<b_l\leq a_{l+1}$ and $a_{l+1}\leq 3b_l$ such that for every x, either $K_{\Psi((1+2\epsilon)\sqrt{L})}(x,y)\chi_l(y,1+2\epsilon)=0$ or $K_{\Psi((1+2\epsilon)\sqrt{L})}(x,y)\chi_l(y,1+2\epsilon)=K_{\Psi((1+2\epsilon)\sqrt{L})}(x,y)$ for all tin each of the intervals complementary to (b_l,a_{l+1}). And for at least one of $\chi_l(y,1+2\epsilon),K_{\Psi((1+2\epsilon)\sqrt{L})}(x,y)\chi_l(y,1+2\epsilon)=0$ for $1+2\epsilon>a_{l+1}$.

Then the same holds for $\chi T_{ij}(y,1+2\epsilon)=\prod_{l=1}^{2n+2}\chi_l(y,1+2\epsilon)K_{\Psi((1+2\epsilon)\sqrt{L})}(x,y)$ in each of the intervals complementary to the union of the intervals (b_l,a_{l+1}), which is what was asserted in the claim. Thus we merely have to prove(P2). To do this, we consider four cases.

Case 1. $\chi_l(y,1+\epsilon)=\chi\{y_l+3(1+\epsilon)\geq c_l\}(y,1+\epsilon)$.

In this case, since $\mathrm{supp} K_{\Psi((1+\epsilon)\sqrt{L})}(x,y) \subseteq \{y: |x-y| \leq (1+\epsilon)\}$, we have that $\mathrm{supp} K_{\Psi((1+\epsilon)\sqrt{L})}(x,y) \subseteq \{y: x_l - (1+\epsilon) \leq y_l \leq x_l + 1 + \epsilon\}$. If $x_l \geq c_l$, then $y_l + 3(1+\epsilon) \geq x_l + 2(1+\epsilon) \geq c_l$ for any $\epsilon \geq 0$. This yields

$$K_{\Psi((1+\epsilon)\sqrt{L})}(x,y)\chi_l(y, 1+\epsilon) = K_{\Psi((1+\epsilon)\sqrt{L})}(x,y), \qquad \epsilon \geq 0.$$

If $x_l < c_l$, then we choose $b_l = \frac{c_l - x_l}{4}$ and $a_{l+1} = \frac{c_l - x_l}{2}$ (see Fig.1).

In the case of $(1+\epsilon) < b_l$, we have $y_l + 3(1+\epsilon) \leq x_l + 4(1+\epsilon) < c_l$, which implies that $K_{\Psi((1+\epsilon)\sqrt{L})}(x,y)\chi T_{ij}(y, 1+\epsilon) = 0$. In the case of $1+\epsilon > a_{l+1}$, we have $y_l + 3(1+\epsilon) \geq x_l + 2(1+\epsilon) > c_l$. This implies that $K_{\Psi((1+\epsilon)\sqrt{L})}(x,y)\chi_l(y, 1+\epsilon) = K_{\Psi((1+\epsilon)\sqrt{L})}(x,y)$.

Case 2. $\chi_l(y, 1+\epsilon) = \chi\{y_l + 3(1+\epsilon) \leq d_l\}(y, 1+\epsilon)$.

Since $\mathrm{supp} K_{\Psi((1+\epsilon)\sqrt{L})}(x,y) \subseteq \{y: |x-y| \leq 1+\epsilon\}$, we have that $K_{\Psi((1+\epsilon)\sqrt{L})}(x,y) \subseteq \{y: x_l - (1+\epsilon) \leq y_l \leq x_l + 1 + \epsilon\}$. When $x_l \geq d_l$, we have that $y_l + 3(1+\epsilon) \geq x_l + 2(1+\epsilon) > dl$ for any $\epsilon \geq 0$. This tells us

$$K_{\Psi((1+\epsilon)\sqrt{L})}(x,y)\chi_l(y, 1+\epsilon) = 0, \qquad \text{for } \epsilon \geq 0$$

When $x_l < d_l$, we choose $b_l = \frac{d_l - x_l}{4}$ and $a_{l+1} = \frac{d_l - x_l}{2}$. If $1+\epsilon < b_l$, then $y_l + 3(1+\epsilon) \leq x_l + 4(1+\epsilon) < d_l$, which implies that $K_{\Psi((1+\epsilon)\sqrt{L})}(x,y)\chi_l(y, 1+\epsilon) = K_{\Psi((1+\epsilon)\sqrt{L})}(x,y)$. If $1+\epsilon > a_{l+1}$, then $y_l + 3(1+\epsilon) \geq x_l + 2(1+\epsilon) > d_l$. From this, we have that $K_{\Psi((1+\epsilon)\sqrt{L})}(x,y)\chi_l(y, 1+\epsilon) = 0$.

Case 3. $\chi_l(y, 1+\epsilon) = \chi + O_i(y, 1+\epsilon)$.

In this case, we choose $b_l = \frac{1}{5\sqrt{n}} d(x, O_i^c)$ and $a_{l+1} = \frac{1}{2\sqrt{n}} d(x, O_i^c)$. Let $|x-y| < 1+\epsilon$. If $1+\epsilon < \frac{1}{5\sqrt{n}} d(x, O_i^c)$, then $d(y, O_i^c) \geq d(x, O_i^c) - |x-y| > 5\sqrt{n}(1+\epsilon) - (1+\epsilon) \geq 4\sqrt{n}(1+\epsilon)$. This tells us

$$K_{\Psi((1+\epsilon)\sqrt{L})}(x,y)\chi_l(y, 1+\epsilon) = K_{\Psi((1+\epsilon)\sqrt{L})}(x,y)$$

for $1+\epsilon < \frac{1}{5\sqrt{n}} d(x, O_i^c)$. If $1+\epsilon > \frac{1}{2\sqrt{n}} d(x, O_i^c)$, then $d(y, O_i^c) \leq d(x, O_i^c) + d(x,y) < (2\sqrt{n}+1)(1+\epsilon) < 4\sqrt{n}(1+\epsilon)$. Hence, if $1+\epsilon > \frac{1}{2\sqrt{n}} d(x, O_i^c)$, then

$$K_{\Psi((1+\epsilon)\sqrt{L})}(x,y)\chi_l(y, 1+\epsilon) = 0$$

Case 4. $\chi_l(y, 1+\epsilon) = \chi(\widehat{O_{l+1}})^c(y, 1+\epsilon)$.

In this case, we can choose $b_l = \frac{1}{5\sqrt{n}} d(x, O_{i+1}^c)$ and $a_{l+1} = \frac{1}{2\sqrt{n}} d(x, O_{i+1}^c)$. The proof can be an adaptation of the proof as in Case 3, and we omit the detail here.

This concludes the proof of the property (P2). We have obtained the claim (P1), and then the proof of Corollary (6.1.13) is complete.

Corollary (6.1.14)[183]: Suppose that an operator L satisfies (H1)–(H2). Fix $0 \leq \epsilon < 1$. For all $\epsilon > 0$ with $0 \leq \epsilon \leq \infty$ and for all integers $M > \frac{n}{2}\left(\frac{\epsilon}{1-\epsilon}\right)$, if $f \in L^2(\mathbb{R}^n)$, then the following conditions on fare equivalent:

(i) $f \in H^{1-\epsilon}_{L,\alpha(1+\epsilon),1+\epsilon,M}(\mathbb{R}^n)$;

(ii) Given $\alpha > 0$, $\varphi^*_{L,\alpha} f = \sup_{|y-x|<\alpha(1+\epsilon)} |\phi((1+\epsilon)\sqrt{L})f(y)| \in L^{1-\epsilon}(\mathbb{R}^n)$ for some even function $\varphi \in S(R)$, $\varphi(0) = 1$;

(iii) $G_L^*(f) = \sup_{\phi\in A} \sup_{|y-x|<1+\epsilon} |\varphi((1+\epsilon)\sqrt{L})f(y)| \in L^{1-\epsilon}(\mathbb{R}^n)$,

$$A = \{\varphi \in L\,(\mathbb{R})\text{: even function with } \varphi(0) \neq 0, \int_{\mathbb{R}} (1+|x|)^N \sum_{k\leq N} \left|\frac{d^k}{dx^k}\varphi(x)\right|^2 dx \leq 0\}$$

where N is a large number depending only on $1-\epsilon$ and n.

Proof: The proof follows the line of (ii)⇒(i) ⇒(iii)⇒(ii). From Corollary (6.1.13) and Corollary (6.1.12), it tells us an implication (ii) ⇒(i). The proof of (i) ⇒(iii) will be an adaptation of the proof of the earlier known implication of (i) ⇒(ii) (see [171]). Obviously, (iii)⇒(ii). This proves Corollary (6.1.14).

We consider an electromagnetic Laplacian

$$L = (i\nabla - A(x))^2 + V(x), \;\; n \geq 3$$

Recall that a measurable function Von $\mathbb{R}^n$ is in the Kato class when

$$\lim_{r\downarrow 0} \sup_x \int_{|x-y|<r} \frac{|V(y)|}{|x-y|^{n-2}} dy = 0$$

while the Kato norm is defined by

$$\|V\|_K = \sup_x \int \frac{|V(y)|}{|x-y|^{n-2}} dy = 0$$

Corollary (6.1.15)[183]: Consider an electromagnetic Laplacian

$$L = (i\nabla - A(x))^2 + V(x), \;\; n \geq 3$$

Assume that $A \in L^2_{loc}(\mathbb{R}^n, \mathbb{R}^n)$, and the positive and negative parts $V_\pm$ of V satisfy V_+ is of Kato class, $\|V_-\|_K < c_n = \pi^{n/2}/\Gamma(n/2\,-1)$. Then for every integer $M > \frac{n}{2}\left(\frac{\epsilon}{1-\epsilon}\right)$, the spaces $H^{1-\epsilon}_{L,max}(\mathbb{R}^n)$ and $H^{1-\epsilon}_{L,\text{at},\infty,\text{M}}(\mathbb{R}^n)$ coincide. In particular,

$$\|f\|_{H^{1-\epsilon}_{L,max}(\mathbb{R}^n)} \approx \|f\|_{H^{1-\epsilon}_{L,\text{at},\infty,\text{M}}(\mathbb{R}^n)}$$

Proof: It is known (see [162]) that under assumptions of Corollary (6.1.15), the operator L has a unique nonnegative self-adjoint extension, and $e^{-(1+\epsilon)L}$ is an integral operator whose kernel satisfies the Gaussian estimate (H2). Now Corollary (6.1.15) is a straightforward consequence of Corollary (6.1.13).

Section (6.2): Weak Factorizations and Commutators

The spaces $H^1(\mathbb{R}^n)$ and $BMO(\mathbb{R}^n)$ are fundamental function spaces in harmonic analysis. The work of Fefferman and Stein, [74], provides a duality relationship between $H^1(\mathbb{R}^n)$and $BMO(\mathbb{R}^n)$. And, further provides characterizations of these spaces in terms of maximal functions, square functions, and Riesz transforms. While the work of Coifman, Rochberg and Weiss, [3], provides a connection between weak factorization of the Hardy spaces, commutators with Riesz transforms and $BMO(\mathbb{R}^n)$. We provide similar connections for $H^1$and BMO spaces adapted to a particular linear differential operator.

There is a substantial literature related to $H^1$and BMO spaces adapted to a linear operator L on $L^2(\mathbb{R}^n)$which generates an analytic semigroup e^{-tL}on $L^1(\mathbb{R}^n)$ with a kernel $p_t(x,y)$ satisfying an upper bound. That is, operators L for which the kernel of the semigroup $p_t(x,y)$ there exists positive constants m andϵsuch that for all $x, y \in \mathbb{R}^n$ and for all$t > 0$:

$$|p_t(x,y)\,| \leq \frac{Ct^{\frac{\epsilon}{m}}}{\left(t^{\frac{1}{m}} + |x-y|\right)^{n+\epsilon}}. \tag{57}$$

In[58], Auscher, Duong, and McIntosh defined a Hardy space $H_L^1(\mathbb{R}^n)$associated with such operators L as the class of all functions $f \in L^1(\mathbb{R}^n)$ for which $S_L \in L^1(\mathbb{R}^n)$, where $S_L(f)$is Littlewood–Paley area function defined as follows.

$$S_L(f)(x) = \left(\int_0^\infty \int_{|y-x|<t} |Q_{t^m} f(y)|^2 \frac{dydt}{t^{n+1}} \right)^{\frac{1}{2}}, \tag{58}$$

with $Q_t = tLe^{-tL}$. The $H_L^1(\mathbb{R}^n)$norm of f is defined as$\|f\|_{H_L^1(\mathbb{R}^n)} = \|S_L(f)\|_{L^1(\mathbb{R}^n)}$. In [70], [71], Duong and Yan defined the function space$BMO_L(\mathbb{R}^n)$associated with an operator$L$. They then go on to prove that if L has a bounded holomorphic functional calculus on$L^2(\mathbb{R}^n)$and the kernel $p_t(x,y)$ of the semigroup e^{-tL} satisfies the upper bound(57), then the space $BMO_{L^*}(\mathbb{R}^n)$ is the dual space of Hardy space $H_L^1(\mathbb{R}^n)$ in which L^* denotes the adjoint operator of L. This gives a generalization of the duality of $H^1(\mathbb{R}^n)$ and $BMO(\mathbb{R}^n)$of Fefferman and Stein [74]. Later, the theory of function spaces associated with operators has been developed and generalized to many other different settings, see [160], [168], [172], [171], [174].

The choice of $L = \Delta$ gives rise to the spaces to the classical spaces$H^1(\mathbb{R}^n)$and $BMO(\mathbb{R}^n)$. While the choice of the semigroup $e^{-tL}$is the Poisson semigroup $e^{-t\sqrt{\Delta}}$(here $m = 2$), given by

$$e^{-t\sqrt{\Delta}} f(x) = \int_{\mathbb{R}^n} p_t(x-y) f(y)\, dy, t > 0, p_t(x) = \frac{c_n t}{(t^2 + |x|^2)^{\frac{n+1}{2}}} \tag{59}$$

yields the spaces $H_{\sqrt{\Delta}}^1(\mathbb{R})$and $BMO_{\sqrt{\Delta}}(\mathbb{R})$coincide with the classical Hardy space and BMO space, respectively (see[58] and[70]).

In [65], Deng, Duong, Sikora, and Yan further considered the comparison of $BMO_L(\mathbb{R}^n)$and $BMO(\mathbb{R}^n)$. By considering the Neumann Laplacian $L = \Delta_N$, they obtained that

$$BMO(\mathbb{R}^n) \subsetneq BMO_{\Delta_N}(\mathbb{R}^n).$$

Recently, in[53] Yan introduced a class of$H_L^p(\mathbb{R}^n)$for a range of $p \le 1$ by using the Littlewood–Paley area function$S_L(f)$. In particular, Yan showed that

$$H_{\Delta_N}^p(\mathbb{R}^n) \subsetneq H^p(\mathbb{R}^n), \frac{n}{n+1} < p \le 1.$$

We carry out a deeper study of the spaces $H_{\Delta_N}^1(\mathbb{R}^n)$ and $BMO_{\Delta_N}(\mathbb{R}^n)$. Interestingly, we show that these spaces behave in an analogous fashion as the standard Hardy space$H^1(\mathbb{R}^n)$and $BMO(\mathbb{R}^n)$.

We first explicitly compute the Riesz transforms$R_N = \nabla \Delta_N^{-\frac{1}{2}}$ associated to the Neumann Laplacian. Because of the close connection between the Laplacian and Neumann Laplacian, we find in Proposition(6.2.4) that the Riesz transforms associated to the Neumann Laplacian are given by an additive perturbation of the standard Riesz transforms. We show that, similar to the classical Hardy space, the space$H_{\Delta_N}^1(\mathbb{R}^n)$can be characterized by the radial and non-tangential maximal functions, by the Riesz transforms, and by atoms, all of which are defined in terms of the Neumann Laplacian Δ_N. To be more precise, we denote by$H_{\Delta_N,max}^1(\mathbb{R}^n)$the Hardy space defined via the radial maximal function associated withΔ_N, and analogously by$H_{\Delta_N,*}^1(\mathbb{R}^n)$, $H_{\Delta_N,Riesz}^1(\mathbb{R}^n)$ and

$H^1_{\Delta_N,atom}(\mathbb{R}^n)$ the Hardy spaces via non-tangential maximal functions, Riesz transforms and atoms, respectively. Then we have the following characterizations.

Theorem (6.2.1)[178]: Let all notation be the same as above. We have

$$H^1_{\Delta_N}(\mathbb{R}^n) = H^1_{\Delta_N,max}(\mathbb{R}^n) = H^1_{\Delta_N,*}(\mathbb{R}^n) = H^1_{\Delta_N,Riesz}(\mathbb{R}^n) = H^1_{\Delta_N,atom}(\mathbb{R}^n)$$

and with equivalent norms

$$\|f\|_{H^1_{\Delta_N}(\mathbb{R}^n)} \approx \|f\|_{H^1_{\Delta_N,max}(\mathbb{R}^n)} \approx \|f\|_{H^1_{\Delta_N,Riesz}(\mathbb{R}^n)} \approx \|f\|_{H^1_{\Delta_N,atom}(\mathbb{R}^n)}$$
$$\approx \|f_{+,e}\|_{H^1(\mathbb{R}^n)} + \|f_{-,e}\|_{H^1(\mathbb{R}^n)}.$$

Here $f_{\pm,e}$is the even extension of the restriction of f from$\mathbb{R}^n_{\pm}$. Namely,$f \in H^1_{\Delta_N}(\mathbb{R}^n)$if and only if $f_{+,e} \in H^1(\mathbb{R}^n)andf_{-,e} \in H^1(\mathbb{R}^n)$.

We also obtain a Fefferman–Stein decomposition of $BMO_{\Delta_N}(\mathbb{R}^n)$in terms of the action of the Riesz transforms associated to the Neumann Laplacian on $L^\infty(\mathbb{R}^n)$ functions.

We then further show the connection between$BMO_{\Delta_N}(\mathbb{R}^n)$,$H^1_{\Delta_N}(\mathbb{R}^n)$, commutators of functions in $BMO_{\Delta_N}(\mathbb{R}^n)$and Riesz transforms$R_N$relative to$\Delta_N$, and a weak factorization of the space$H^1_{\Delta_N}(\mathbb{R}^n)$. We have the following theorem.

Theorem (6.2.2)[178]: For $1 \le l \le n$, let$\Pi_l(h,g) = h.R^*_{N,l}(g) - g.R_{N,l}(h)$, where$R_{N,l} = \frac{\partial}{\partial x_1}\Delta_N^{-\frac{1}{2}}$is the l-th Riesz transform associated to the Neumann Laplacian and$R^*_{N,l}$ is the adjoint operator of$R_{N,l}$. Then for any$f \in H^1_{\Delta_N}(\mathbb{R}^n)$ there exists sequences$\{\lambda^*_j\} \in \ell^1$and functions $g^k_j, h^k_j \in L^\infty(\mathbb{R}^n)$with compact supports such that$f = \sum_{k=1}^\infty \sum_{j=1}^\infty \lambda^k_j \Pi_l(g^k_j, h^k_j)$. Moreover, we have that:

$$\|f\|_{H^1_{\Delta_N}(\mathbb{R}^n)} \approx \inf\left\{\sum_{k=1}^\infty\sum_{j=1}^\infty |\lambda^k_j| \|g^k_j\|_{L^2(\mathbb{R}^n)} \|h^k_j\|_{L^2(\mathbb{R}^n)} : f \sum_{k=1}^\infty\sum_{j=1}^\infty \lambda^k_j \Pi_l(g_j, h_j)\right\}.$$

We then obtain the following new characterization of$BMO_{\Delta_N}(\mathbb{R}^n)$in terms of the commutators with the Riesz transforms associated toΔ_N.

We point out that Theorem(6.2.2) and Theorem(6.2.27) can be extended to work for $L^p(\mathbb{R}^n)$ when $1 < p < \infty$.

For$0 < \alpha < n$, the fractional operator $\Delta_N^{-\alpha/2}$of the operator$\Delta_N$is defined by

$$\Delta_N^{-\alpha/2} f(x) = \frac{1}{\Gamma(\alpha/2)} \int_0^\infty e^{-t\Delta_N}(f)(x) \frac{dt}{t^{1-\alpha/2}}.$$

We collect the background for the Neumann Laplacian and the associated Riesz transforms. The related Hardy and BMO spaces associated to $\Delta_N$are studied and their basic properties are collected. In particular, we demonstrate a collection of equivalent norms for $H^1_{\Delta_N}(\mathbb{R}^n)$, Theorem (6.2.18), and show the Fefferman–Stein decomposition of $BMO_{\Delta_N}(\mathbb{R}^n)$holds, Corollary (6.2.19). We provide the proof of Theorems (6.2.2) and (6.2.27). The letter "C" will denote, possibly different, constants that are independent of the essential variables.

We now recall some notation and basic facts introduced in [65]. For any subset $A \subset \mathbb{R}^n$and a function$f: \mathbb{R}^n \to \mathbb{C}$ by$f|_A$we denote the restriction of f to A. Next we set$\mathbb{R}^n_+ = \{(x', x_n) \in \mathbb{R}^n : x' = (x_1, \dots, x_{n-1}) \in \mathbb{R}^{n-1}, x_n > 0\}$. For any function f on$\mathbb{R}^n$, we set

$$f_+ = f|_{\mathbb{R}^n_+} \text{and} f_- = f|_{\mathbb{R}^n_-}.$$

For any $x = (x', x_n) \in \mathbb{R}^n$ we set $\tilde{x} = (x' - x_n)$. If f is any function defined on $\mathbb{R}^n_+$, its even extension defined on $\mathbb{R}^n$ is

$$f_e(x) = f(x), \text{if } x \in \mathbb{R}^n_+; f_e(x) = f(\tilde{x}), \text{if } x \in \mathbb{R}^n_-. \quad (60)$$

We denote by Δ_n the Laplacian on $\mathbb{R}^n$. Next we recall the Neumann Laplacian on $\mathbb{R}^n_+$ and $\mathbb{R}^n_-$.

Consider the Neumann problem on the half line $(0, \infty)$ (see [181]):

$$\begin{cases} w_t - w_{xx} = 0 \text{ for } 0 < x < \infty, 0 < t < \infty, \\ w(x, 0) = \phi(x), \\ w_x(0, t) = 0. \end{cases} \quad (61)$$

Denote this corresponding Laplacian by Δ_{1,N_+}. According to [181], we see that

$$w(x, t) = e^{-t\Delta_{1,N_+}}(\phi)(x).$$

For $n > 1$, we write $\mathbb{R}^n_+ = \mathbb{R}^{n-1} \times \mathbb{R}_+$. And we define the Neumann Laplacian on $\mathbb{R}^n_+$ by

$$\Delta_{n,N_+} = \Delta_{n-1} + \Delta_{1,N_+},$$

where Δ_{n-1} is the Laplacian on $\mathbb{R}^{n-1}$ and Δ_{1,N_+} is the Laplacian corresponding to (61). Similarly we can define Neumann Laplacian Δ_{n,N_-} on $\mathbb{R}^n_-$.

We skip the index n, we denote by Δ the Laplacian on $\mathbb{R}^n$, denote the Neumann Laplacian on $\mathbb{R}^n_+$ by Δ_{N_+}, and Neumann Laplacian on $\mathbb{R}^n_-$ by Δ_{N_-}.

The Laplacian and Neumann Laplacian $\Delta_{N_\pm}$ are positive definite self-adjoint operators. By the spectral theorem one can define the semigroups generated by these operators $\{\exp(-t\Delta), t \geq 0\}$ and $\{\exp(-t\Delta_{N_\pm}), t \geq 0\}$. By $p_t(x, y)$, $p_{t,\Delta_{N_+}}(x, y)$ and $p_{t,\Delta_{N_-}}(x, y)$ we denote the heat kernels corresponding to the semigroups generated by Δ, Δ_{N_+}, and Δ_{N_-}, respectively. Then we have

$$p_t(x, y) = \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x-y|^2}{4t}}.$$

From the reflection method (see [181]), we get

$$p_{t,\Delta_{N_+}}(x, y) = \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x'-y'|^2}{4t}} \left(e^{-\frac{|x_n - y_n|^2}{4t}} + e^{-\frac{|x_n - y_n|^2}{4t}} \right), x, y \in \mathbb{R}^n_+;$$

$$p_{t,\Delta_{N_-}}(x, y) = \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x'-y'|^2}{4t}} \left(e^{-\frac{|x_n - y_n|^2}{4t}} + e^{-\frac{|x_n - y_n|^2}{4t}} \right), x, y \in \mathbb{R}^n_-.$$

For any function f on $\mathbb{R}^n_+$, we have

$$\exp(-t\Delta_{N_+}) f(x) = \exp(-t\Delta) f_e(x)$$

for all $t \geq 0$ and $x \in \mathbb{R}^n_+$. Similarly, for any function f on $\mathbb{R}^n_-$,

$$\exp(-t\Delta_{N_-}) f(x) = \exp(-t\Delta) f_e(x)$$

for all $t \geq 0$ and $x \in \mathbb{R}^n_-$.

Now let Δ_N be the uniquely determined unbounded operator acting on $L^2(\mathbb{R}^n)$ such that

$$(\Delta_N f)_+ = \Delta_{N_+} f_+ \text{ and } (\Delta_N f)_- = \Delta_{N_-} f_- \quad (62)$$

for all $f: \mathbb{R}^n \to \mathbb{R}$ such that $f_+ \in W^{1,2}(\mathbb{R}^n_+)$ and $f_- \in W^{1,2}(\mathbb{R}^n_-)$. Then Δ_N is a positive self-adjoint operator and

$$(\exp(-t\Delta_N) f)_+ = \exp(-t\Delta_{N_+}) f_+ \text{ and } (\exp(-t\Delta_N) f)_- = \exp(-t\Delta_{N_-}) f_-.$$

The heat kernel of $\exp(-t\Delta_N)$, denoted by $p_{t,\Delta_N}(x, y)$, is then given as:

$$p_{t,\Delta_N}(x,y)=\frac{1}{(4\pi t)^{\frac{n}{2}}}e^{-\frac{|x'-y'|^2}{4t}}\left(e^{-\frac{|x_n-y_n|^2}{4t}}+e^{-\frac{|x_n-y_n|^2}{4t}}\right)H(x_ny_n),\tag{63}$$

where$H$: $\mathbb{R}\{0,1\}$is the Heaviside function given by

$$H(t)=0,\text{if } t<0; H(t)=1\text{ ,if } t\geq 0.\tag{64}$$

Let us note that

(α) All the operatorsΔ, Δ_{N_+}, Δ_{N_-}, and$\Delta_N$are self-adjoint and they generate bounded analytic positive semigroups acting on all$L^p(\mathbb{R}^n)$spaces for$1\leq p\leq\infty$; (β) Suppose that$p_{t,L}(x,y)$is the kernel corresponding to the semigroup generated by one of the operators L listed in(α). Then the kernel $p_{t,L}(x,y)$ satisfies Gaussian bounds:

$$\left|p_{t,L}(x,y)\right|\leq\frac{C}{t^{\frac{n}{2}}}e^{-c\frac{|x-y|^2}{t}},\tag{65}$$

for all$x,y\in\Omega$, where$\Omega=\mathbb{R}^n$forΔ, Δ_N; $\Omega=\mathbb{R}^n_+$for$\Delta_{N_+}$and $\Omega=\mathbb{R}^n_-$forΔ_{N_-}.

Next we consider the smoothness property of the heat kernel forΔ,Δ_{N_+}, andΔ_{N_-}.

Proposition (6.2.3)[178]: Suppose that L is one of the operatorsΔ_{N_+}, Δ_{N_-}andΔ_N. Then for $x,x',y\in\mathbb{R}^n_+$(or$\in\mathbb{R}^n_-$) with$|x-x'|\leq\frac{1}{2}|x-y|$, we have

$$\left|p_{t,L}(x,y)-p_{t,L}(x',y)\right|\leq C\frac{|x-x'|}{\left(\sqrt{t}+|x-y|\right)}\frac{\sqrt{t}}{\left(\sqrt{t}+|x-y|\right)^{n+1}};\tag{66}$$

Symmetrically, for $x,y,y'\in\mathbb{R}^n_+$(or$\in\mathbb{R}^n_-$) with$|y-y'|\leq\frac{1}{2}|x-y|$,we have

$$\left|p_{t,L}(x,y)-p_{t,L}(x,y')\right|\leq C\frac{|y-y'|}{\left(\sqrt{t}+|x-y|\right)}\frac{\sqrt{t}}{\left(\sqrt{t}+|x-y|\right)^{n+1}}.\tag{67}$$

Proof: Suppose$x,y\in\mathbb{R}^n_+$. Then for$i=1,\dots,n-1$, we have

$$\frac{\partial}{\partial x_i}p_{t,\Delta_{N_+}}(x,y)=-\frac{(x_i-y_i)}{2t}\frac{1}{(4\pi t)^{\frac{n}{2}}}e^{-\frac{|x'-y'|^2}{4t}}\left(e^{-\frac{|x_n-y_n|^2}{4t}}+e^{-\frac{|x_n-y_n|^2}{4t}}\right).$$

Moreover,

$$\frac{\partial}{\partial x_n}p_{t,\Delta_{N_+}}(x,y)=-\frac{1}{(4\pi t)^{\frac{n}{2}}}e^{-\frac{|x'-y'|^2}{4t}}\left(e^{-\frac{|x_n-y_n|^2}{4t}}\frac{(x_n-y_n)}{2t}+e^{-\frac{|x_n-y_n|^2}{4t}}\frac{(x_n-y_n)}{2t}\right).$$

Then we obtain that

$$\begin{aligned}&\left|\nabla_x p_{t,\Delta_{N_+}}(x,y)\right|^2\\&=\sum_{i=1}^{n-1}\left|\frac{\partial}{\partial x_i}p_{t,\Delta_{N_+}}(x,y)\right|^2+\left|\frac{\partial}{\partial x_n}p_{t,\Delta_{N_+}}(x,y)\right|^2\\&\leq\sum_{i=1}^{n-1}\frac{(x_i-y_i)^2}{4t^2}\frac{1}{(4\pi t)^n}e^{-\frac{|x'-y'|^2}{2t}}\left(e^{-\frac{|x_n-y_n|^2}{4t}}+e^{-\frac{|x_n-y_n|^2}{4t}}\right)^2\\&+2\frac{1}{(4\pi t)^n}e^{-\frac{|x'-y'|^2}{2t}}\left(e^{-\frac{|x_n-y_n|^2}{4t}}\frac{(x_n-y_n)}{2t}\right)^2\\&+2\frac{1}{(4\pi t)^n}e^{-\frac{|x'-y'|^2}{2t}}\left(e^{-\frac{|x_n-y_n|^2}{4t}}\frac{(x_n-y_n)}{2t}\right)^2\end{aligned}$$

$$\leq C\sum_{i=1}^{n}\frac{(x_i-y_i)^2}{t^2}\frac{1}{(4\pi t)^n}e^{-\frac{|x-y|^2}{2t}}+2\frac{1}{(4\pi t)^n}e^{-\frac{|x'-y'|^2}{2t}}\left(e^{-\frac{|x_n-y_n|^2}{4t}}\frac{(x_n-y_n)}{2t}\right)^2$$

$$\leq C\frac{|x-y|^2}{t^2}\frac{1}{(4\pi t)^n}e^{-\frac{|x-y|^2}{2t}}+2\frac{1}{(4\pi t)^n}e^{-\frac{|x-y|^2}{4t}}\left(e^{-\frac{|x_n-y_n|^2}{8t}}\frac{(x_n-y_n)}{2t}\right)^2$$

$$\leq C\frac{t}{(t+|x-y|^2)^{n+2}}.$$

Hence, it is easy to verify that

$$\left|\nabla_x p_{t,\Delta_{N_+}}(x,y)\right|\leq C\frac{\sqrt{t}}{\left(\sqrt{t}+|x-y|\right)^{n+2}}$$

and similarly we can obtain that

$$\left|\nabla_x p_{t,\Delta_{N_+}}(x,y)\right|\leq C\frac{\sqrt{t}}{\left(\sqrt{t}+|x-y|\right)^{n+2}},$$

which implies that

$$\left|p_{t,\Delta_{N_+}}(x,y)-p_{t,\Delta_{N_+}}(x',y)\right|\leq C\frac{|x'-x|}{\left(\sqrt{t}+|x-y|\right)}\frac{\sqrt{t}}{\left(\sqrt{t}+|x-y|\right)^{n+1}}$$

for $x,x',y\in\mathbb{R}_+^n$ with $|x-x'|\leq\frac{1}{2}|x-y|$, and

$$\left|p_{t,\Delta_{N_+}}(x,y)-p_{t,\Delta_{N_+}}(x,y')\right|\leq C\frac{|y-y'|}{\left(\sqrt{t}+|x-y|\right)}\frac{\sqrt{t}}{\left(\sqrt{t}+|x-y|\right)^{n+1}}$$

for $x,x',y\in\mathbb{R}_+^n$ with $|y-y'|\leq\frac{1}{2}|x-y|$.

We can obtain similar estimates for the heat semigroup of Δ_{N_-} and Δ_N.

A fundamental object in our study are the Riesz transforms associated to the Neumann Laplacian. Recall that the Riesz transforms associated to the Neumann Laplacian are given by: $R_N=\nabla\Delta_N^{-\frac{1}{2}}$. We collect the formula for these kernels in the following proposition.

Proposition (6.2.4)[178]: Denote by $R_{N,j}(x,y)$ the kernel of the j-th Riesz transform $\frac{\partial}{\partial_{x_j}}\Delta_N^{-\frac{1}{2}}$ of Δ_N. Then for $1\leq j\leq n-1$ and for $x,y\in\mathbb{R}_+^n$ we have:

$$R_{N,j}(x,y)=-C_n\left(\frac{x_i-y_i}{|x-y|^{n+1}}+\frac{x_i-y_i}{(|x'-y'|^2+|x_n-y_n|^2)^{\frac{n+1}{2}}}\right)$$

And

$$R_{N,n}(x,y)=-C_n\left(\frac{x_i-y_i}{|x-y|^{n+1}}+\frac{x_n+y_n}{(|x'-y'|^2+|x_n+y_n|^2)^{\frac{n+1}{2}}}\right),$$

where $C_n=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\pi^{\frac{n+1}{2}}}$. Similar expressions also hold for $R_{N,j}(x,y), j=1,\dots,n$, when $x,y\in\mathbb{R}_-^n$.

Proof: Working from the definition of the square root of Δ_N, i.e.,

$$\Delta_N^{-\frac{1}{2}} = \frac{1}{\Gamma\left(\frac{1}{2}\right)} \int_0^\infty e^{-t\Delta_N} \frac{dt}{\sqrt{t}},$$

we have that for $1 \le j \le n-1$:

$$R_{N,j}(x,y) = \frac{1}{\Gamma\left(\frac{1}{2}\right)} \frac{\partial}{\partial_{x_j}} \int_0^\infty p_{t,\Delta_N}(x,y) \frac{dt}{\sqrt{t}}$$

$$= \frac{1}{\Gamma\left(\frac{1}{2}\right)} \frac{\partial}{\partial_{x_j}} \left(\int_0^\infty \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x-y|^2}{4t}} \frac{dt}{\sqrt{t}} + \int_0^\infty \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x'-y'|^2}{4t}} e^{-\frac{|x_n-y_n|^2}{4t}} \frac{dt}{\sqrt{t}} \right)$$

$$= -\frac{\Gamma\left(\frac{n+1}{2}\right)}{(\pi)^{\frac{n+1}{2}}} \left(\frac{x_i - y_i}{|x-y|^{n+1}} + \frac{x_i - y_i}{(|x'-y'|^2 + |x_n + y_n|^2)^{\frac{n+1}{2}}} \right).$$

For $j = n$ and for $x, y \in \mathbb{R}_+^n$ we again observe:

$$R_{N,n}(x,y) = \frac{\sqrt{\pi}}{2} \frac{\partial}{\partial_{x_n}} \int_0^\infty p_{t,\Delta_N}(x,y) \frac{dt}{\sqrt{t}}$$

$$= -\frac{\Gamma\left(\frac{n+1}{2}\right)}{(\pi)^{\frac{n+1}{2}}} \left(\frac{x_n - y_n}{|x-y|^{n+1}} + \frac{x_n - y_n}{(|x'-y'|^2 + |x_n + y_n|^2)^{\frac{n+1}{2}}} \right).$$

We next make the observation that kernels $R_{N,j}(x,y)$ are Calderón–Zygmund kernels.

Proposition (6.2.5)[178]: Denote by $R_N(x,y)$ the kernel of the vector of Riesz transforms $\nabla\Delta_N^{-\frac{1}{2}}$. Then:

$$R_N(x,y) = \left(R_{N,1}(x,y), \dots, R_{N,n}(x,y)\right) H(x_n, y_n), \tag{68}$$

with $H(t)$ the Heaviside function defined in (64). Moreover, we have that

$$|R_N(x,y)| \le C_n \frac{1}{|x-y|^2},$$

And

$$|R_N(x,y) - R_N(x_0,y)| + |R_N(x,y) - R_N(y,x_0)| \le C \frac{|x-x_0|}{|x-y|^{n+1}}$$

for $x, x_0 \in \mathbb{R}_+^n$ (or $x, x_0, y \in \mathbb{R}_-^n$) with $|x - x_0| \le \frac{1}{2}|x-y|$.

Proof: We first claim that for $j = 1, \dots, n$, and $x, y \in \mathbb{R}_+^n$ (or $x, y \in \mathbb{R}_-^n$)

$$\left|R_{N,j}(x,y)\right| \le C_n \frac{1}{|x-y|^n}.$$

In fact, from Proposition (6.2.4), it is direct that for $1 \le j \le n-1$,

$$\frac{|x_j - y_j|}{(|x'-y'|^2 + |x_n + y_n|^2)^{\frac{n+1}{2}}} \le \frac{|x_j - y_j|}{(|x'-y'|^2 + |x_n + y_n|^2)^{\frac{n+1}{2}}} \le \frac{1}{|x-y|^n}$$

and for $j = n$,

$$\frac{|x_n - y_n|}{(|x'-y'|^2 + |x_n + y_n|^2)^{\frac{n+1}{2}}} \le \frac{|x_j - y_j|}{(|x'-y'|^2 + |x_n + y_n|^2)^{\frac{n+1}{2}}} \le \frac{1}{|x-y|^n},$$

where we use the fact that $x, x_0, y \in \mathbb{R}_+^n$ (or $x, x_0, y \in \mathbb{R}_-^n$)and hence$x_j - y_j > |x_j - y_j|for1 \leq j \leq n$.

Similarly, by considering the estimates for the terms$\frac{\partial}{\partial_{x_j}} R_{N,j}(x,y)$and $\frac{\partial}{\partial_{x_j}} R_{N,j}(x,y)$, we obtain that

$$|R_{N,j}(x,y) - R_{N,j}(x_0,y)| + |R_{N,j}(x,y) - R_{N,j}(y,x_0)| \leq C \frac{|x-x_0|}{|x-y|^{n+1}}$$

for $x, x_0, y \in \mathbb{R}_+^n$ (or $x, x_0, y \in \mathbb{R}_-^n$)with$|x - x_0| \leq \frac{1}{2}|x-y|$.

For $0 < \alpha < n$, denote by$K(x,y)$ the kernel of the classical fractional operator $\Delta^{-\alpha/2}$, which is defined by

$$\Delta^{-\alpha/2} f(x) = \frac{1}{\Gamma(\alpha/2)} \int_0^\infty e^{-t\Delta} f(x) \frac{dt}{t^{1-\alpha/2}}.$$

We know that

$$K(x,y) = \frac{C_{n,\alpha}}{|x-y|^{n-\alpha}},$$

Where$C_{n,\alpha} = \frac{\Gamma\left(\frac{n}{2}-\frac{\alpha}{2}\right)}{\Gamma\left(\frac{\alpha}{2}\right)} \frac{1}{\pi^{\frac{n}{2}} 2^{\alpha}}$. It is well known that when$b \in BMO(\mathbb{R}^n)$, the commutator$[b, \Delta^{-\alpha/2}]$ is bounded from $L^p(\mathbb{R}^n)$to$L^q(\mathbb{R}^n)$for$1 < p < n/\alpha$ and$1/q = 1/p - \alpha/n$. See[179].

Proposition (6.2.6)[178]: Denote by $K_N(x,y)$the kernel of the fractional operator$\Delta_N^{-\alpha/2}$. The $x, y \in \mathbb{R}_+^n$we have:

$$K_N(x,y) = K(x,y) + \widetilde{K}_N(x,y)$$

with

$$\widetilde{K}_N(x,y) := C_{n,\alpha} = \frac{1}{(|x'-y'|^2 + |x_n + y_n|^2)^{\frac{n}{2}-\frac{\alpha}{2}}}.$$

Similar expressions for $K_N(x,y)$ when $x, y \in \mathbb{R}_-^n$ also hold.

Proof: For $x, y \in \mathbb{R}_+^n$, working from the fraction of the square root of $\Delta_N$we have that:

$$K_N(x,y) = \frac{1}{\Gamma(\alpha/2)} \int_0^\infty p_{t,\Delta_N}(x,y) \frac{dt}{t^{1-\alpha/2}}$$

$$= \frac{1}{\Gamma(\alpha/2)} \int_0^\infty \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x-y|^2}{4t}} \frac{dt}{t^{1-\alpha/2}} + \frac{1}{\Gamma(\alpha/2)} \int_0^\infty \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x'-y'|^2}{4t}} e^{-\frac{|x_n-y_n|^2}{4t}} \frac{dt}{t^{1-\alpha/2}}$$

$$= C_{n,\alpha}\left(\frac{1}{|x-y|^{n-\alpha}} + \frac{1}{(|x'-y'|^2 + |x_n+y_n|^2)^{\frac{n}{2}-\frac{\alpha}{2}}}\right)$$

$$= K(x,y) + \widetilde{K}_N(x,y),$$

where we set

$$\widetilde{K}_N(x,y) = C_{n,\alpha} \frac{1}{(|x'-y'|^2 + |x_n+y_n|^2)^{\frac{n}{2}-\frac{\alpha}{2}}}.$$

We now recall the definition and some fundamental properties of$BMO_{\Delta_N}(\mathbb{R}^n)$ from[5].

Define

$$\mathcal{M} = \left\{ f \in L^1_{loc}(\mathbb{R}^n) : \exists d > 0 s.t. \int_{\mathbb{R}^n} \frac{|f(x)|^2}{1+|n|^{n+d}} dx < \infty \right\}.$$

Definition (6.2.7)[178]: ([65]). We say that$f \in \mathcal{M}$ is of bounded mean oscillation associated withΔ_N, abbreviated as$BMO_{\Delta_N}(\mathbb{R}^n)$, if

$$\|f\|_{BMO_{\Delta_N}(\mathbb{R}^n)} = \sup_{B(y,r)} \frac{1}{|B(y,r)|} \int_{B(y,r)} |f(x) - \exp(-r^2\Delta_N) f(x)| \, dx <, \qquad (69)$$

where the supremum is taken over all balls$B(y,r)$ in$\mathbb{R}^n$. The smallest bound for which(69) is satisfied is then taken to be the norm of f in this space, and is denoted by$\|f\|_{BMO_{\Delta_N}(\mathbb{R}^n)}$.

Definition (6.2.8)[178]: ([65]). A function$f$ on$\mathbb{R}^n_+$ said to be in$BMO_r(\mathbb{R}^n_+)$if there exists $F \in BMO(\mathbb{R}^n)$such that$F|_{\mathbb{R}^n_+} = f$. If$f \in BMO_r(\mathbb{R}^n_+)$, then we set

$$\|f\|_{BMO_r(\mathbb{R}^n_+)} = \inf\left\{ \|F\|_{BMO_r(\mathbb{R}^n_+)} : F|_{\mathbb{R}^n_+} = f \right\}.$$

Definition (6.2.9)[178]: ([65]). For any function$f \in L^1_{loc}(\mathbb{R}^n_+)$, define

$$\|f\|_{BMO_e(\mathbb{R}^n_+)} = \|f_e\|_{BMO(\mathbb{R}^n_+)},$$

where$f_e$is defined in(60). We denote by$BMO_e(\mathbb{R}^n_+)$the corresponding Banach space. Similarly we can define the spaces$BMO_r(\mathbb{R}^n_-)$and$BMO_e(\mathbb{R}^n_-)$.

Proposition (6.2.10)[178]: ([65]). The spaces$BMO_r(\mathbb{R}^n_+)$and$BMO_e(\mathbb{R}^n_+)$ coincide, and their norms are equivalent. Similar result holds for$BMO_r(\mathbb{R}^n_-)$and $BMO_e(\mathbb{R}^n_+)$.

Proposition (6.2.11)[178]: ([65]). The Neumann BMO space$BMO_{\Delta_N}(\mathbb{R}^n)$ can be described in the following way:

$$BMO_{\Delta_N}(\mathbb{R}^n) = \{ f \in \mathcal{M} : f_+ \in BMO_r(\mathbb{R}^n_+) and f_- \in BMO_r(\mathbb{R}^n_-) \}.$$

As a consequence of the results from[65]listed above, we obtain that$f \in BMO_{\Delta_N}(\mathbb{R}^n)$if and only if$f_{+,e}, f_{-,e} \in BMO(\mathbb{R}^n)$. A final key fact that plays a role in our analysis is the duality between$BMO_{\Delta_N}(\mathbb{R}^n)$ and $H^1_{\Delta_N}(\mathbb{R}^n)$.

Proposition (6.2.12)[178]: ([65]). The dual space of$H^1_{\Delta_N}(\mathbb{R}^n)isBMO_{\Delta_N}(\mathbb{R}^n)$.

We provide a deeper study of the space$H^1(\mathbb{R}^n)$. We first provide several equivalent characterizations of$H^1_{\Delta_N}(\mathbb{R}^n)$. To do so, we need the following definitions of the Hardy space associated toΔ_N in terms of the radial maximal function, the non-tangential maximal function, the Riesz transforms, and atoms. As one might expect, these definitions all turn out to be equivalent as shown below in Theorem(6.2.18).

Definition (6.2.13)[178]: We define$H^1_{\Delta_N,max}(\mathbb{R}^n) = \left\{ f L^1(\mathbb{R}^n) : f^+_{\Delta_N} \in L^1(\mathbb{R}^n) \right\}$with the norm $\|f\|_{H^1_{\Delta_N,max}(\mathbb{R}^n)} = \left\| f^+_{\Delta_N}(\mathbb{R}^n) \right\|_{L^1(\mathbb{R}^n)}$, where$f^+_{\Delta_N}(x) = \sup_{t>0} |\exp(-t^2\Delta_N) f(x)|$.

Definition (6.2.14)[178]: We define$H^1_{\Delta_N,*}(\mathbb{R}^n) = \left\{ f L^1(\mathbb{R}^n) : f^*_{\Delta_N} \in L^1(\mathbb{R}^n) \right\}$with the norm $\|f\|_{H^1_{\Delta_N,*}(\mathbb{R}^n)} = \left\| f^+_{\Delta_N}(\mathbb{R}^n) \right\|_{L^1(\mathbb{R}^n)}$, where$f^*_{\Delta_N}(x) = \sup_{t>0} |\exp(-t^2\Delta_N) f(x)|$.

Definition (6.2.15)[178]: We define

$$H^1_{\Delta_N,Riesz}(\mathbb{R}^n) = \left\{ f L^1(\mathbb{R}^n) : \frac{\partial}{\partial_{x_l}} \Delta^+_N(\mathbb{R}^n) f \in L^1(\mathbb{R}^n) \text{for } 1 \le l \le n \right\}.$$

with the norm$H^1_{\Delta_N,Riesz}(\mathbb{R}^n) = \|f\|_{L^1(\mathbb{R}^n)} + \sum_{i=1}^n \left\| \frac{\partial}{\partial_{x_l}} \Delta^+_N(\mathbb{R}^n) \right\|_{L^1(\mathbb{R}^n)}$

Next we define the atoms for $H^1_{\Delta_N,max}(\mathbb{R}^n)$, which we adapt from a very recent result of Song and Yan[158].

Definition (6.2.16)[178]: Given $M \in \mathbb{N}$. We say that a function $a(x) \in L^\infty(\mathbb{R}^n)$ is an $H^1_{\Delta_N,max}(\mathbb{R}^n)$-atom, if there exist a function b in the domain of Δ_N^M and a ball $B \subset \mathbb{R}^n$ such that

(i) $a = \Delta_N^M b$;

(ii) $\operatorname{supp} \Delta_N^k b \subset B, k = 0,1,\dots,M$;

(iii) $\left\|(r_B^2\Delta_N)^k b\right\|_{L^\infty(\mathbb{R}^n)} r_B^{2M}|B|^{-1}, k = 0,1,\dots,M$;

Definition (6.2.17)[178]: We say that $f = \sum_j \lambda_j a_j$ is an atomic representation of f if $\{\lambda_j\} \in \ell^1$, each a_j is an $H^1_{\Delta_N,max}(\mathbb{R}^n)$ atom, and the sum converges in $L^2(\mathbb{R}^n)$. Set

$$\widetilde{H}^1_{\Delta_N,atom}(\mathbb{R}^n) = \{f \in L^2(\mathbb{R}^n): f \text{ has an atomic representation }\}$$

with the norm $\|f\|_{\widetilde{H}^1_{\Delta_N,atom}(\mathbb{R}^n)}$ given by

$$\inf\left\{\sum_j |\lambda_j| : f = \sum_j \lambda_j a_j \text{ is an atomic representation }\right\}.$$

The space $H^1_{\Delta_N,atom}(\mathbb{R}^n)$ is defined as the completion of $\widetilde{H}^1_{\Delta_N,atom}(\mathbb{R}^n)$ with respect to this norm.

We now collection the equivalence of all these definitions and moreover provide a link between $H^1(\mathbb{R}^n)$ and $H^1_{\Delta_N}(\mathbb{R}^n)$.

Theorem (6.2.18)[178]: Let all the notation be as above. Then,

$$H^1_{\Delta_N}(\mathbb{R}^n) = H^1_{\Delta_N,max}(\mathbb{R}^n) = H^1_{\Delta_N,*}(\mathbb{R}^n) = H^1_{\Delta_N,Riesz}(\mathbb{R}^n) = H^1_{\Delta_N,atom}(\mathbb{R}^n)$$

and they have equivalent norms

$$\|f\|_{H^1_{\Delta_N}(\mathbb{R}^n)} \approx \|f\|_{H^1_{\Delta_N,max}(\mathbb{R}^n)} \|f\|_{H^1_{\Delta_N,Riesz}(\mathbb{R}^n)} \|f\|_{H^1_{\Delta_N,atom}(\mathbb{R}^n)} \approx \left\|f_{+,e}\right\|_{H^1(\mathbb{R}^n)} + \left\|f_{-,e}\right\|_{H^1(\mathbb{R}^n)}.$$

Namely, $f \in H^1_{\Delta_N}(\mathbb{R}^n)$ if and only if $f_{+,e} \in H^1(\mathbb{R}^n)$ and $f_{-,e} \in H^1(\mathbb{R}^n)$.

Proof: We recall that the Hardy space associated with Δ_N is defined as the set of functions $\left\{f \in L^1(\mathbb{R}^n): \left\|S_{\Delta_N}(f)\right\|_{L^1(\mathbb{R}^n)} < \infty\right\}$ in the norm of $\|f\|_{H^1_{\Delta_N}(\mathbb{R}^n)} = \left\|S_{\Delta_N}(f)\right\|_{L^1(\mathbb{R}^n)}$ where $S_{\Delta_N}(f)(x) = \left(\int_0^\infty \int_{|y-x|<t} |Q_{t^2} f(y)|^2 \frac{dydt}{t^{n+1}}\right)^{\frac{1}{2}}$, and $Q_{t^2} = t^2\Delta_N \exp(-t^2\Delta_N)$.

We now consider the operator $Q_t = t\Delta_N \exp(-t\Delta_N) = -t\frac{d}{dt}\exp(-t\Delta_N)$ for any $t > 0$ (see [71]). Then we have

$$Q_{t^2} f(x) = t^2\Delta_N \exp(-t^2\Delta_N) f(x) = \int_{\mathbb{R}^n} -\frac{t}{2}\frac{\partial}{\partial t} p_{t^2,\Delta_N}(x,y) f(y)dy.$$

From the definition of $p_{t,\Delta_N}(x,y)$, see (63), we have that for any $x \in \mathbb{R}^n_+$,

$$t^2\Delta_N \exp(-t^2\Delta_N) f(x) = \int_{\mathbb{R}^n_+} -\frac{t}{2}\frac{\partial}{\partial t} p_{t^2,\Delta_N}(x,y) f_+(y)dy$$

$$= \int_{\mathbb{R}^n} -\frac{t}{2}\frac{\partial}{\partial t} p_{t^2}(x,y) f_{+,e}(y)dy$$

$$= t^2\Delta \exp(-t^2\Delta_N) f_{+,e}(x).$$

Similarly, for any $x \in \mathbb{R}^n_-$, we have $t^2\Delta_N \exp(-t^2\Delta_N) f(x) = t^2\Delta \exp(-t^2\Delta) f_{-,e}(x)$. Moreover, by a change of variable,

$$t^2\Delta_N \exp(-t^2\Delta_N) f(x) = -t^2\Delta \exp(-t^2\Delta) f_{+,e}(\tilde{x}) \text{for any } t > 0, x \in \mathbb{R}^n_+; \qquad (70)$$

$$t^2\Delta_N \exp(-t^2\Delta_N) f(x) = -t^2\Delta \exp(-t^2\Delta) f_{-,e}(\tilde{x}) \text{for any } t > 0, x \in \mathbb{R}^n_-$$

Then from(70) we have

$$S_{\Delta_N}(f)(x)^2 = \int_0^\infty \int_{|x-y|<t, y\in\mathbb{R}^n_+} |t^2\Delta_N \exp(-t^2\Delta_N) f(x)|^2 \frac{dydt}{t^n}$$
$$+ \int_0^\infty \int_{|x-y|<t, y\in\mathbb{R}^n_+} |t^2\Delta_N \exp(-t^2\Delta_N) f(x)|^2 \frac{dydt}{t^n}$$
$$= \int_0^\infty \int_{|x-y|<t, y\in\mathbb{R}^n_+} \left|t^2\Delta \exp(-t^2\Delta) f_{+,e}(y)\right|^2 \frac{dydt}{t^n}$$
$$+ \int_0^\infty \int_{|x-y|<t, y\in\mathbb{R}^n_-} \left|t^2\Delta \exp(-t^2\Delta) f_{-,e}(y)\right|^2 \frac{dydt}{t^n}$$
$$= \frac{1}{2}\left(\int_0^\infty \int_{|x-y|<t} \left|t^2\Delta \exp(-t^2\Delta) f_{+,e}(y)\right|^2 \frac{dydt}{t^n}\right.$$
$$\left. + \int_0^\infty \int_{|x-y|<t} \left|t^2\Delta \exp(-t^2\Delta) f_{-,e}(y)\right|^2 \frac{dydt}{t^n}\right),$$

which implies that $S_{\Delta_N}(f)(x) \le \frac{\sqrt{2}}{2}\left(S(f_{+,e})(x) + S(f_{-,e})(y)\right)$. Conversely,

$$S(f_{+,e})(x)^2 = \int_0^\infty \int_{|x-y|<t} \left|t^2\Delta \exp(-t^2\Delta) f_{+,e}(y)\right|^2 \frac{dydt}{t^n}$$
$$= 2\int_0^\infty \int_{|x-y|<t, y\in\mathbb{R}^n_+} \left|t^2\Delta \exp(-t^2\Delta) f_{+,e}(y)\right|^2 \frac{dydt}{t^n}$$
$$\le 2S_{\Delta_N}(f)(x)^2.$$

Similarly we have $S(f_{-,e})(x)^2 \le 2S_{\Delta_N}(f)(x)^2$. Hence, we obtain that $S(f_{+,e})(x) + S(f_{-,e})(x) \le 2\sqrt{2}S_{\Delta_N}(f)(x)$. As a consequence, we have

$$\|f\|_{H^1_{\Delta_N}(\mathbb{R}^n)} = \int_{\mathbb{R}^n} \left|S_{\Delta_N}(f)(x)\right| dx \approx \int_{\mathbb{R}^n} \left|S(f_{+,e})(x)\right| dx + \int_{\mathbb{R}^n} \left|S(f_{-,e})(x)\right| dx$$
$$= \left\|f_{+,e}\right\|_{H^1(\mathbb{R}^n)} + \left\|f_{-,e}\right\|_{H^1(\mathbb{R}^n)}. \qquad (71)$$

Next we turn to $H^1_{\Delta_D,max}(\mathbb{R}^n)$. From(63) we can see that for any $t \ge 0$ and $x \in \mathbb{R}^n_+$,

$$\exp(-t^2\Delta_N) f(x) = \int_{\mathbb{R}^n} p_{t^2,\Delta_N}(x,y) f(y) dy = \int_{\mathbb{R}^n_+} p_{t^2,\Delta_N}(x,y) f_+(y) dy.$$

$$= \int_{\mathbb{R}^n} p_{t^2}(x,y) f_{+,e}(y)dy = \exp(-t^2\Delta) f_{+,e}(x).$$

Similarly, $\exp(-t^2\Delta_N) f(x) = \exp(-t^2\Delta) f_{+,e}(x)$ for any $t \geq 0$ and $x \in \mathbb{R}^n_-$. Thus,

$$\sup_{t>0} |\exp(-t^2\Delta_N) f(x)| = \sup_{t>0} \left|\exp(-t^2\Delta) f_{+,e}(x)\right| \text{ for any } x \in \mathbb{R}^n_+; \tag{72}$$

$$\sup_{t>0} |\exp(-t^2\Delta_N) f(x)| = \sup_{t>0} \left|\exp(-t^2\Delta) f_{-,e}(x)\right| \text{ for any } x \in \mathbb{R}^n_-.$$

Again, by a change of variable, we have that

$$\exp(-t^2\Delta_N) f(x) = -\exp(-t^2\Delta) f_{+,e}(\tilde{x}) \text{ for any } t > 0, x \in \mathbb{R}^n_+; \tag{73}$$

$$\exp(-t^2\Delta_N) f(x) = -\exp(-t^2\Delta) f_{-,e}(\tilde{x}) \text{ for any } t > 0, x \in \mathbb{R}^n_-.$$

Then, for any $f \in H^1_{\Delta_N,max}(\mathbb{R}^n)$, from(72) and(73) we can obtain that

$$\|f\|_{H^1_{\Delta_N,max}(\mathbb{R}^n)} = \int_{\mathbb{R}^n_+} \left|f^+_{\Delta_N}(x)\right| dx + \int_{\mathbb{R}^n_-} \left|f^+_{\Delta_N}(x)\right| dx \tag{74}$$

$$= \int_{\mathbb{R}^n_+} \sup_{t>0} |\exp(-t^2\Delta_N) f(x)|\, dx + \int_{\mathbb{R}^n_-} \sup_{t>0} |\exp(-t^2\Delta_N) f(x)|\, dx$$

$$= \int_{\mathbb{R}^n_+} \sup_{t>0} \left|\exp(-t^2\Delta) f_{+,e}(x)\right| dx + \int_{\mathbb{R}^n_-} \sup_{t>0} \left|\exp(-t^2\Delta) f_{-,e}(x)\right| dx$$

$$= \frac{1}{2}\left(\int_{\mathbb{R}^n} \sup_{t>0} \left|\exp(-t^2\Delta) f_{+,e}(x)\right| dx + \int_{\mathbb{R}^n} \sup_{t>0} \left|\exp(-t^2\Delta) f_{-,e}(x)\right| dx \right)$$

$$= \frac{1}{2}\left(\left\|\left(f_{+,e}\right)^+\right\|_{L^1(\mathbb{R}^n)} + \left\|\left(f_{-,e}\right)^+\right\|_{L^1(\mathbb{R}^n)} \right)$$

$$= \frac{1}{2}\left(\left\|f_{+,e}\right\|_{H^1(\mathbb{R}^n)} + \left\|f_{-,e}\right\|_{H^1(\mathbb{R}^n)} \right),$$

where $f^+(x) = \sup_{t>0} |p_{t^2} * f(x)|$ is the classical maximal function as defined in (74). Thus(74) yields that $f \in H^1_{\Delta_N,max}(\mathbb{R}^n)$ if and only if $f_{+,e} \in H^1(\mathbb{R}^n)$ and $f_{-,e} \in H^1(\mathbb{R}^n)$. We now consider the Hardy space $H^1_{\Delta_N,*}(\mathbb{R}^n)$ via the non-tangential maximal function. Note that

$$f^*_{\Delta_N}(x) = \sup_{|x-y|<t} |\exp(-t^2\Delta_N) f(y)|$$

$$\leq \sup_{|x-y|<t, y\in\mathbb{R}^n_+} |\exp(-t^2\Delta_N) f(y)| + \sup_{|x-y|<t, y\in\mathbb{R}^n_-} |\exp(-t^2\Delta_N) f(y)|$$

$$\leq \sup_{|x-y|<t, y\in\mathbb{R}^n_+} \left|\exp(-t^2\Delta) f_{+,e}(y)\right| + \sup_{|x-y|<t, y\in\mathbb{R}^n_-} \left|\exp(-t^2\Delta) f_{-,e}(y)\right|$$

$$\leq \sup_{|x-y|<t} \left|\exp(-t^2\Delta) f_{+,e}(y)\right| + \sup_{|x-y|<t} \left|\exp(-t^2\Delta) f_{-,e}(y)\right|$$

$$= \left(f_{+,e}\right)^*(x) + \left(f_{-,e}\right)^*(x),$$

where $f^*(x) = \sup_{|x-y|<t} |p_{t^2} * f(y)|$ is the classical non-tangential maximal function. Hence $\left\|f^*_{\Delta_N}(x)\right\|_{L^1(\mathbb{R}^n)} \leq \left\|\left(f_{+,e}\right)^*\right\|_{L^1(\mathbb{R}^n)} + \left\|\left(f_{-,e}\right)^*\right\|_{L^1(\mathbb{R}^n)}$. Moreover, we have

$$\left(f_{+,e}\right)^*(x) = \sup_{|x-y|<t} \left|\exp(-t^2\Delta) f_{+,e}(y)\right|$$

$$\leq \sup_{|x-y|<t, y\in\mathbb{R}^n_+} \left|\exp(-t^2\Delta) f_{+,e}(y)\right| + \sup_{|x-y|<t, y\in\mathbb{R}^n_-} \left|\exp(-t^2\Delta) f_{-,e}(y)\right|$$

$$\leq 2 \sup_{|x-y|<t} |\exp(-t^2\Delta_N) f(y)|$$
$$\leq 2f^*_{\Delta_N}(x)$$
Thus, $\|(f_{+,e})^*\|_{L^1(\mathbb{R}^n)} \leq 2\|f^*_{\Delta_N}(x)\|_{L^1(\mathbb{R}^n)}$. Similarly we obtain $\|(f_{-,e})^*\|_1 \leq 2\|f^*_{\Delta_N}(x)\|_{L^1(\mathbb{R}^n)}$. This implies that
$$\|f^*_{\Delta_N}(x)\|_1 \approx \|(f_{+,e})^*\|_{L^1(\mathbb{R}^n)} + \|(f_{-,e})^*\|_{L^1(\mathbb{R}^n)}. \quad (75)$$
Thus,(75) yields that $f \in H^1_{\Delta_N,*}(\mathbb{R}^n)$if and only if $f_{+,e} \in H^1(\mathbb{R}^n)$and $f_{-,e} \in H^1(\mathbb{R}^n)$.As for the Riesz transform characterization of the Hardy space$H^1_{\Delta_N}(\mathbb{R}^n)$, it suffices to note that when $x \in \mathbb{R}^n_+$,
$$\nabla\Delta_N^{-\frac{1}{2}} f(x) = \int_{\mathbb{R}^n} K_N(x,y)f(y)dy = \int_{\mathbb{R}^n_+} R_N(x,y)f_+(y)dy = \int_{\mathbb{R}^n} R(x,y)f_{+,e}(y)dy$$
$$= \nabla\Delta_N^{-\frac{1}{2}} f_{+,e}(x)$$
and that when $x \in \mathbb{R}^n_-$,
$$\nabla\Delta_N^{-\frac{1}{2}} f(x) = \nabla\Delta_N^{-\frac{1}{2}} f_{-,e}(x).$$
Thus, $f \in H^1_{\Delta_N,Riesz}(\mathbb{R}^n)$if and only if $f_{+,e} \in H^1(\mathbb{R}^n)andf_{-,e} \in H^1(\mathbb{R}^n)$.
Finally, for the atomic decomposition, in Song and Yan[158], they already obtained that$H^1_{\Delta_N,*}(\mathbb{R}^n) = H^1_{\Delta_N,atom}(\mathbb{R}^n)$. See [158] for this fact.
We now prove the Fefferman–Stein type representation for the space $BMO_{\Delta_N}(\mathbb{R}^n)$.

Corollary (6.2.19)[178]: The following are equivalent for a function$b$:

(i) $b \in BMO_{\Delta_N}(\mathbb{R}^n)$;

(ii) There exists $b_0, b_1, \dots, b_n \in L^\infty(\mathbb{R}^n)$such that $b = b_0 + \sum_{j=1}^n R^*_{N,j} b_j$, where $R^*_{N,j}$is the adjoint operator of $R_{N,j}$.

Proof. The proof is as in[74]. Let $B = \oplus_{j=0}^n L^1(\mathbb{R}^n)$and norm B by $\sum_{j=0}^n \|f\|_{L^1(\mathbb{R}^n)}$. We have that $B^* = \oplus_{j=0}^n L^\infty(\mathbb{R}^n)$. Let S be the subspace of B given by
$$S = \{(f, R_{N,1}f, \dots, R_{N,n}f) : f \in L^1(\mathbb{R}^n)\}.$$
We have that S is a closed subspace and that$f \to (f, R_{N,1}f, \dots, R_{N,n}f)$is a isometry of $H^1_{\Delta_N}(\mathbb{R}^n)$ to S. Linear functionals on S and $H^1_{\Delta_N}(\mathbb{R}^n)$can be identified in an obvious way, hence any continuous linear functional on $H^1_{\Delta_N}(\mathbb{R}^n)$can be extended by Hahn-Banach to a continuous linear functional on B and can be identified with a vector of functions$(b_0, b_1, \dots, b_n)$with each $b_j \in L^\infty(\mathbb{R}^n)$.
We use this conclusion in the following way. Let ℓ be a continuous linear functional on $H^1_{\Delta_N}(\mathbb{R}^n)$. Then by Proposition(6.2.12) there is a function $b \in BMO_{\Delta_N}(\mathbb{R}^n)$so that:
$$\int_{\mathbb{R}^n} f(x)\overline{b(x)}dx = \ell(f).$$
However, by the discussion above, and by restricting the extended linear functional back to $H^1_{\Delta_N}(\mathbb{R}^n)$we have for$(f, R_{N,1}f, \dots, R_{N,n}f) = (f_0, \dots, f_n)$:
$$\ell(f) = \sum_{j=0}^n \int_{\mathbb{R}^n} f_j(x)\overline{b_j(x)}dx.$$

Using the definition of the $f_j = R_{N,j}f$ we see that:

$$\ell(f) \int_{\mathbb{R}^n} f(x) \overline{\left(b_0(x) + \sum_{j=1}^{n} R_{N,j}^* b_j(x) \right)} dx.$$

This then gives the decomposition that any $b \in BMO_{\Delta_N}(\mathbb{R}^n)$can be written as:

$$b = b_0 + \sum_{j=1}^{n} R_{N,j}^* b_j$$

with $b_j \in L^\infty(\mathbb{R}^n)$.

For the converse, we simply observe that from our Theorem(6.2.18), we obtained that $\mathbb{R}^n$ maps $H^1_{\Delta_N}(\mathbb{R}^n)$ to $L^1(\mathbb{R}^n)$. Hence, the boundedness of the Riesz transform R_N^* from $L^\infty(\mathbb{R}^n)$ to $BMO_{\Delta_N}(\mathbb{R}^n)$follows from duality of $H^1_{\Delta_N}(\mathbb{R}^n)$ with $BMO_{\Delta_N}(\mathbb{R}^n)$. We then have that any b that can be written as:

$$b = b_0 + \sum_{j=1}^{n} R_{N,j}^* b_j$$

with$b_j \in L^\infty(\mathbb{R}^n)$must belong to$BMO_{\Delta_N}(\mathbb{R}^n)$.

We next note that$H^1_{\Delta_N}(\mathbb{R}^n)$is a proper subspace of the classical $H^1(\mathbb{R}^n)$, which was proved by Yan in [53] from the viewpoint of the semigroup generated by Δ_N. And we now give a direct proof and provide a specific function f which lies in $H^1(\mathbb{R}^n)$but does not belong to $H^1_{\Delta_N}(\mathbb{R}^n)$. A related claim is made in [65].

Theorem (6.2.20)[178]: ([53]). $H^1_{\Delta_N}(\mathbb{R}^n) \subsetneq H^1(\mathbb{R}^n)$.

Proof: We first show that the containment $H^1_{\Delta_N}(\mathbb{R}^n) \subset H^1(\mathbb{R}^n)$holds. This follows directly from the fact that corresponding BMO spaces norm theH^1spaces, namely that:

$$\|f\|_{H^1_{\Delta_N}(\mathbb{R}^n)} \approx \sup_{\|b\|_{BMO_{\Delta_N}(\mathbb{R}^n)} \le 1} |\langle f, b\rangle_{L^2}(\mathbb{R}^n)|.$$

An identical statement holds for$H^1(\mathbb{R}^n)$and$BMO(\mathbb{R}^n)$. As shown in[65], $BMO(\mathbb{R}^n) \subsetneq BMO_{\Delta_N}(\mathbb{R}^n)$, and so we have

$$\|f\|_{H^1(\mathbb{R}^n)} \approx \sup_{\|b\|_{BMO(\mathbb{R}^n)} \le 1} |\langle f, b\rangle_{L^2}(\mathbb{R}^n)| \le \sup_{\|b\|_{BMO_{\Delta_N}(\mathbb{R}^n)} \le 1} |\langle f, b\rangle_{L^2}(\mathbb{R}^n)| \approx \|f\|_{H^1_{\Delta_N}(\mathbb{R}^n)}.$$

This gives the containment,$H^1_{\Delta_N}(\mathbb{R}^n) \subset H^1(\mathbb{R}^n)$.

We now show that there exists a function $f \in H^1(\mathbb{R}^n)$but $f \notin H^1_{\Delta_N}(\mathbb{R}^n)$. For the sake of simplicity, we just consider the example in dimension 1.

Define

$$f(x) := \frac{\chi_{[0,1]}(x)}{\sqrt{2}} - \frac{\chi_{[-1,0]}(x)}{\sqrt{2}}.$$

It is easy to see that$f(x)$ is supported in$[-1, 1]$, and$\int_{\mathbb{R}} f(x)dx = 0$.Moreover, we have

$$\|f\|_{L^2(\mathbb{R}^n)} = 1.$$

These implies that f is an atom of $H^1(\mathbb{R})$, which shows that $f \in H^1(\mathbb{R})$. From the definition of f, we obtain that $f_+(x) \frac{\chi_{[0,1]}(x)}{\sqrt{2}}$, and the even extension is

$$f_{+,e}(x) \frac{\chi_{[-1,1]}(x)}{\sqrt{2}}.$$

But, then it is immediate that $f_{+,e} \notin H^1(\mathbb{R})$since$\int_{\mathbb{R}^n} f_{+,e}(x)dx \neq 0$. One can also prove this by using the equivalent definition of $H^1(\mathbb{R})$via the radial maximal function. Similarly we have these estimates for $f_{-,e}$. Hence,$f_{+,e} \notin H^1(\mathbb{R})$and $f_{-,e} \notin H^1(\mathbb{R})$, which, combining the result in Theorem(6.2.18), implies that $f \notin H^1_{\Delta_N}(\mathbb{R})$.

Finally, we provide a description of the atoms in $H^1_{\Delta_N}(\mathbb{R}^n)$that connects back to the atom in $H^1(\mathbb{R}^n)$.

Proposition (6.2.21)[178]: Suppose $a(x)$ is an $H^1_{\Delta_N}(\mathbb{R}^n)$-atom supported in $B \subset \mathbb{R}^n$. Then we have

$$\int_{\mathbb{R}^n} a(x)dx = 0. \tag{76}$$

Moreover, if $B \cap \{x \in \mathbb{R}^n : x_n = 0\} \neq \emptyset$, we denote$B_+ = B \cap \mathbb{R}^n_+andB_- = B \cap \mathbb{R}^n_-$. Then we have

$$\int_{B_+} a(x)dx = \int_{B_-} a(x)dx = 0. \tag{77}$$

Proof: First note that from Theorem(6.2.20),$H^1_{\Delta_N}(\mathbb{R}^n) \subsetneq H^1(\mathbb{R}^n)$. Since $a(x)$ is an $H^1_{\Delta_N}(\mathbb{R}^n)$ atom, we have$a(x) \in H^1(\mathbb{R}^n)$, and hence(76) holds, where we use [180].

Second, suppose $B \cap \{x \in \mathbb{R}^n : x_n = 0\} \neq \emptyset$. Then we define $a_+(x) = a(x)|_{B_+}$and $a_-(x) = a(x)|_{B_-}$. Since $a(x) \in H^1_{\Delta_N}(\mathbb{R}^n)$, from Theorem(6.2.18) we obtain that both $a_{+,e}(x)$ and $a_{-,e}(x)$are in $H^1(\mathbb{R}^n)$, which implies that

$$\int_{\mathbb{R}^n} a_{+,e}(x)dx = \int_{\mathbb{R}^n} a_{-,e}(x)dx = 0.$$

Next we claim that$\int_{\mathbb{R}^n} a_{+,e}(x)dx = 0$. In fact,

$$\int_{\mathbb{R}^n} a_{+,e}(x)dx = \int_{\mathbb{R}^n_+} a_{+,e}(x)dx + \int_{\mathbb{R}^n_-} a_{+,e}(x)dx = 2\int_{\mathbb{R}^n_+} a_{+,e}(x)dx.$$

Hence,$\int_{\mathbb{R}^n} a_{+,e}(x)dx = 0$ implies that$\int_{\mathbb{R}^n} a_{+,e}(x)dx = 0$, i.e., $\int_{B_+} a(x)dx = 0$. Similarly we obtain that$\int_{B_-} a(x)dx = 0$. Hence(77) holds.

We turn to proving Theorem(6.2.2). There are two parts to this Theorem, and upper and lower bound, and we focus first on the (easier) upper bound.

Recall that, for notational simplicity, we are letting

$$\Pi_l(h, g) \coloneqq h.R^*_{N,l}(g) - g.R_{N,l}(h),$$

where$R_{N,l} = \frac{\partial}{\partial x_l}\Delta_N^{-\frac{1}{2}}$ for$1 \leq l \leq n$. We now prove the following theorem.

Theorem (6.2.22)[178]: If $b \in BMO_{\Delta_N}(\mathbb{R}^n)$, then for $1 \leq l \leq n$, the commutator

$$[b, \mathrm{R}_{N,l}](f)(x) = b(x)b, \mathrm{R}_{N,l}(f)(x) - \mathrm{R}_{N,l}(bf)(x)$$

is a bounded map on$L^2(\mathbb{R}^n)$, with operator norm

$$\left\|[b, \mathrm{R}_{N,l}]: L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)\right\| \leq C\|b\|_{BMO_{\Delta_N}(\mathbb{R}^n)}.$$

Proof: Suppose b is in $BMO_{\Delta_N}(\mathbb{R}^n)$. Then according to [65], we have that $b_{+,e}(x)BMO(\mathbb{R}^n)$ and $b_{-,e}(x)BMO(\mathbb{R}^n)$, and moreover,

$$\|b\|_{BMO_{\Delta_N}(\mathbb{R}^n)} \approx \|b_{+,e}\|_{BMO(\mathbb{R}^n)}\|b_{-,e}\|_{BMO(\mathbb{R}^n)}.$$

For every$f \in L^2(\mathbb{R}^n)$, we have

$$\left\|[b,\mathrm{R}_{N,l}](f)\right\|^2_{L^2(\mathbb{R}^n)} = \int_{\mathbb{R}^n_+}[b,\mathrm{R}_{N,l}](f)(x)^2dx + \int_{\mathbb{R}^n_-}[b,\mathrm{R}_{N,l}](f)(x)^2dx =: I + II.$$

For the term$I$, note that when $x\in\mathbb{R}^n_+$, we have

$$\begin{aligned}[b,\mathrm{R}_{N,l}](f)(x) &= b(x)b,\mathrm{R}_{N,l}(f)(x) - \mathrm{R}_{N,l}(bf)(x)\\ &= b_{+,e}(x)R_l(f_{+,e})(x) - R_l(b_{+,e}f_{+,e})(x) = [b_{+,e},\mathrm{R}_l](f_{+,e})(x),\end{aligned}$$

which implies that

$$\begin{aligned}I = \int_{\mathbb{R}^n_+}[b,\mathrm{R}_{N,l}](f)(x)^2dx &= \int_{\mathbb{R}^n_+}[b_{+,e},\mathrm{R}_l](f_{+,e})(x)^2dx\\ &\le \int_{\mathbb{R}^n}[b_{+,e},\mathrm{R}_l](f_{+,e})(x)^2dx\\ &\le C\left\|b_{+,e}\right\|^2_{BMO(\mathbb{R}^n)}\left\|f_{+,e}\right\|^2_{L^2(\mathbb{R}^n)},\end{aligned}$$

whereR$_l$is the classical l-th Riesz transform$\frac{\partial}{\partial \mathrm{x}_l}\Delta^{-\frac{1}{2}}$.

For the last estimate we use the result [3], which applies since we know from that R$_{N,l}$is a Calderón–Zygmund kernel. Similarly we can obtain that

$$II \le C\left\|b_{-,e}\right\|^2_{BMO(\mathbb{R}^n)}\left\|f_{-,e}\right\|^2_{L^2(\mathbb{R}^n)}.$$

Combining the estimates for I and II above, we obtain that

$$\begin{aligned}\left\|[b,\mathrm{R}_{N,l}](f)\right\|^2_{L^2(\mathbb{R}^n)} &\le C\left\|b_{+,e}\right\|^2_{BMO(\mathbb{R}^n)}\left\|f_{+,e}\right\|^2_{L^2(\mathbb{R}^n)} + C\left\|b_{+,e}\right\|^2_{BMO(\mathbb{R}^n)}\left\|f_{-,e}\right\|^2_{L^2(\mathbb{R}^n)}\\ &\le C\|b\|^2_{BMO_{\Delta_N}(\mathbb{R}^n)}\left(\left\|f_{+,e}\right\|^2_{L^2(\mathbb{R}^n)} + \left\|f_{-,e}\right\|^2_{L^2(\mathbb{R}^n)}\right)\\ &\le C\|b\|^2_{BMO_{\Delta_N}(\mathbb{R}^n)}\|f\|^2_{L^2(\mathbb{R}^n)},\end{aligned}$$

which yields that$\left\|[b,\mathrm{R}_{N,l}]: L^2(\mathbb{R}^n)\to L^2(\mathbb{R}^n)\right\| \le C\|b\|_{BMO_{\Delta_N}(\mathbb{R}^n)}$.

Theorem (6.2.23)[178]: Let $g,h\in L^\infty(\mathbb{R}^n)$with compact supports. Then for$1\le l\le n$,

$$\|\Pi_l(h,g)\|_{H^1_{\Delta_N}(\mathbb{R}^n)} \le C\|g\|_{L^2(\mathbb{R}^n)}\|h\|_{L^2(\mathbb{R}^n)}.$$

Proof. By the duality result of [65], stated in Proposition(6.2.12), we know that $H^1_{\Delta_N}(\mathbb{R}^n)^* = BMO_{\Delta_N}(\mathbb{R}^n)$. A simple duality computation shows for $b\in BMO_{\Delta_N}(\mathbb{R}^n)$and for any $g,h\in L^\infty(\mathbb{R}^n)$with compact supports:

$$\langle b,\Pi_l(g,h)\rangle_{L^2(\mathbb{R}^n)} = \langle b,\mathrm{R}^*_{N,l}(g)h - \mathrm{R}_{N,l}(h)g\rangle_{L^2(\mathbb{R}^n)} = \langle g,[b,\mathrm{R}_{N,l}(h)]h\rangle_{L^2(\mathbb{R}^n)}.$$

Thus, from Theorem(6.2.22), we obtain that

$$\left|\langle b,\Pi_l(g,h)\rangle_{L^2(\mathbb{R}^n)}\right| \le C\|b\|_{BMO_{\Delta_N}(\mathbb{R}^n)}\|g\|_{L^2(\mathbb{R}^n)}\|h\|_{L^2(\mathbb{R}^n)}.$$

This, together with the duality of $H^1_{\Delta_N}(\mathbb{R}^n)$with $BMO_{\Delta_N}(\mathbb{R}^n)$shows that $\Pi_l(g,h)$is in $H^1_{\Delta_N}(\mathbb{R}^n)$. And then by testing $\Pi_l(g,h)$against $b\in BMO_{\Delta_N}(\mathbb{R}^n)$functions, we find:

$$\begin{aligned}\|\Pi_l(g,h)\|_{H^1_{\Delta_N}(\mathbb{R}^n)} &\approx \sup_{\|b\|_{BMO_{\Delta_N}(\mathbb{R}^n)}\le 1}\left|\langle \Pi_l(g,h),b\rangle_{L^2(\mathbb{R}^n)}\right|\\ &\le C\|g\|_{L^2(\mathbb{R}^n)}\|h\|_{L^2(\mathbb{R}^n)}\sup_{\|b\|_{BMO_{\Delta_N}(\mathbb{R}^n)}\le 1}\|b\|_{BMO_{\Delta_N}(\mathbb{R}^n)}\\ &\le C\|g\|_{L^2(\mathbb{R}^n)}\|h\|_{L^2(\mathbb{R}^n)}.\end{aligned}$$

The proof of the lower bound is more algorithmic in nature and follows a proof strategy developed by Uchiyama in[182]. We begin with a fact that will play a prominent role in the algorithm below. It is a modification of a related fact for the standard Hardy space $H^1(\mathbb{R}^n)$.

Lemma (6.2.24)[178]: Suppose f is a function satisfying: $\int_{\mathbb{R}^n} f(x)dx = 0$, and $|f(x)| \le \chi_{B(x_0,1)}(x) + \chi_{B(y_0,1)}(x)$, where $|x_0 - y_0| := M > 10$. Then we have

$$\|f\|_{H^1_{\Delta_N}(\mathbb{R}^n)} \le C_n \log M. \tag{78}$$

Proof: First note that

$$f^+_{\Delta_N}(x) = \sup_{t>0}\left|e^{-t\Delta_N} f(x)\right| = \sup_{t>0}\left|\int_{\mathbb{R}^n} p_{t,\Delta_N}(x,y)f(y)dy\right| \sup_{t>0}\int_{\mathbb{R}^n}\left|p_{t,\Delta_N}(x,y)\right|dy \le C.$$

Hence, we obtain that

$$\int_{B(x_0,5)} f^+_{\Delta_N}(x)dx + \int_{B(y_0,5)} f^+_{\Delta_N}(x)dx \le C_n.$$

Now it suffices to estimate

$$\int_{\mathbb{R}^n\setminus(B(x_0,5)\cup B(y_0,5))} f^+_{\Delta_N}(x)dx.$$

To see this, we write it as

$$\int_{\mathbb{R}^n/B(x_0,2M)} f^+_{\Delta_N}(x)dx + \int_{B(x_0,2M)\setminus(B(x_0,5)\cup B(y_0,5))} f^+_{\Delta_N}(x)dx =: I + II.$$

We now estimate the term I. First note that from Hölder's regularity (67) of the heat kernel $p_{t,\Delta_N}(x,y)$, we have

$$\left|p_{t,\Delta_N}(x,y) - p_{t,\Delta_N}(x,x_0)\right| \le C\left(\frac{|y-x_0|}{\sqrt{t}+|x-x_0|}\right)\frac{\sqrt{t}}{\left(\sqrt{t}+|x-x_0|\right)^{n+1}}$$

for $|y-x_0| < \sqrt{t}$. Moreover, when $|y-x_0| \ge \sqrt{t}$, we have

$$\begin{aligned}\left|p_{t,\Delta_N}(x,y) - p_{t,\Delta_N}(x,x_0)\right| &\le \left|p_{t,\Delta_N}(x,y)\right| + \left|p_{t,\Delta_N}(x-x_0)\right| \\ &\le C\frac{e^{-|x,x_0|^2/ct}}{t^{n/2}} + \frac{e^{-|x-y|^2/ct}}{t^{n/2}} \\ &\le C\left(\frac{|y-x_0|}{\sqrt{t}}\right)\frac{e^{-|x,x_0|^2/ct}}{t^{n/2}} \\ &\le C\left(\frac{|y-x_0|}{\sqrt{t}+|x-x_0|}\right)\frac{\sqrt{t}}{\left(\sqrt{t}+|x-x_0|\right)^{n+1}}.\end{aligned}$$

Now note that from the cancellation condition of f and Hölder's regularity of the heat kernel $p_t(x,y)$ as above, we have

$$\begin{aligned}f^+_{\Delta_N}(x) &= \sup_{t>0}\left|\int_{\mathbb{R}^n}\left[p_{t,\Delta_N}(x,y) - p_{t,\Delta_N}(x,x_0)\right]f(y)dy\right| \\ &\le C\sup_{t>0}\int_{B(x_0,1)\cup B(y_0,1)}\left(\frac{|y-x_0|}{\sqrt{t}+|x-x_0|}\right)\frac{\sqrt{t}}{\left(\sqrt{t}+|x-x_0|\right)^{n+1}}dy \\ &\le C_n\frac{|y_0-x_0|}{|x-x_0|^{n+1}} = C_n\frac{M}{|x-x_0|^{n+1}}.\end{aligned}$$

As a consequence, we obtain that

$$I \leq \int_{\mathbb{R}^n/B(x_0,2M)} C_n \frac{M}{|x-x_0|^{n+1}} dx \leq C_n.$$

We now turn to the term II. Note that when $x \in B(x_0,2M)\backslash\big(B(x_0,5) \cup B(y_0,5)\big)$, we have

$$\left|\int_{\mathbb{R}^n} p_{t,\Delta_N}(x,y) f(y) dy\right| \leq \int_{B(x_0,1)} \left|p_{t,\Delta_N}(x,y)\right| dy + \int_{B(y_0,1)} \left|p_{t,\Delta_N}(x,y)\right| dy.$$

When $t > 1$, from the size estimate of the heat kernel$p_{t,\Delta_N}(x,y)$, we have

$$\left|\int_{\mathbb{R}^n} p_{t,\Delta_N}(x,y) f(y) dy\right| \leq C \frac{1}{|x-x_0|^n} + C \frac{1}{|x-y_0|^n}.$$

When $t \leq 1$, similarly we obtain that

$$\left|\int_{\mathbb{R}^n} p_{t,\Delta_N}(x,y) f(y) dy\right| \leq C \frac{1}{|x-x_0|^{n+1}} + C \frac{1}{|x-y_0|^{n+1}} \leq C \frac{1}{|x-x_0|^n} + C \frac{1}{|x-y_0|^n}.$$

Thus,

$$\begin{aligned} II &\leq \int_{B(x_0,2M)\backslash(B(x_0,5)\cup B(y_0,5))} f^+_{\Delta_N}(x) dx \\ &\leq C \int_{B(x_0,2M)\backslash(B(x_0,5)\cup B(y_0,5))} \frac{1}{|x-x_0|^n} + \frac{1}{|x-y_0|^n} dx \\ &\leq C_n \log M. \end{aligned}$$

Combining all the estimates above, we obtain that

$$\|f\|_{H^1_{\Delta_N}(\mathbb{R}^n)} = \left\|f^+_{\Delta_N}\right\|_{L^1(\mathbb{R}^n)} \leq C_n \log M.$$

Suppose $1 \leq l \leq n$. Ideally, given an $H^1_{\Delta_N}(\mathbb{R}^n)$-atom a, we would like to find $g,h \in L^2(\mathbb{R}^n)$such that $\Pi_l(g,h) = a$pointwise. While this can't be accomplished in general, the Theorem below shows that it is "almost" true.

Theorem (6.2.25)[178]: Suppose$1 \leq l \leq n$. For every $H^1_{\Delta_N}(\mathbb{R}^n)$-atom $a(x)$ and for all $\varepsilon > 0$ there exist a large positive number M and $g,h \in L^\infty(\mathbb{R}^n)$with compact supports such that:

$$\|a - \Pi_l(h,g)\|_{H^1_{\Delta_N}(\mathbb{R}^n)} < \varepsilon$$

and$\|g\|_{L^2(\mathbb{R}^n)}\|h\|_{L^2(\mathbb{R}^n)} \leq CM^n$.

Proof: Let $a(x)$ be an $H^1_{\Delta_N}(\mathbb{R}^n)$-atom, supported in$B(x_0,r)$.We first consider the construction of the bilinear form $\Pi_l(h,g)$for$1 \leq l \leq n-1$ and the approximation to$a(x)$. To begin with, for the ball $B(x_0,r)$, we now consider the following cases: Case1:$x_{0,n} \geq 0$; Case 2:$x_{0,n} < 0$.

We first consider Case1. To begin with, fix$\varepsilon > 0$. Choose $M \in [100,\infty)$sufficiently large so that $\frac{\log M}{M} < \varepsilon$. Now select$y_0 \in \mathbb{R}^n_+$in the following way: for $1 \leq i \leq n$, choose $y_{0,i} > 0$ such that$y_{0,i} - x_{0,i} = \frac{Mr}{\sqrt{n}}$, where $x_{0,i}$(reps. $y_{0,i}$) is the ith coordinate of x_0(reps y_0).

Note that for this y_0, it is clear that $B(y_0,r) \subset \mathbb{R}^n_+$and we have$|x_0 - y_0| = Mr$. Moreover, for any $y \in B(y_0,r)$, we also have$|x_0 - y| > \frac{Mr}{2}$. We set

$$g(x) := \chi_{B(y_0,r)}(x) \text{and} h(x) := \frac{a(x)}{R^*_{N,l}g(x_0)}. \tag{79}$$

We first claim that

$$\left|R^*_{N,l}g(x_0)\right| \geq CM^{-n}, \qquad 1 \leq l \leq n-1. \tag{80}$$

In fact, for$l = 1, \dots, n-1$, from Proposition (6.2.4), we have

$$R^*_{N,l}g(x_0) = \left|\int_{B(y_0,r)} R_{N,l}(y,x_0)dy\right|$$

$$= C_n \left|\int_{B(y_0,r)} \left(\frac{y_l - x_{0,l}}{|x_0 - y|^{n+1}} + \frac{y_l - x_{0,l}}{\left(|x'_0 - y'|^2 + |x_{0,n} + y_n|^2\right)^{\frac{n+1}{2}}}\right)dy\right|$$

$$= C_n|y_l - x_{0,l}| \left|\int_{B(y_0,r)} \left(\frac{1}{|x_0 - y|^{n+1}} + \frac{1}{\left(|x'_0 - y'|^2 + |x_{0,n} + y_n|^2\right)^{\frac{n+1}{2}}}\right)dy\right|$$

$$\geq CMr \int_{B(y_0,r)} \frac{1}{|x_0 - y|^{n+1}}dy \geq CM^{-n}.$$

As a consequence, we get that the claim(80) holds.

As for Case2, we handle it in a symmetric way as follows. Fix $\varepsilon > 0$. Choose $M \in [100, \infty)$ sufficiently large so that$\frac{\log M}{M} < \varepsilon$. Now select $y_0 \in \mathbb{R}^n_+$ in the following way: for $1 \leq i \leq n$, choose $y_{0,i} > 0$ such that $y_{0,i} - x_{0,i} = -\frac{Mr}{\sqrt{n}}$. Note that for this y_0, it is clear that $B(y_0,r) \subset \mathbb{R}^n_-$and we have$|x_0 - y_0| = Mr$. Moreover, for any $y \in B(y_0,r)$, we also have $|x_0 - y| > \frac{Mr}{2}$. We now define the functions g and h as in(79), and the following the same estimates, we can obtain that the claim(80) holds.

From the definitions of the functions g and h, we obtain that $\text{supp}g(x) = B(y_0,r)$ and $\text{supp}h(x) = B(x_0,r)$. Moreover, from(80) we obtain that

$$\|g\|_{L^2(\mathbb{R}^n)} \approx r^{\frac{n}{2}} \text{and} \|h\|_{L^2(\mathbb{R}^n)} = \frac{1}{\left|R_{N,l}g(x_0)\right|}\|a\|_{L^2(\mathbb{R}^n)} \leq CM^n r^{-\frac{n}{2}}.$$

Hence$\|g\|_{L^2(\mathbb{R}^n)}\|h\|_{L^2(\mathbb{R}^n)} \leq CM^n$. Now write

$$a(x) - \left(h(x)R^*_{N,l}g(x) - g(x)R_{N,l}h(x)\right) = a(x)\frac{R^*_{N,l}g(x_0) - R^*_{N,l}g(x_0)}{R^*_{N,l}g(x_0)} - g(x)R_{N,l}h(x)$$

$$=: W_1(x) + W_2(x).$$

By definition, it is obvious that $W_1(x)$ is supported on $B(x_0,r)$ and $W_2(x)$ is supported on $B(y_0,r)$.

We first turn to $W_1(x)$. For$x \in B(x_0,r)$,

$$|W_1(x)| = |a(x)|\frac{\left|R^*_{N,l}g(x_0) - R^*_{N,l}g(x)\right|}{R^*_{N,l}g(x_0)}$$

$$\leq CM^n \|a\|_{L^\infty(\mathbb{R}^n)} \int_{B(y_0,r)} \left|R_{N,l}(y,x_0) - R_{N,l}(y,x)\right| dy$$

$$\leq C\frac{M^n}{r^n} \int_{B(y_0,r)} \frac{|x-x_0|}{|x-y|^{n+1}} dy$$

$$\leq C\frac{1}{Mr^n}.$$

Hence $|W_1(x)| \leq C\frac{1}{Mr^n}\chi_{B(x_0,r)}(x)$.

We next estimate $W_2(x)$. From the definition of $g(x)$, we have

$$|W_2(x)| = \chi_{B(y_0,r)}(x)\left|R_{N,l}h(x)\right|$$

$$= \chi_{B(y_0,r)}(x)\frac{1}{\left|R_{N,l}^* g(x_0)\right|}\left|\int_{B(y_0,r)} R_{N,l}h(x,y)a(y)dy\right|$$

$$= \chi_{B(y_0,r)}(x)\frac{1}{\left|R_{N,l}^* g(x_0)\right|}\left|\int_{B(y_0,r)} R_{N,l}(x,y)a_+(y)dy\right|,$$

where the last equality follows from the fact that $x \in B(y_0,r) \subset \mathbb{R}^n_+$and from the definition of the Riesz kernel $R_N(x,y)$ as in(68). Hence, from the cancellation property of $a_+(y)$, we get

$$|W_2(x)| = \chi_{B(y_0,r)}(x)\frac{1}{\left|R_{N,l}^* g(x_0)\right|}\left|\int_{B(x_0,r)} \left(R_{N,l}(x,y) - R_{N,l}(x,x_0)\right) a_+(y)dy\right|$$

$$\leq C\chi_{B(y_0,r)}(x)M^n \int_{B(x_0,r)} \|a\|_{L^\infty(\mathbb{R}^n)} \frac{|y-x_0|}{|x-x_0|^{n+1}} dy$$

$$\leq \frac{C}{M^n}\chi_{B(y_0,r)}(x).$$

Combining the estimates of W_1 and W_2, we obtain that

$$\left|a(x) - \left(h(x)R_{N,l}^* g(x) - g(x)R_{N,l}h(x)\right)\right| \leq \frac{C}{M^n}\left(\chi_{B(x_0,r)}(x) + \chi_{B(y_0,r)}(x)\right). \quad (81)$$

Next we point out that

$$\int \left[a(x) - \left(h(x)R_{N,l}^* g(x) - g(x)R_{N,l}h(x)\right)\right] dx$$

$$= \int a(x)dx - \int \left(h(x)R_{N,l}^* g(x) - g(x)R_{N,l}h(x)\right) dx = 0, \quad (82)$$

since $a(x)$ has cancellation (Proposition(6.2.21)) and the second integral equals 0 just by the definitions of g and h.

Then the size estimate(81) and the cancellation(82), together with Lemma(6.2.24), imply that

$$\left\|a(x) - \left(h(x)R_{N,l}^* g(x) - g(x)R_{N,l}h(x)\right)\right\|_{H^1_{\Delta_N}(\mathbb{R}^n)} \leq C\frac{\log M}{M} \leq C\epsilon.$$

This proves the result for $1 \leq l \leq n-1$. We now consider the bilinear form $\Pi_l(g,h)$and its approximation to $a(x)$. Again, for the ball $B(x_0,r)$, we now consider the following cases: Case1:$x_0,n \geq 0$;Case2: $x_{0,n}, < 0$. It suffices to consider the Case 1 since the other can be handled symmetrically. In this case, for $x_0$with$x_{0,n} \geq 0$, choose y_0 such that $y_{0,i} - x_{0,i} = \frac{Mr}{\sqrt{n}}$ for $i = 1, \ldots, n$. We now define the functions g and h as in(79). This, together with Proposition(6.2.4), yields

$$R_{N,l}^{*} g(x) \left| \int_{B(x_0,r)} R_{N,l}(x, x_0) dy \right|$$

$$= C_n \left| \int_{B(y_0,r)} \left(\frac{y_n - x_{0,n}}{|x_0 - y|^{n+1}} + \frac{x_{0,n} - y_n}{\left(\left|x_0' - y'\right|^2 + \left|x_{0,n} + y_n\right|^2\right)^{\frac{n+1}{2}}} \right) dy \right|$$

$$\geq C_n \left| \int_{B(y_0,r)} \frac{y_n - x_{0,n}}{|x_0 - y|^{n+1}} dy \right|$$

$$= C_n \left| y_n - x_{0,n} \right| \left| \int_{B(y_0,r)} \frac{1}{|x_0 - y|^{n+1}} dy \right|$$

$$\geq C M^n.$$

Here, we obtain that the claim(80) holds for these g and$h$.

Now following the approximation as that for$R_{N,l}$with$1 \leq l \leq n - 1$, we obtain that

$$\left\| a(x) - \left(h(x) R_{N,l}^{*} g(x) - g(x) R_{N,l} h(x) \right) \right\|_{H_{\Delta_N}^1(\mathbb{R}^n)} \leq C \frac{\log M}{M} \leq C\epsilon. \qquad (83)$$

With this approximation result, we can now prove the main Theorem(6.2.2), restated below for the convenience of the reader.

Theorem (6.2.26)[178]: Suppose$1 \leq l \leq n$. For any $f \in H_{\Delta_N}^1(\mathbb{R}^n)$there exists sequences $\{\lambda_j^k\} \in \ell^1$ and functions $g_j^k, h_j^k \in L^\infty(\mathbb{R}^n)$with compact supports such that

$$f = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \lambda_j^k \Pi_l\left(g_j^k, h_j^k\right).$$

Moreover, we have that:

$$\|f\|_{H_{\Delta_N}^1(\mathbb{R}^n)} \approx \inf \left\{ \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \left|\lambda_j^k\right| \left\|g_j^k\right\|_{L^2(\mathbb{R}^n)} \left\|h_j^k\right\|_{L^2(\mathbb{R}^n)} : f = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \lambda_j^k \Pi_l\left(g_j, h_j\right) \right\}.$$

Proof: By Theorem(6.2.23) we have that $\|\Pi_l(g,h)\|_{H_{\Delta_N}^1(\mathbb{R}^n)} \leq C \|g\|_{L^2(\mathbb{R}^n)} \|h\|_{L^2(\mathbb{R}^n)}$, it is immediate that we have for any representation of $f = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \lambda_j^k \Pi_l\left(g_j, h_j\right)$ that

$$\|f\|_{H_{\Delta_N}^1(\mathbb{R}^n)} \leq C \inf \left\{ \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \left|\lambda_j^k\right| \left\|g_j^k\right\|_{L^2(\mathbb{R}^n)} \left\|h_j^k\right\|_{L^2(\mathbb{R}^n)} : f = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \lambda_j^k \Pi_l\left(g_j^k, h_j^k\right) \right\}.$$

We turn to show that the other inequality hold and that it is possible to obtain such a decomposition for any $f \in H_{\Delta_N}^1(\mathbb{R}^n)$. By the atomic decomposition for $H_{\Delta_N}^1(\mathbb{R}^n)$, Theorem(6.2.18), for any $f \in H_{\Delta_N}^1(\mathbb{R}^n)$we can find a sequence$\{\lambda_j^k\} \in \ell^1$and sequence of $H_{\Delta_N}^1(\mathbb{R}^n)$-atoms a_j^1so that $f = \sum_{j=1}^{\infty} \lambda_j^k a_j^1and\sum_{j=1}^{\infty} \left|\lambda_j^k\right| \leq C_0 \|f\|_{H_{\Delta_N}^1(\mathbb{R}^n)}$.

We explicitly track the implied absolute constant C_0 appearing from the atomic decomposition since it will play a role in the convergence of the approach. Fix $\varepsilon > 0$ so that $\varepsilon C_0 < 1$. Then we also have a large positive number M with $\frac{\log M}{M} < \varepsilon$. We apply Theorem (6.2.25) to each atom a_j^1. So there exists $g_j^k, h_j^k \in L^\infty(\mathbb{R}^n)$ with compact supports and satisfying $\|g_j^k\|_{L^2(\mathbb{R}^n)} \|h_j^k\|_{L^2(\mathbb{R}^n)} \le CM^n$ and

$$\left\|a_j^1 - \Pi_l(g_j^1, h_j^1)\right\|_{H^1_{\Delta_N}(\mathbb{R}^n)} < \varepsilon \forall_j.$$

Now note that we have

$$f = \sum_{j=1}^{\infty} \lambda_j^1 a_j^1 = \sum_{j=1}^{\infty} \lambda_j^1 \Pi_l(g_j^1, h_j^1) + \sum_{j=1}^{\infty} \lambda_j^1 \left(a_j^1 - \Pi_l(g_j^1, h_j^1)\right) := M_1 + E_1.$$

Observe that we have

$$\|\mathrm{E}_1\|_{H^1_{\Delta_N}(\mathbb{R}^n)} \le \sum_{j=1}^{\infty} |\lambda_j^1| \left\|a_j^1 - \Pi_l(g_j^1, h_j^1)\right\|_{H^1_{\Delta_N}(\mathbb{R}^n)} \le \varepsilon \sum_{j=1}^{\infty} |\lambda_j^1| \le \varepsilon C_0 \|f\|_{H^1_{\Delta_N}(\mathbb{R}^n)}.$$

We now iterate the construction on the function E_1. Since $\mathrm{E}_1 \in H^1_{\Delta_N}(\mathbb{R}^n)$, we can apply the atomic decomposition in $H^1_{\Delta_N}(\mathbb{R}^n)$, Theorem(6.2.20), to find a sequence $\{\lambda_j^k\} \in \ell^1$ and a sequence of $H^1_{\Delta_N}(\mathbb{R}^n)$-atoms $\{a_j^2\}$ so that $E_1 = \sum_{j=1}^{\infty} \lambda_j^k a_j^2$ and

$$\sum_{j=1}^{\infty} |\lambda_j^2| \le C_0 \|E_1\|_{H^1_{\Delta_N}(\mathbb{R}^n)} \le \varepsilon C_0^2 \|f\|_{H^1_{\Delta_N}(\mathbb{R}^n)}.$$

Again, we will apply Theorem (6.2.25) to each atom a_j^2. So there exist $g_j^2, h_j^2 \in L^\infty(\mathbb{R}^n)$ with compact supports and satisfying $\|g_j^2\|_{L^2(\mathbb{R}^n)} \|h_j^2\|_{L^2(\mathbb{R}^n)} \le CM^n$ and

$$\left\|a_j^2 - \Pi_l(g_j^2, h_j^2)\right\|_{H^1_{\Delta_N}(\mathbb{R}^n)} < \varepsilon, \forall_j.$$

We then have that:

$$E_1 = \sum_{j=1}^{\infty} \lambda_j^2 a_j^2 = \sum_{j=1}^{\infty} \lambda_j^2 \Pi_l(g_j^2, h_j^2) + \sum_{j=1}^{\infty} \lambda_j^2 \left(a_j^2 - \Pi_l(g_j^2, h_j^2)\right) := M_2 + E_2.$$

But, as before observe that

$$\|\mathrm{E}_2\|_{H^1_{\Delta_N}(\mathbb{R}^n)} \le \sum_{j=1}^{\infty} |\lambda_j^2| \left\|a_j^2 - \Pi_l(g_j^2, h_j^2)\right\|_{H^1_{\Delta_N}(\mathbb{R}^n)} \le \varepsilon \sum_{j=1}^{\infty} |\lambda_j^2| \le (\varepsilon C_0)^2 \|f\|_{H^1_{\Delta_N}(\mathbb{R}^n)}.$$

And, this implies for f that we have:

$$f = \sum_{j=1}^{\infty} \lambda_j^1 a_j^1 = \sum_{j=1}^{\infty} \lambda_j^1 \Pi_l(g_j^1, h_j^1) + \sum_{j=1}^{\infty} \lambda_j^1 \left(a_j^1 - \Pi_l(g_j^1, h_j^1)\right)$$

$$= M_1 + E_1 = M_1 + M_2 + E_2 = \sum_{k=1}^{2} \sum_{j=1}^{\infty} \lambda_j^k \Pi_l(g_j^k, h_j^k) + E_2.$$

Repeating this construction for each $1 \le k \le K$ produces functions $g_j^k, h_j^k \in L^\infty(\mathbb{R}^n)$ with compact supports and satisfying $\|g_j^k\|_{L^2(\mathbb{R}^n)} \|h_j^k\|_{L^2(\mathbb{R}^n)} \le CM^n$ for all j, sequences $\{\lambda_j^k\} \in \ell^1$ with $\left\|\{\lambda_j^k\}\right\|_{\ell^1} \le \varepsilon^{k-1} C_0^2 \|f\|_{H^1_{\Delta_N}(\mathbb{R}^n)}$, and a function $E_K \in H^1_{\Delta_N}(\mathbb{R}^n)$ with $\|E_K\|_{H^1_{\Delta_N}(\mathbb{R}^n)} (\varepsilon C_0)^K \|f\|_{H^1_{\Delta_N}(\mathbb{R}^n)}$ so that

$$f=\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\lambda_j^k\Pi_l\left(g_j^k,h_j^k\right)+E_K.$$

Passing $K\to\infty$gives the desired decomposition of $f=\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\lambda_j^k\Pi_l\left(g_j^k,h_j^k\right)$. We also have that:

$$\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\left|\lambda_j^k\right|\leq\sum_{k=1}^{\infty}\varepsilon^{-1}(\varepsilon C_0)^k\|f\|_{H^1_{\Delta_N}(\mathbb{R}^n)}=\frac{C_0}{1-\varepsilon C_0}\|f\|_{H^1_{\Delta_N}(\mathbb{R}^n)}.$$

Finally, we have

Theorem (6.2.27)[178]: Suppose $b\in\cup_{p\geq1}L^p_{loc}(\mathbb{R}^n)$.

If b is in $BMO_{\Delta_N}(\mathbb{R}^n)$, then for$1\leq l\leq n$, the commutator

$$\left[b,R_{N,l}\right](f)(x)=b(x)R_{N,l}(f)(x)-R_{N,l}(bf)(x)$$

is a bounded map on $L^2(\mathbb{R}^n)$, with operator norm

$$\left\|\left[b,R_{N,l}\right]:L^2(\mathbb{R}^n)\to L^2(\mathbb{R}^n)\right\|\leq C\|b\|_{BMO_{\Delta_N}(\mathbb{R}^n)}.$$

Conversely, for$1\leq l\leq n$, if $\left[b,R_{N,l}\right]$are bounded on $L^2(\mathbb{R}^n)$ then b is in $BMO_{\Delta_N}(\mathbb{R}^n)and\|b\|_{BMO_{\Delta_N}(\mathbb{R}^n)}\leq C\left\|\left[b,R_{N,l}\right]:L^2(\mathbb{R}^n)\to L^2(\mathbb{R}^n)\right\|$.

Proof. The upper bound in this theorem is contained in Theorem (6.2.23). For the lower bound, we first note that from Theorem(6.2.18), $H^1_{\Delta_N}(\mathbb{R}^n)$has equivalent characterizations via atoms, which shows that $H^1_{\Delta_N}(\mathbb{R}^n)\cap L^\infty_c(\mathbb{R}^n)$is dense in $H^1_{\Delta_N}(\mathbb{R}^n)$with respect to the $H^1_{\Delta_N}(\mathbb{R}^n)$norm, where we use $L^\infty_c(\mathbb{R}^n)$to denote the L^∞function with compact supports. Then using the weak factorization in Theorem(6.2.2) we have that for $f\in H^1_{\Delta_N}(\mathbb{R}^n)\cap L^\infty_c(\mathbb{R}^n)$,

$$\left|\langle b,f\rangle_{L^2(\mathbb{R}^n)}\right|\leq\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\left|\lambda_j^k\right|\left|\langle b,\Pi_l\left(g_j^k,h_j^k\right)\rangle_{L^2(\mathbb{R}^n)}\right|=\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\left|\lambda_j^k\right|\left|\langle g_j^k,\left[b,R_{N,l}\right]h_j^k\rangle_{L^2(\mathbb{R}^n)}\right|.$$

Hence we have that

$$\begin{aligned}\left|\langle b,f\rangle_{L^2(\mathbb{R}^n)}\right|&\leq\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\left|\lambda_j^k\right|\left\|\left[b,R_{N,l}\right]\left(h_j^k\right)\right\|_{L^2(\mathbb{R}^n)}\left\|g_j^k\right\|_{L^2(\mathbb{R}^n)}\\&\leq\left\|\left[b,R_{N,l}\right]:L^2(\mathbb{R}^n)\to L^2(\mathbb{R}^n)\right\|\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\left|\lambda_j^k\right|\left\|g_j^k\right\|_{L^2(\mathbb{R}^n)}\left\|h_j^k\right\|_{L^2(\mathbb{R}^n)}\\&\leq C\left\|\left[b,R_{N,l}\right]:L^2(\mathbb{R}^n)\to L^2(\mathbb{R}^n)\right\|\|f\|_{H^1_{\Delta_N}(\mathbb{R}^n)}.\end{aligned}$$

By the duality between $BMO_{\Delta_N}(\mathbb{R}^n)$and $H^1_{\Delta_N}(\mathbb{R}^n)$we have that:

$$\|b\|_{BMO_{\Delta_N}(\mathbb{R}^n)}\approx\sup_{\|f\|_{H^1_{\Delta_N}(\mathbb{R}^n)\leq1}}\left|\langle b,f\rangle_{L^2(\mathbb{R}^n)}\right|\leq C\left\|\left[b,R_{N,l}\right]:L^2(\mathbb{R}^n)\to L^2(\mathbb{R}^n)\right\|$$

Theorem (6.2.28)[178]: If b is in $BMO_{\Delta_N}(\mathbb{R}^n)$, then for$1<\alpha<n$, the commutator

$$\left[b,\Delta_N^{-\alpha/2}\right](f)(x)=b(x)\Delta_N^{-\alpha/2}(f)(x)-\Delta_N^{-\alpha/2}(bf)(x)$$

is a bounded map from$L^p(\mathbb{R}^n)$to $L^q(\mathbb{R}^n)$with operator norm

$$\left\|\left[b,\Delta_N^{-\alpha/2}\right]:L^p(\mathbb{R}^n)\to L^q(\mathbb{R}^n)\right\|\leq C\|b\|_{BMO_{\Delta_N}(\mathbb{R}^n)},$$

where $1<p<\frac{n}{\alpha}$ and$\frac{1}{q}=\frac{1}{p}-\frac{\alpha}{n}$.

Proof. Suppose b is in $BMO_{\Delta_N}(\mathbb{R}^n)$. Then according to [65], we have that $b_{+,e} \in BMO_{\Delta_N}(\mathbb{R}^n)$ and $b_{-,e} \in BMO_{\Delta_N}(\mathbb{R}^n)$, and moreover,

$$\|b\|_{BMO_{\Delta_N}(\mathbb{R}^n)} \approx \|b_{+,e}\|_{BMO_{\Delta_N}(\mathbb{R}^n)} \|b_{-,e}\|_{BMO_{\Delta_N}(\mathbb{R}^n)}.$$

For every $f \in L^p(\mathbb{R}^n)$, we have

$$\left\|\left[b, \Delta_N^{-\alpha/2}\right](f)\right\|_{L^q(\mathbb{R}^n)}^q = \int_{\mathbb{R}_+^n} \left[b, \Delta_N^{-\alpha/2}\right](f)(x)^q \, dx + \int_{\mathbb{R}_-^n} \left[b, \Delta_N^{-\alpha/2}\right](f)(x)^q \, dx =: I + II.$$

For the term I, note that when $x \in \mathbb{R}_+^n$, we have

$$\begin{aligned}\left[b, \Delta_N^{-\alpha/2}\right](f)(x) &= b(x)\Delta_N^{-\alpha/2}(f)(x) - \Delta_N^{-\alpha/2}(bf)(x) \\ &= b_{+,e}(x)\Delta^{-\alpha/2}(f_{+,e})(x) - \Delta^{-\alpha/2}(b_{+,e}f_{+,e})(x) \\ &= \left[b_{+,e}, \Delta^{-\alpha/2}\right](f_{+,e})(x),\end{aligned}$$

which implies that

$$\begin{aligned} I = \int_{\mathbb{R}_+^n} [b, \Delta^{-\alpha/2}](f)(x)^q \, dx &= \int_{\mathbb{R}_+^n} [b_{+,e}, \Delta^{-\alpha/2}](f_{+,e})(x)^q \, dx \\ &\le \int_{\mathbb{R}^n} [b_{+,e}, \Delta^{-\alpha/2}](f_{+,e})(x)^q \, dx \\ &\le C\|b_{+,e}\|_{BMO(\mathbb{R}^n)}^q \|f_{+,e}\|_{L^p(\mathbb{R}^n)}^q. \end{aligned}$$

For the last estimate we use the result [3], which applies since we know from that $R_{N,l}$is a Calderón–Zygmund kernel. Similarly we can obtain that

$$II \le C\|b_{+,e}\|_{BMO(\mathbb{R}^n)}^q \|f_{+,e}\|_{L^p(\mathbb{R}^n)}^q.$$

Combining the estimates for I and II above, we obtain that

$$\begin{aligned}\left\|\left[b, \Delta_N^{-\alpha/2}\right](f)\right\|_{L^p(\mathbb{R}^n)}^q &\le C\|b_{+,e}\|_{BMO(\mathbb{R}^n)}^q \|f_{+,e}\|_{L^p(\mathbb{R}^n)}^q + C\|b_{-,e}\|_{BMO(\mathbb{R}^n)}^q \|f_{-,e}\|_{L^p(\mathbb{R}^n)}^q \\ & C\|b\|_{BMO_{\Delta_N}(\mathbb{R}^n)}^q \left(\|f_{+,e}\|_{L^p(\mathbb{R}^n)}^q + \|f_{-,e}\|_{L^p(\mathbb{R}^n)}^q\right) \\ & C\|b\|_{BMO_{\Delta_N}(\mathbb{R}^n)}^q \|f\|_{L^p(\mathbb{R}^n)}^q,\end{aligned}$$

which yields that$\left\|[b, R_{N,l}]: L^p(\mathbb{R}^n) \to L^q(\mathbb{R}^n)\right\| \le C\|b\|_{BMO_{\Delta_N}(\mathbb{R}^n)}$.

Corollary (6.2.29)[183]: Suppose that L is one of the operators $\Delta_{N_+}, \Delta_{N_-}$and Δ_N.

Then for $x, x', (x+\epsilon) \in \mathbb{R}_+^{(1+2\epsilon)}$(or $\in \mathbb{R}_-^{(1+2\epsilon)}$) with $|x - x'| \le \frac{1}{2}|\epsilon|$, we have

$$\left|(1+\epsilon)_{(1+\epsilon),L}(x, x+\epsilon) - (1+\epsilon)_{(1+\epsilon),L}(x', x+\epsilon)\right| \le C\frac{|x - x'|\sqrt{1+\epsilon}}{\left(\sqrt{1+\epsilon} + |\epsilon|\right)^{(3+2\epsilon)}}; \quad (84)$$

Symmetrically, for $x, (x+\epsilon), (x+\epsilon)' \in \mathbb{R}_+^{(1+2\epsilon)}$(or$\in \mathbb{R}_-^{(1+2\epsilon)}$) with $|(x+\epsilon) - (x+\epsilon)'| \le \frac{1}{2}|\epsilon|$, we have

$$\begin{aligned}&\left|(1+\epsilon)_{(1+\epsilon),L}(x, x+\epsilon) - (1+\epsilon)_{(1+\epsilon),L}(x', x+\epsilon)\right| \\ &\le C\frac{|(x+\epsilon) - (x+\epsilon)'|\sqrt{(1+\epsilon)}}{\left(\sqrt{(1+\epsilon)} + |\epsilon|\right)^{3+2\epsilon}}. \qquad (85)\end{aligned}$$

Proof: Suppose $x, (x+\epsilon) \in \mathbb{R}_+^{(1+2\epsilon)}$. Then for $i = 1, \ldots, 2\epsilon$, we have

$$\frac{\partial}{\partial x_i}(1+\epsilon)_{(1+\epsilon),\Delta_{N_+}}(x,x+\epsilon)$$

$$= -\frac{(x_i-(x+\epsilon)_i)}{2(1+\epsilon)}\frac{1}{(4\pi(1+\epsilon))^{\frac{(1+2\epsilon)}{2}}}e^{-\frac{|x'-(x+\epsilon)'|^2}{4(1+\epsilon)}}\left(e^{-\frac{|x_{(1+2\epsilon)}-(x+\epsilon)_{(1+2\epsilon)}|^2}{4(1+\epsilon)}}+e^{-\frac{|x_{(1+2\epsilon)}-(x+\epsilon)_{(1+2\epsilon)}|^2}{4(1+\epsilon)}}\right).$$

Moreover,

$$\frac{\partial}{\partial x_{(1+2\epsilon)}}(1+\epsilon)_{(1+\epsilon),\Delta_{N_+}}(x,x+\epsilon)$$

$$= -\frac{1}{(4\pi(1+\epsilon))^{(\frac{1+2\epsilon}{2})}}e^{-\frac{|x'-(x+\epsilon)'|^2}{4(1+\epsilon)}}\left(e^{-\frac{|x_{(1+2\epsilon)}-(x+\epsilon)_{(1+2\epsilon)}|^2}{4(1+\epsilon)}}\frac{\left(x_{(1+2\epsilon)}-(x+\epsilon)_{(1+2\epsilon)}\right)}{2(1+\epsilon)}+e^{-\frac{|x_{(1+2\epsilon)}+(x+\epsilon)_{(1+2\epsilon)}|^2}{4(1+\epsilon)}}\frac{\left(x_{(1+2\epsilon)}+(x+\epsilon)_{(1+2\epsilon)}\right)}{2(1+\epsilon)}\right).$$

Then we obtain that

$$\left|\nabla_x(1+\epsilon)_{(1+\epsilon),\Delta_{N_+}}(x,(x+\epsilon))\right|^2$$

$$= \sum_{i=1}^{2\epsilon}\left|\frac{\partial}{\partial x_i}(1+\epsilon)_{(1+\epsilon),\Delta_{N_+}}(x,(x+\epsilon))\right|^2+\left|\frac{\partial}{\partial x_{(1+2\epsilon)}}(1+\epsilon)_{(1+\epsilon),\Delta_{N_+}}(x,(x+\epsilon))\right|^2$$

$$\leq \sum_{i=1}^{2\epsilon}\frac{(x_i-(x+\epsilon)_i)^2}{4(1+\epsilon)^2}\frac{1}{(4\pi(1+\epsilon))^{(1+2\epsilon)}}e^{-\frac{|x'-(x+\epsilon)'|^2}{2(1+\epsilon)}}\left(e^{-\frac{|x_{(1+2\epsilon)}-(x+\epsilon)_{(1+2\epsilon)}|^2}{4(1+\epsilon)}}+e^{-\frac{|x_{(1+2\epsilon)}+(x+\epsilon)_{(1+2\epsilon)}|^2}{4(1+\epsilon)}}\right)^2$$

$$+2\frac{1}{(4\pi(1+\epsilon))^{(1+2\epsilon)}}e^{-\frac{|x'-(x+\epsilon)'|^2}{2(1+\epsilon)}}\left(e^{-\frac{|x_{(1+2\epsilon)}-(x+\epsilon)_{(1+2\epsilon)}|^2}{4(1+\epsilon)}}\frac{\left(x_{(1+2\epsilon)}-(x+\epsilon)_{(1+2\epsilon)}\right)}{2(1+\epsilon)}\right)^2$$

$$+2\frac{1}{(4\pi(1+\epsilon))^{(1+2\epsilon)}}e^{-\frac{|x'-(x+\epsilon)'|^2}{2(1+\epsilon)}}\left(e^{-\frac{|x_{(1+2\epsilon)}+(x+\epsilon)_{(1+2\epsilon)}|^2}{4(1+\epsilon)}}\frac{\left(x_{(1+2\epsilon)}+(x+\epsilon)_{(1+2\epsilon)}\right)}{2(1+\epsilon)}\right)^2$$

$$\leq C\sum_{i=1}^{1+2\epsilon}\frac{(x_i-(x+\epsilon)_i)^2}{(1+\epsilon)^2}\frac{1}{(4\pi(1+\epsilon))^{(1+2\epsilon)}}e^{-\frac{|\epsilon|^2}{2(1+\epsilon)}}$$

$$+2\frac{1}{(4\pi(1+\epsilon))^{(1+2\epsilon)}}e^{-\frac{|x'-(x+\epsilon)'|^2}{2(1+\epsilon)}}\left(e^{-\frac{|x_{(1+2\epsilon)}+(x+\epsilon)_{(1+2\epsilon)}|^2}{4(1+\epsilon)}}\frac{\left(x_{(1+2\epsilon)}+(x+\epsilon)_{(1+2\epsilon)}\right)}{2(1+\epsilon)}\right)^2$$

$$\leq C\frac{|\epsilon|^2}{(1+\epsilon)^2}\frac{1}{(4\pi(1+\epsilon))^{(1+2\epsilon)}}e^{-\frac{|\epsilon|^2}{2(1+\epsilon)}}$$
$$+2\frac{1}{(4\pi(1+\epsilon))^{(1+2\epsilon)}}e^{-\frac{|\epsilon|^2}{4(1+\epsilon)}}\left(e^{-\frac{\left|x_{(1+2\epsilon)}+(x+\epsilon)_{(1+2\epsilon)}\right|^2}{8(1+\epsilon)}}\frac{\left(x_{(1+2\epsilon)}+(x+\epsilon)_{(1+2\epsilon)}\right)}{2(1+\epsilon)}\right)^2$$
$$\leq C\frac{(1+\epsilon)}{(1+\epsilon+\epsilon^2)}.$$

Hence, it is easy to verify that

$$\left|\nabla_x(1+\epsilon)_{(1+\epsilon),\Delta_{N_+}}(x,(x+\epsilon))\right|\leq C\frac{\sqrt{(1+\epsilon)}}{\left(\sqrt{(1+\epsilon)}+|\epsilon|\right)^{(3+2\epsilon)}}$$

and similarly we can obtain that

$$\left|\nabla_{(x+\epsilon)}(1+\epsilon)_{(1+\epsilon),\Delta_{N_+}}(x,(x+\epsilon))\right|\leq C\frac{\sqrt{(1+\epsilon)}}{\left(\sqrt{(1+\epsilon)}+|\epsilon|\right)^{(3+2\epsilon)}},$$

which implies that

$$\left|(1+\epsilon)_{(1+\epsilon),\Delta_{N_+}}(x,(x+\epsilon))-(1+\epsilon)_{(1+\epsilon),\Delta_{N_+}}(x',(x+\epsilon))\right|$$
$$\leq C\frac{|x'-x|\sqrt{(1+\epsilon)}}{\left(\sqrt{(1+\epsilon)}+|\epsilon|\right)^{(4+2\epsilon)}}$$

for $x,x',(x+\epsilon)\in\mathbb{R}_+^{(1+2\epsilon)}$ with $|x-x'|\leq\frac{1}{2}|\epsilon|$, and

$$\left|(1+\epsilon)_{(1+\epsilon),\Delta_{N_+}}(x,(x+\epsilon))-(1+\epsilon)_{(1+\epsilon),\Delta_{N_+}}(x,(x+\epsilon)')\right|$$
$$\leq C\frac{|(x+\epsilon)-(x+\epsilon)'|\sqrt{(x+\epsilon)}}{\left(\sqrt{x+\epsilon}+|\epsilon|\right)^{(4+2\epsilon)}}$$

for $x,x',(x+\epsilon)\in\mathbb{R}_+^{(1+2\epsilon)}$ with $|(x+\epsilon)-(x+\epsilon)'|\leq\frac{1}{2}|\epsilon|$.

Corollary (6.2.30)[183]: Denote by $R_{N,j}(x,(x+\epsilon))$the kernel of the j-th Riesz transform $\frac{\partial}{\partial_{x_j}}\Delta_N^{-\frac{1}{2}}$ of Δ_N. Then for $1\leq\ j\leq 2\epsilon$ and for $x,(x+\epsilon)\in\mathbb{R}_+^{(1+2\epsilon)}$we have:

$R_{N,j}(x,x+\epsilon)$
$$=-C_{(1+2\epsilon)}\left(\frac{x_i-(x+\epsilon)_i}{|\epsilon|^{(2+2\epsilon)}}+\frac{x_i-(x+\epsilon)_i}{\left(|x'-(x+\epsilon)'|^2+\left|x_{(1+2\epsilon)}-(x+\epsilon)_{1+2\epsilon)}\right|^2\right)^{(1+\epsilon)}}\right)$$

and

$$R_{N,(1+2\epsilon)}(x,(x+\epsilon))$$
$$=-C_{(1+2\epsilon)}\left(\frac{x_i-(x+\epsilon)_i}{|\epsilon|^{(2+2\epsilon)}}\right.$$
$$\left.+\frac{x_{(1+2\epsilon)}+(x+\epsilon)_{(1+2\epsilon)}}{\left(|x'-(x+\epsilon)'|^2+\left|x_{(1+2\epsilon)}+(x+\epsilon)_{(1+2\epsilon)}\right|^2\right)^{(1+\epsilon)}}\right),$$

where $C_{(1+2\epsilon)} = \frac{\Gamma(1+\epsilon)}{\pi^{(1+\epsilon)}}$. Similar expressions also hold for $R_{N,j}(x, x+\epsilon), j = 1, \dots, (1+2\epsilon)$, when $x, (x+\epsilon) \in \mathbb{R}_{-}^{(1+2\epsilon)}$.

Proof: Working from the definition of the square root of Δ_N, i.e.,

$$\Delta_N^{-\frac{1}{2}} = \frac{1}{\Gamma\left(\frac{1}{2}\right)} \int_0^\infty e^{-(1+\epsilon)\Delta_N} \frac{d(1+\epsilon)}{\sqrt{(1+\epsilon)}},$$

we have that for $1 \le j \le 2\epsilon$:

$$R_{N,j}(x,(x+\epsilon)) = \frac{1}{\Gamma\left(\frac{1}{2}\right)} \frac{\partial}{\partial_{x_j}} \int_0^\infty (1+\epsilon)_{(1+\epsilon),\Delta_N}(x,(x+\epsilon)) \frac{d(1+\epsilon)}{\sqrt{(1+\epsilon)}}$$

$$= \frac{1}{\Gamma\left(\frac{1}{2}\right)} \frac{\partial}{\partial_{x_j}} \left(\int_0^\infty \frac{1}{(4\pi(1+\epsilon))^{\frac{(1+2\epsilon)}{2}}} e^{-\frac{|\epsilon|^2}{4(1+\epsilon)}} \frac{d(1+\epsilon)}{\sqrt{(1+\epsilon)}} \right.$$
$$\left. + \int_0^\infty \frac{1}{(4\pi(1+\epsilon))^{\frac{(1+2\epsilon)}{2}}} e^{-\frac{|x'-(x+\epsilon)'|^2}{4(1+\epsilon)}} e^{-\frac{\left|x_{(1+2\epsilon)}-(x+\epsilon)_{(1+2\epsilon)}\right|^2}{4(1+\epsilon)}} \frac{d(1+\epsilon)}{\sqrt{(1+\epsilon)}} \right)$$

$$= -\frac{\Gamma(1+\epsilon)}{(\pi)^{(1+\epsilon)}} \left(\frac{x_i - (x+\epsilon)_i}{|\epsilon|^{(2+2\epsilon)}} + \frac{x_i - (x+\epsilon)_i}{\left(|x' - (x+\epsilon)'|^2 + \left|x_{(1+2\epsilon)} + (x+\epsilon)_{(1+2\epsilon)}\right|^2\right)^{(1+\epsilon)}} \right).$$

For $j = (1+2\epsilon)$ and for $x, (x+\epsilon) \in \mathbb{R}_{+}^{(1+2\epsilon)}$ we again observe:

$$R_{N,(1+2\epsilon)}(x,(x+\epsilon)) = \frac{\sqrt{\pi}}{2} \frac{\partial}{\partial_{x_{(1+2\epsilon)}}} \int_0^\infty (1+\epsilon)_{(1+\epsilon),\Delta_N}(x,(x+\epsilon)) \frac{d(1+\epsilon)}{\sqrt{(1+\epsilon)}}$$

$$= -\frac{\Gamma(1+\epsilon)}{(\pi)^{(1+\epsilon)}} \left(\frac{x_{(1+2\epsilon)} - (x+\epsilon)_{(1+2\epsilon)}}{|\epsilon|^{(2+2\epsilon)}} \right.$$
$$\left. + \frac{x_{(1+2\epsilon)} + (x+\epsilon)_{(1+2\epsilon)}}{\left(|x' - (x+\epsilon)'|^2 + \left|x_{(1+2\epsilon)} + (x+\epsilon)_{(1+2\epsilon)}\right|^2\right)^{(1+\epsilon)}} \right).$$

We next make the observation that kernels $R_{N,j}(x,(x+\epsilon))$ are Calderón–Zygmund kernels.

Corollary (6.2.31)[183]: Denote by $R_N(x,(x+\epsilon))$ kernel of the vector of Riesz transforms $\nabla\Delta_N^{-\frac{1}{2}}$. Then:

$$\begin{aligned} R_N(x,(x+\epsilon)) &= \left(R_{N,1}(x,(x+\epsilon)), \dots, R_{N,(1+2\epsilon)}(x,(x+\epsilon))\right) \\ &\times H\left(x_{(1+2\epsilon)}, (x+\epsilon)_{(1+2\epsilon)}\right), \end{aligned} \tag{86}$$

with $H(1+\epsilon)$ the Heaviside function defined in (64). Moreover, we have that

$$|R_N(x,(x+\epsilon))| \le C_{(1+2\epsilon)} \frac{1}{|\epsilon|^{(1+2\epsilon)}},$$

and

$$|R_N(x,(x+\epsilon)) - R_N(x_0,(x+\epsilon))| + |R_N(x,(x+\epsilon)) - R_N((x+\epsilon),x_0)| \le C\frac{|x-x_0|}{|\epsilon|^{(2+2\epsilon)}}$$

for $x, x_0, (x+\epsilon) \in \mathbb{R}_+^{(1+2\epsilon)}$ (or $x, x_0, (x+\epsilon) \in \mathbb{R}_-^{(1+2\epsilon)}$) with $|x-x_0| \le \frac{1}{2}|\epsilon|$.

Proof: We first claim that for $j = 1, \ldots, (1+2\epsilon)$, and $x, (1+\epsilon) \in \mathbb{R}_+^{(1+2\epsilon)}$(or $x, (x+\epsilon) \in \mathbb{R}_-^{(1+2\epsilon)}$)

$$\left|R_{N,j}(x,(x+\epsilon))\right| \le C_{(1+2\epsilon)}\frac{1}{|\epsilon|^{(1+2\epsilon)}}.$$

In fact, from Corollary (6.2.30), it is direct that for $1 \le j \le 2\epsilon$,

$$\frac{\left|x_j-(x+\epsilon)_j\right|}{\left(|x'-(x+\epsilon)'|^2+\left|x_{(1+2\epsilon)}+(x+\epsilon)_{(1+2\epsilon)}\right|^2\right)^{(1+\epsilon)}} \le \frac{\left|x_j-(x+\epsilon)_j\right|}{\left(|x'-(x+\epsilon)'|^2+\left|x_{(1+2\epsilon)}+(x+\epsilon)_{(1+2\epsilon)}\right|^2\right)^{(1+\epsilon)}} \le \frac{1}{|\epsilon|^{(1+2\epsilon)}}$$

and for $j = (1+2\epsilon)$,

$$\frac{\left|x_{(1+2\epsilon)}-(x+\epsilon)_{(1+2\epsilon)}\right|}{\left(|x'-(x+\epsilon)'|^2+\left|x_{(1+2\epsilon)}+(x+\epsilon)_{(1+2\epsilon)}\right|^2\right)^{(1+\epsilon)}} \le \frac{\left|x_j-(x+\epsilon)_j\right|}{\left(|x'-(x+\epsilon)'|^2+\left|x_{(1+2\epsilon)}+(x+\epsilon)_{(1+2\epsilon)}\right|^2\right)^{(1+\epsilon)}} \le \frac{1}{|\epsilon|^{(1+2\epsilon)}},$$

where we use the fact that $x, x_0, (x+\epsilon) \in \mathbb{R}_+^{(1+2\epsilon)}$ (or $x, x_0, (x+\epsilon) \in \mathbb{R}_-^{(1+2\epsilon)}$) and hence $x_j-(x+\epsilon)_j > \left|x_j-(x+\epsilon)_j\right|$ for $1 \le j \le 1+2\epsilon$.

Similarly, by considering the estimates for the terms $\frac{\partial}{\partial_{x_j}}R_{N,j}(x,(x+\epsilon))$ and $\frac{\partial}{\partial_{x_j}}R_{N,j}(x,(x+\epsilon))$, we obtain that

$$\left|R_{N,j}(x,(x+\epsilon)) - R_{N,j}(x_0,(x+\epsilon))\right| + \left|R_{N,j}\big((x+\epsilon),x\big) - R_{N,j}((x+\epsilon),x_0)\right| \le C\frac{|x-x_0|}{|\epsilon|^{(2+2\epsilon)}}$$

For $x, x_0, (x+\epsilon) \in \mathbb{R}_+^{(2+2\epsilon)}$ (or $x, x_0, (x+\epsilon) \in \mathbb{R}_-^{(2+2\epsilon)}$) with $|x-x_0| \le \frac{1}{2}|\epsilon|$.

Corollary (6.2.32)[183]: Denote by $K_N(x,(x+\epsilon))$the kernel of the fractional operator $\Delta_N^{-(1+\epsilon)/2}$. Then $x, (x+\epsilon) \in \mathbb{R}_+^{(1+2\epsilon)}$ we have:

$$K_N(x,(x+\epsilon)) = K(x,(x+\epsilon)) + \widetilde{K}_N(x,(x+\epsilon))$$

with

$$\widetilde{K}_N(x,(x+\epsilon)) \coloneqq C_{(1+2\epsilon),(1+\epsilon)} = \frac{1}{\left(|x'-(x+\epsilon)'|^2+\left|x_{(1+2\epsilon)}+(x+\epsilon)_{(1+2\epsilon)}\right|^2\right)^{\frac{\epsilon}{2}}}.$$

Similar expressions for $K_N(x,(x+\epsilon))$ when $x, (x+\epsilon) \in \mathbb{R}_-^{(1+2\epsilon)}$ also hold.

Proof: For $x, (x+\epsilon) \in \mathbb{R}_+^{(1+2\epsilon)}$, working from the fraction of the square root of $\Delta_N$we have that:

$$K_N(x,(x+\epsilon)) = \frac{1}{\Gamma((1+\epsilon)\epsilon/2)}\int_0^\infty (1+\epsilon)_{(1+\epsilon),\Delta_N}(x,(x+\epsilon))\frac{d(1+\epsilon)}{(1+\epsilon)^{-\frac{\epsilon}{2}}}$$

$$= \frac{1}{\Gamma((1+\epsilon)/2)}\int_0^\infty \frac{1}{(4\pi(1+\epsilon))^{\frac{(1+2\epsilon)}{2}}} e^{-\frac{|\epsilon|^2}{4(1+\epsilon)}} \frac{d(1+\epsilon)}{(1+\epsilon)^{-\frac{\epsilon}{2}}}$$

$$+ \frac{1}{\Gamma((1+\epsilon)/2)}\int_0^\infty \frac{1}{(4\pi(1+\epsilon))^{\frac{(1+2\epsilon)}{2}}} e^{-\frac{|x'-(x+\epsilon)'|^2}{4(1+\epsilon)}} e^{-\frac{|x_{(1+2\epsilon)}-(x+\epsilon)_{(1+2\epsilon)}|^2}{4(1+\epsilon)}} \frac{d(1+\epsilon)}{(1+\epsilon)^{-\frac{\epsilon}{2}}}$$

$$= C_{(1+2\epsilon),(1+\epsilon)}\left(\frac{1}{|\epsilon|^\epsilon} + \frac{1}{\left(|x'-(x+\epsilon)'|^2 + \left|x_{(1+2\epsilon)} + (x+\epsilon)_{(1+2\epsilon)}\right|^2\right)^{\frac{\epsilon}{2}}}\right)$$

$$= K(x,(x+\epsilon)) + \widetilde{K}_N(x,(x+\epsilon)),$$

where we set

$$\widetilde{K}_N(x,(x+\epsilon)) = C_{(1+2\epsilon),(1+\epsilon)} \frac{1}{\left(|x'-(x+\epsilon)'|^2 + \left|x_{(1+2\epsilon)} + (x+\epsilon)_{(1+2\epsilon)}\right|^2\right)^{\frac{\epsilon}{2}}}.$$

Corollary (6.2.33)[183]: Let all the notation be as above. Then,

$$H^1_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right) = H^1_{\Delta_N,max}\left(\mathbb{R}^{(1+2\epsilon)}\right) = H^1_{\Delta_N,*}\left(\mathbb{R}^{(1+2\epsilon)}\right) = H^1_{\Delta_N,Riesz}\left(\mathbb{R}^{(1+2\epsilon)}\right) = H^1_{\Delta_N,atom}\left(\mathbb{R}^{(1+2\epsilon)}\right)$$

and they have equivalent norms

$$\|f\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})} \approx \|f\|_{H^1_{\Delta_N,max}(\mathbb{R}^{(1+2\epsilon)})}\|f\|_{H^1_{\Delta_N,Riesz}(\mathbb{R}^{(1+2\epsilon)})}\|f\|_{H^1_{\Delta_N,atom}(\mathbb{R}^{(1+2\epsilon)})} \approx \left\|f_{+,e}\right\|_{H^1(\mathbb{R}^{(1+2\epsilon)})} + \left\|f_{-,e}\right\|_{H^1(\mathbb{R}^{(1+2\epsilon)})}.$$

Namely, $f \in H^1_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)$if and only if $f_{+,e} \in H^1\left(\mathbb{R}^{(1+2\epsilon)}\right)$and $f_{-,e} \in H^1\left(\mathbb{R}^{(1+2\epsilon)}\right)$.

Proof: We recall that the Hardy space associated with Δ_N is defined as the set of functions $\left\{f \in L^1\left(\mathbb{R}^{(1+2\epsilon)}\right): \left\|S_{\Delta_N}(f)\right\|_{L^1(\mathbb{R}^{(1+2\epsilon)})} < \infty\right\}$ in the norm of $\|f\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})} = \left\|S_{\Delta_N}(f)\right\|_{L^1(\mathbb{R}^{(1+2\epsilon)})}$ where $S_{\Delta_N}(f)(x) = \left(\int_0^\infty \int_{|\epsilon|>-\frac{1}{2}} \left|Q_{(1+\epsilon)^2} f((x+\epsilon))\right|^2 \frac{d(x+\epsilon)d(1+\epsilon)}{(1+\epsilon)^{(2+2\epsilon)}}\right)^{\frac{1}{2}}$, and $Q_{(1+\epsilon)^2} = (1+\epsilon)^2\Delta_N \exp(-(1+\epsilon)^2\Delta_N)$.

We now consider the operator $Q_{(1+\epsilon)} = (1+\epsilon)\Delta_N \exp(-(1+\epsilon)\Delta_N) = -(1+\epsilon)\frac{d}{d(1+\epsilon)}\exp(-(1+\epsilon)\Delta_N)$ for any $\epsilon > -1$ (see [71]). Then we have

$$Q_{(1+\epsilon)^2} f(x) = (1+\epsilon)^2\Delta_N \exp(-(1+\epsilon)^2\Delta_N) f(x)$$

$$= \int_{\mathbb{R}^{(1+2\epsilon)}} -\frac{(1+\epsilon)}{2}\frac{\partial}{\partial(1+\epsilon)}(1+\epsilon)_{(1+\epsilon)^2,\Delta_N}\left(x,(x+\epsilon)\right) f\left((x+\epsilon)\right) d(x+\epsilon).$$

From the definition of $(1+\epsilon)_{(1+\epsilon)\epsilon,\Delta_N}(x,(x+\epsilon))$, see(2.4), we have that for any $x \in \mathbb{R}^{(1+2\epsilon)}_+$,

$$(1+\epsilon)^2\Delta_N \exp(-(1+\epsilon)^2\Delta_N) f(x)$$
$$= \int_{\mathbb{R}_+^{(1+2\epsilon)}} -\frac{(1+\epsilon)}{2}\frac{\partial}{\partial(1+\epsilon)}(1+\epsilon)_{(1+\epsilon)^2,\Delta_N}\big(x,(x+\epsilon)\big)f_+\big((x+\epsilon)\big)d(x+\epsilon)$$
$$= \int_{\mathbb{R}^{(1+2\epsilon)}} -\frac{(1+\epsilon)}{2}\frac{\partial}{\partial(1+\epsilon)}(1+\epsilon)_{(1+\epsilon)^2}\big(x,(x+\epsilon)\big)f_{+,e}\big((x+\epsilon)\big)d(x+\epsilon)$$
$$= (1+\epsilon)^2\Delta \exp(-(1+\epsilon)^2\Delta_N) f_{+,e}(x).$$

Similarly, for any $x \in \mathbb{R}_-^{(1+2\epsilon)}$, we have

$$(1+\epsilon)^2\Delta_N \exp(-(1+\epsilon)^2\Delta_N) f(x) = (1+\epsilon)^2\Delta \exp(-(1+\epsilon)^2\Delta) f_{-,e}(x).$$

Moreover, by a change of variable,

$$(1+\epsilon)^2\Delta_N \exp(-(1+\epsilon)^2\Delta_N) f(x)$$
$$= -(1+\epsilon)^2\Delta \exp(-(1+\epsilon)^2\Delta) f_{+,e}(\tilde{x}) \text{ for any } \epsilon > -1,$$
$$x \in \mathbb{R}_+^{(1+2\epsilon)}; \tag{87}$$
$$(1+\epsilon)^2\Delta_N \exp(-(1+\epsilon)^2\Delta_N) f(x) = -(1+\epsilon)^2\Delta \exp(-(1+\epsilon)^2\Delta) f_{-,e}(\tilde{x}) \text{ for any } \epsilon > -1, x \in \mathbb{R}_-^{(1+2\epsilon)}$$

Then from(87) we have

$$S_{\Delta_N}(f)(x)^2 = \int_0^\infty \int_{|\epsilon|>-\frac{1}{2},(x+\epsilon)\in\mathbb{R}_+^{(1+2\epsilon)}} |(1+\epsilon)^2\Delta_N \exp(-(1+\epsilon)^2\Delta_N) f(x+\epsilon)|^2 \frac{d(x+\epsilon)d(1+\epsilon)}{(1+\epsilon)^{(1+2\epsilon)}}$$
$$+ \int_0^\infty \int_{|\epsilon|>-\frac{1}{2},(x+\epsilon)\in\mathbb{R}_+^{(1+2\epsilon)}} |(1+\epsilon)^2\Delta_N \exp(-(1+\epsilon)^2\Delta_N) f(x+\epsilon)|^2 \frac{d(x+\epsilon)d(1+\epsilon)}{(1+\epsilon)^{(1+2\epsilon)}}$$
$$= \int_0^\infty \int_{|\epsilon|>-\frac{1}{2},(x+\epsilon)\in\mathbb{R}_+^{(1+2\epsilon)}} \left|(1+\epsilon)^2\Delta \exp(-(1+\epsilon)^2\Delta) f_{+,e}(x+\epsilon)\right|^2 \frac{d(x+\epsilon)d(1+\epsilon)}{(1+\epsilon)^{(1+2\epsilon)}}$$
$$+ \int_0^\infty \int_{|\epsilon|>-\frac{1}{2},(x+\epsilon)\in\mathbb{R}_-^{(1+2\epsilon)}} \left|(1+\epsilon)^2\Delta \exp(-(1+\epsilon)^2\Delta) f_{-,e}((x+\epsilon))\right|^2 \frac{d(x+\epsilon)d(1+\epsilon)}{(1+\epsilon)^{(1+2\epsilon)}}$$
$$= \frac{1}{2}\left(\int_0^\infty \int_{|\epsilon|>-\frac{1}{2}} \left|(1+\epsilon)^2\Delta \exp(-(1+\epsilon)^2\Delta) f_{+,e}((x+\epsilon))\right|^2 \frac{d(x+\epsilon)d(1+\epsilon)}{(1+\epsilon)^{(1+2\epsilon)}}\right.$$
$$\left.+ \int_0^\infty \int_{|\epsilon|>-\frac{1}{2}} \left|(1+\epsilon)^2\Delta \exp(-(1+\epsilon)^2\Delta) f_{-,e}((x+\epsilon))\right|^2 \frac{d(x+\epsilon)d(1+\epsilon)}{(1+\epsilon)^{(1+2\epsilon)}}\right),$$

which implies that $S_{\Delta_N}(f)(x) \leq \frac{\sqrt{2}}{2}\Big(S\big(f_{+,e}\big)(x) + S\big(f_{-,e}\big)((x+\epsilon))\Big)$. Conversely,

$$S(f_{+,e})(x)^2 = \int_0^\infty \int_{|\epsilon|>-\frac{1}{2}} \left|(1+\epsilon)^2\Delta \exp(-(1+\epsilon)^2\Delta) f_{+,e}((x+\epsilon))\right|^2 \frac{d(x+\epsilon)d(1+\epsilon)}{(1+\epsilon)^{(1+2\epsilon)}}$$

$$= 2\int_0^\infty \int_{|\epsilon|>-\frac{1}{2},(x+\epsilon)\in\mathbb{R}_+^{(1+2\epsilon)}} \left|(1+\epsilon)^2\Delta \exp(-(1+\epsilon)^2\Delta) f_{+,e}((x+\epsilon))\right|^2 \frac{d(x+\epsilon)d(1+\epsilon)}{(1+\epsilon)^{(1+2\epsilon)}}$$

$$\le 2S_{\Delta_N}(f)(x)^2.$$

Similarly we have $S(f_{-,e})(x)^2 \le 2S_{\Delta_N}(f)(x)^2$. Hence, we obtain that $S(f_{+,e})(x) + S(f_{-,e})(x) \le 2\sqrt{2}S_{\Delta_N}(f)(x)$. As a consequence, we have

$$\|f\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})} = \int_{\mathbb{R}^{(1+2\epsilon)}} \left|S_{\Delta_N}(f)(x)\right| dx \tag{88}$$

$$\approx \int_{\mathbb{R}^{(1+2\epsilon)}} \left|S(f_{+,e})(x)\right| dx + \int_{\mathbb{R}^{(1+2\epsilon)}} \left|S(f_{-,e})(x)\right| dx$$

$$= \left\|f_{+,e}\right\|_{H^1(\mathbb{R}^{(1+2\epsilon)})} + \left\|f_{-,e}\right\|_{H^1(\mathbb{R}^{(1+2\epsilon)})}.$$

Next we turn to $H^1_{\Delta_D,max}(\mathbb{R}^{(1+2\epsilon)})$. From (63) we can see that for any $\epsilon > -1$ and $x \in \mathbb{R}_+^{(1+2\epsilon)}$,

$$\exp(-(1+\epsilon)^2\Delta_N) f(x) = \int_{\mathbb{R}^{(1+2\epsilon)}} (1+\epsilon)_{(1+\epsilon)^2,\Delta_N}(x,(x+\epsilon))f((x+\epsilon))d(x+\epsilon)$$

$$= \int_{\mathbb{R}_+^{(1+2\epsilon)}} (1+\epsilon)_{(1+\epsilon)^2,\Delta_N}(x,(x+\epsilon))f_+((x+\epsilon))d(x+\epsilon).$$

$$= \int_{\mathbb{R}^{(1+2\epsilon)}} (1+\epsilon)_{(1+\epsilon)^2}(x,(x+\epsilon))f_{+,e}((x+\epsilon))d(x+\epsilon) = \exp(-(1+\epsilon)^2\Delta) f_{+,e}(x).$$

Similarly, $\exp(-(1+\epsilon)^2\Delta_N) f(x) = \exp(-(1+\epsilon)^2\Delta) f_{+,e}(x)$for any $\epsilon > -1$ and $x \in \mathbb{R}_-^{(1+2\epsilon)}$. Thus,

$$\sup_{\epsilon>-1}|\exp(-(1+\epsilon)^2\Delta_N) f(x)| = \sup_{\epsilon>-1}\left|\exp(-(1+\epsilon)^2\Delta) f_{+,e}(x)\right| \text{ for any, } x \in \mathbb{R}_+^{(1+2\epsilon)}; \tag{89}$$

$$\sup_{\epsilon>-1}|\exp(-(1+\epsilon)^2\Delta_N) f(x)| = \sup_{\epsilon>-1}\left|\exp(-(1+\epsilon)^2\Delta) f_{-,e}(x)\right| \text{ for any } x \in \mathbb{R}_-^{(1+2\epsilon)}.$$

Again, by a change of variable, we have that

$$\exp(-(1+\epsilon)^2\Delta_N) f(x) = -\exp(-(1+\epsilon)^2\Delta) f_{+,e}(\tilde{x}) \text{ for any } \epsilon > -1, x \in \mathbb{R}_+^{(1+2\epsilon)}; \tag{90}$$

$$\exp(-(1+\epsilon)^2\Delta_N) f(x) = -\exp(-(1+\epsilon)^2\Delta) f_{-,e}(\tilde{x})\text{for any } \epsilon > -1, x \in \mathbb{R}_-^{(1+2\epsilon)}.$$

Then, for any$f \in H^1_{\Delta_N,max}(\mathbb{R}^{(1+2\epsilon)})$, from(89) and(90) we can obtain that

$$\|f\|_{H^1_{\Delta_N,max}(\mathbb{R}^{(1+2\epsilon)})} = \int_{\mathbb{R}_+^{(1+2\epsilon)}} \left|f^+_{\Delta_N}(x)\right| dx + \int_{\mathbb{R}_-^{(1+2\epsilon)}} \left|f^+_{\Delta_N}(x)\right| dx \tag{91}$$

$$= \int_{\mathbb{R}_+^{(1+2\epsilon)}} \sup_{\epsilon>-1} |\exp(-(1+\epsilon)^2\Delta_N) f(x)|\, dx + \int_{\mathbb{R}_-^{(1+2\epsilon)}} \sup_{\epsilon>-1} |\exp(-(1+\epsilon)^2\Delta_N) f(x)|\, dx$$

$$= \int_{\mathbb{R}_+^{(1+2\epsilon)}} \sup_{\epsilon>-1} \left|\exp(-(1+\epsilon)^2\Delta) f_{+,e}(x)\right| dx + \int_{\mathbb{R}_-^{(1+2\epsilon)}} \sup_{\epsilon>-1} \left|\exp(-(1+\epsilon)^2\Delta) f_{-,e}(x)\right| dx$$

$$= \frac{1}{2}\left(\int_{\mathbb{R}^{(1+2\epsilon)}} \sup_{\epsilon>-1} \left|\exp(-(1+\epsilon)^2\Delta) f_{+,e}(x)\right| dx + \int_{\mathbb{R}^{(1+2\epsilon)}} \sup_{\epsilon>-1} \left|\exp(-(1+\epsilon)^2\Delta) f_{-,e}(x)\right| dx \right)$$

$$= \frac{1}{2}\left(\left\|\left(f_{+,e}\right)^+\right\|_{L^1\left(\mathbb{R}^{(1+2\epsilon)}\right)} + \left\|\left(f_{-,e}\right)^+\right\|_{L^1\left(\mathbb{R}^{(1+2\epsilon)}\right)} \right)$$

$$= \frac{1}{2}\left(\left\|f_{+,e}\right\|_{H^1\left(\mathbb{R}^{(1+2\epsilon)}\right)} + \left\|f_{-,e}\right\|_{H^1\left(\mathbb{R}^{(1+2\epsilon)}\right)} \right),$$

Where $f^+(x) = \sup_{\epsilon>-1}\left|(1+\epsilon)_{(1+\epsilon)^2} * f(x)\right|$ is the classical maximal function as defined. Thus (91) yields that $f \in H^1_{\Delta_N,max}\left(\mathbb{R}^{(1+2\epsilon)}\right)$ if and only if $f_{+e} \in H^1\left(\mathbb{R}^{(1+2\epsilon)}\right)$ and $f_{-,e} \in H^1\left(\mathbb{R}^{(1+2\epsilon)}\right)$.

We now consider the Hardy space $H^1_{\Delta_N,*}\left(\mathbb{R}^{(1+2\epsilon)}\right)$ via the non-tangential maximal function. Note that

$$f^*_{\Delta_N}(x) = \sup_{|\epsilon|>-\frac{1}{2}} |\exp(-(1+\epsilon)^2\Delta_N) f((x+\epsilon))|$$

$$\le \sup_{|\epsilon|>-\frac{1}{2}, (x+\epsilon)\in\mathbb{R}_+^{(1+2\epsilon)}} |\exp(-(1+\epsilon)^2\Delta_N) f((x+\epsilon))| + \sup_{|\epsilon|<(1+\epsilon), (x+\epsilon)\in\mathbb{R}_-^{(1+2\epsilon)}} |\exp(-(1+\epsilon)^2\Delta_N) f((x+\epsilon))|$$

$$\le \sup_{|\epsilon|>-\frac{1}{2}, (x+\epsilon)\in\mathbb{R}_+^{(1+2\epsilon)}} \left|\exp(-(1+\epsilon)^2\Delta) f_{+,e}((x+\epsilon))\right| + \sup_{|\epsilon|>-\frac{1}{2}, (x+\epsilon)\in\mathbb{R}_-^{(1+2\epsilon)}} \left|\exp(-(1+\epsilon)^2\Delta) f_{-,e}((x+\epsilon))\right|$$

$$\le \sup_{|\epsilon|>-\frac{1}{2}} \left|\exp(-(1+\epsilon)^2\Delta) f_{+,e}((x+\epsilon))\right| + \sup_{|\epsilon|>-\frac{1}{2}} \left|\exp(-(1+\epsilon)^2\Delta) f_{-,e}((x+\epsilon))\right|$$

$$= \left(f_{+,e}\right)^*(x) + \left(f_{-,e}\right)^*(x),$$

where $f^*(x) = \sup_{|\epsilon|>-\frac{1}{2}} \left|(1+\epsilon)_{(1+\epsilon)^2} * f((x+\epsilon))\right|$ is the classical non-tangential maximal function. Hence $\left\|f^*_{\Delta_N}(x)\right\|_{L^1\left(\mathbb{R}^{(1+2\epsilon)}\right)} \le \left\|\left(f_{+,e}\right)^*\right\|_{L^1\left(\mathbb{R}^{(1+2\epsilon)}\right)} + \left\|\left(f_{-,e}\right)^*\right\|_{L^1\left(\mathbb{R}^{(1+2\epsilon)}\right)}$.

Moreover, we have

$$\left(f_{+,e}\right)^*(x) = \sup_{|\epsilon|>-\frac{1}{2}} \left|\exp(-(1+\epsilon)^2\Delta) f_{+,e}((x+\epsilon))\right|$$

$$\leq \sup_{|\epsilon|>-\frac{1}{2},(x+\epsilon)\in\mathbb{R}_+^{(1+2\epsilon)}} \left|\exp(-(1+\epsilon)^2\Delta) f_{+,e}((x+\epsilon))\right|$$

$$+ \sup_{|\epsilon|>-\frac{1}{2},(x+\epsilon)\in\mathbb{R}_-^{(1+2\epsilon)}} \left|\exp(-(1+\epsilon)^2\Delta) f_{+,e}((x+\epsilon))\right|$$

$$\leq 2 \sup_{|\epsilon|>-\frac{1}{2}} (1+\epsilon)^2 |\exp(-\Delta_N) f((x+\epsilon))| \leq 2 f^*_{\Delta_N}(x)$$

Thus, $\left\|(f_{+,e})^*\right\|_{L^1(\mathbb{R}^{(1+2\epsilon)})} \leq 2\left\|f^*_{\Delta_N}(x)\right\|_{L^1(\mathbb{R}^{(1+2\epsilon)})}$.

Similarly we obtain $\left\|(f_{-,e})^*\right\|_1 \leq 2\left\|f^*_{\Delta_N}(x)\right\|_{L^1(\mathbb{R}^{(1+2\epsilon)})}$. This implies that

$$\left\|f^*_{\Delta_N}(x)\right\|_1 \approx \left\|(f_{+,e})^*\right\|_{L^1(\mathbb{R}^{(1+2\epsilon)})} + \left\|(f_{-,e})^*\right\|_{L^1(\mathbb{R}^{(1+2\epsilon)})}. \tag{92}$$

Thus,(92) yields that $f \in H^1_{\Delta_N,*}(\mathbb{R}^{(1+2\epsilon)})$ if and only if $f_{+,e} \in H^1(\mathbb{R}^{(1+2\epsilon)})$ and $f_{-,e} \in H^1(\mathbb{R}^{(1+2\epsilon)})$.

As for the Riesz transform characterization of the Hardy space $H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$, it suffices to note that when $x \in \mathbb{R}_+^{(1+2\epsilon)}$,

$$\nabla\Delta_N^{-\frac{1}{2}} f(x) = \int_{\mathbb{R}^{(1+2\epsilon)}} K_N(x,(x+\epsilon)) f((x+\epsilon)) d(x+\epsilon)$$

$$= \int_{\mathbb{R}_+^{(1+2\epsilon)}} R_N(x,(x+\epsilon)) f_+((x+\epsilon)) d(x+\epsilon)$$

$$= \int_{\mathbb{R}^{(1+2\epsilon)}} R(x,(x+\epsilon)) f_{+,e}((x+\epsilon)) d(x+\epsilon)$$

$$= \nabla\Delta_N^{-\frac{1}{2}} f_{+,e}(x)$$

and that when $x \in \mathbb{R}_-^{(1+2\epsilon)}$,

$$\nabla\Delta_N^{-\frac{1}{2}} f(x) = \nabla\Delta_N^{-\frac{1}{2}} f_{-,e}(x).$$

Thus, $f \in H^1_{\Delta_N,Riesz}(\mathbb{R}^{(1+2\epsilon)})$if and only if $f_{+,e} \in H^1(\mathbb{R}^{(1+2\epsilon)})$and $f_{-,e} \in H^1(\mathbb{R}^{(1+2\epsilon)})$.

Corollary (6.2.34)[183]: The following are equivalent for a function b:

(i) $b \in BMO_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$;

(ii) There exists $b_0, b_1, \ldots, b_{(1+2\epsilon)} \in L^\infty(\mathbb{R}^{(1+2\epsilon)})$such that $b = b_0 + \sum_{j=1}^{(1+2\epsilon)} R^*_{N,j} b_j$, where $R^*_{N,j}$is the adjoint operator of $R_{N,j}$.

Proof. The proof is as in[74]. Let$B = \oplus_{j=0}^{(1+2\epsilon)} L^1(\mathbb{R}^{(1+2\epsilon)})$ and norm B by $\sum_{j=0}^{(1+2\epsilon)} \|f\|_{L^1(\mathbb{R}^{(1+2\epsilon)})}$. We have that $B^* = \oplus_{j=0}^{(1+2\epsilon)} L^\infty(\mathbb{R}^{(1+2\epsilon)})$. Let S be the subspace of B given by

$$S = \left\{\left(f, R_{N,1}f, \ldots, R_{N,(1+2\epsilon)}f\right) : f \in L^1(\mathbb{R}^{(1+2\epsilon)})\right\}.$$

We have that S is a closed subspace and that $f \to \left(f, R_{N,1}f, \ldots, R_{N,(1+2\epsilon)}f\right)$ is a isometry of $H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$ to S. Linear functionals on S and $H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$can be identified in an obvious way, hence any continuous linear functional on $H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$can be extended by Hahn-Banach to a continuous linear functional on B and can be identified with a vector of functions $(b_0, b_1, \ldots, b_{(1+2\epsilon)})$ with each $b_j \in L^\infty(\mathbb{R}^{(1+2\epsilon)})$.

We use this conclusion in the following way. Let ℓ be a continuous linear functional on $H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$. Then by Proposition (6.2.12) there is a function $b \in \mathrm{BMO}_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$ so that:

$$\int_{\mathbb{R}^{(1+2\epsilon)}} f(x)\overline{b(x)}dx = \ell(f).$$

However, by the discussion above, and by restricting the extended linear functional back to $H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$ we have for $(f, R_{N,1}f, \dots, R_{N,(1+2\epsilon)}f) = (f_0, \dots, f_{(1+2\epsilon)})$:

$$\ell(f) = \sum_{j=0}^{(1+2\epsilon)} \int_{\mathbb{R}^{(1+2\epsilon)}} f_j(x)\overline{b_j(x)}dx.$$

Using the definition of the $f_j = R_{N,j}f$ we see that:

$$\ell(f) = \int_{\mathbb{R}^{(1+2\epsilon)}} f(x)\overline{\left(b_0(x) + \sum_{j=1}^{(1+2\epsilon)} R^*_{N,j}b_j(x)\right)}dx.$$

This then gives the decomposition that any $b \in BMO_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$can be written as:

$$b = b_0 + \sum_{j=1}^{1+2\epsilon} R^*_{N,j}b_j$$

with $b_j \in L^\infty(\mathbb{R}^{(1+2\epsilon)})$.

For the converse, we simply observe that from our Corollary (6.2.33), we obtained that R_N maps $H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$to $L^1(\mathbb{R}^{(1+2\epsilon)})$. Hence, the boundedness of the Riesz transform R^*_N from $L^\infty(\mathbb{R}^{(1+2\epsilon)})$ to $BMO_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$ follows from duality of $H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$ with $\mathrm{BMO}_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$. We then have that any b that can be written as:

$$b = b_0 + \sum_{j=1}^{(1+2\epsilon)} R^*_{N,j}b_j$$

with $b_j \in L^\infty(\mathbb{R}^{(1+2\epsilon)})$ must belong to $\mathrm{BMO}_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$.

Corollary (6.2.35)[183]: ([53]).$H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)}) \subsetneq H^1(\mathbb{R}^{(1+2\epsilon)})$.

Proof: We first show that the containment $H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)}) \subset H^1(\mathbb{R}^{(1+2\epsilon)})$holds.

This follows directly from the fact that corresponding BMO spaces norm the H^1spaces, namely that:

$$\|f\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})} \approx \sup_{\|b\|_{BMO_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})}\le 1} \left|\langle f, b\rangle_{L^2}(\mathbb{R}^{(1+2\epsilon)})\right|.$$

An identical statement holds for $H^1(\mathbb{R}^{(1+2\epsilon)})$and $BMO(\mathbb{R}^{(1+2\epsilon)})$. As shown in[65], $BMO(\mathbb{R}^{(1+2\epsilon)}) \subsetneq BMO_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$, and so we have

$$\begin{aligned}\|f\|_{H^1(\mathbb{R}^{(1+2\epsilon)})} &\approx \sup_{\|b\|_{BMO(\mathbb{R}^{(1+2\epsilon)})}\le 1} \left|\langle f, b\rangle_{L^2}(\mathbb{R}^{(1+2\epsilon)})\right| \\ &\le \sup_{\|b\|_{BMO_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})}\le 1} \left|\langle f, b\rangle_{L^2}(\mathbb{R}^{(1+2\epsilon)})\right| \approx \|f\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})}.\end{aligned}$$

This gives the containment, $H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)}) \subset H^1(\mathbb{R}^{(1+2\epsilon)})$.

We now show that there exists a function $f \in H^1(\mathbb{R}^{(1+2\epsilon)})$ but $f \notin H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$. For the sake of simplicity, we just consider the example in dimension 1.
Define

$$f(x) := \frac{\chi_{[0,1]}(x)}{\sqrt{2}} - \frac{\chi_{[-1,0]}(x)}{\sqrt{2}}.$$

It is easy to see that $f(x)$ is supported in $[-1,1]$, and $\int_{\mathbb{R}} f(x)dx = 0$. Moreover, we have

$$\|f\|_{L^2(\mathbb{R}^{(1+2\epsilon)})} = 1.$$

These implies that f is an atom of $H^1(\mathbb{R})$, which shows that $f \in H^1(\mathbb{R})$.
From the definition of f, we obtain that $f_+(x) \frac{\chi_{[0,1]}(x)}{\sqrt{2}}$, and the even extension is

$$f_{+,e}(x) \frac{\chi_{[-1,1]}(x)}{\sqrt{2}}.$$

But, then it is immediate that $f_{+,e} \notin H^1(\mathbb{R})$ since $\int_{\mathbb{R}^{(1+2\epsilon)}} f_{+,e}(x)dx \neq 0$. One can also prove this by using the equivalent definition of $H^1(\mathbb{R})$ via the radial maximal function. Similarly we have these estimates for $f_{-,e}$. Hence, $f_{+,e} \notin H^1(\mathbb{R})$ and $f_{-,e} \notin H^1(\mathbb{R})$, which, combining the result in Corollary (6.2.33) implies that $f \notin H^1_{\Delta_N}(\mathbb{R})$.
Finally, we provide a description of the atoms in $H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$ that connects back to the atom in $H^1(\mathbb{R}^{(1+2\epsilon)})$ (see [178]).

Corollary (6.2.36)[183]: Suppose $a(x)$ is an $H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$-atom supported in $B \subset \mathbb{R}^{(1+2\epsilon)}$. Then we have

$$\int_{\mathbb{R}^{(1+2\epsilon)}} a(x)dx = 0. \tag{93}$$

Moreover, if $B \cap \{x \in \mathbb{R}^{(1+2\epsilon)} : x_{(1+2\epsilon)} = 0\} \neq \emptyset$, we denote $B_+ = B \cap \mathbb{R}^{(1+2\epsilon)}_+$ and $B_- = B \cap \mathbb{R}^{(1+2\epsilon)}_-$. Then we have

$$\int_{B_+} a(x)dx = \int_{B_-} a(x)dx = 0. \tag{94}$$

Proof: First note that from Corollary (6.2.35), $H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)}) \subsetneq H^1(\mathbb{R}^{(1+2\epsilon)})$. Since $a(x)$ is an $H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$ atom, we have $a(x) \in H^1(\mathbb{R}^{(1+2\epsilon)})$, and hence (93) holds, where we use [180].
Second, suppose $B \cap \{x \in \mathbb{R}^{(1+2\epsilon)} : x_{(1+2\epsilon)} = 0\} \neq \emptyset$.
Then we define $a_+(x) = a(x)|_{B_+}$ and $a_-(x) = a(x)|_{B_-}$.
Since $a(x) \in H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$, from Corollary (6.2.33) we obtain that both $a_{+,e}(x)$ and $a_{-,e}(x)$ are in $H^1(\mathbb{R}^{(1+2\epsilon)})$,
which implies that

$$\int_{\mathbb{R}^{(1+2\epsilon)}} a_{+,e}(x)dx = \int_{\mathbb{R}^{(1+2\epsilon)}} a_{-,e}(x)dx = 0.$$

Next we claim that $\int_{\mathbb{R}^{(1+2\epsilon)}} a_{+,e}(x)dx = 0$. In fact,

$$\int_{\mathbb{R}^{(1+2\epsilon)}} a_{+,e}(x)dx = \int_{\mathbb{R}^{(1+2\epsilon)}_+} a_{+,e}(x)dx + \int_{\mathbb{R}^{(1+2\epsilon)}_-} a_{+,e}(x)dx = 2\int_{\mathbb{R}^{(1+2\epsilon)}_+} a_{+,e}(x)dx.$$

Hence,

$$\int_{\mathbb{R}^{(1+2\epsilon)}} a_{+,e}(x)dx = 0$$

implies that $\int_{\mathbb{R}^{(1+2\epsilon)}} a_{+,e}(x)dx = 0$, i.e. $\int_{B_+} a(x)dx = 0$. Similarly we obtain that $\int_{B_-} a(x)dx = 0$. Hence(94) holds.

Corollary (6.2.37)[183]: If $b \in BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)$, then for $0 \le \epsilon \le 2\epsilon$, the commutator

$$\left[b, \mathrm{R}_{N,(1+\epsilon)}\right](f)(x) = b(x)b, \mathrm{R}_{N,(1+\epsilon)}(f)(x) - \mathrm{R}_{N,(1+\epsilon)}(bf)(x)$$

is a bounded map on $L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)$, with operator norm

$$\left\|\left[b, \mathrm{R}_{N,(1+\epsilon)}\right]: L^2\left(\mathbb{R}^{(1+2\epsilon)}\right) \to L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)\right\| \le C\|b\|_{BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)}.$$

Proof: Suppose b is in $BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)$. Then according to [65], we have that $b_{+,e}(x)BMO\left(\mathbb{R}^{(1+2\epsilon)}\right)$and $b_{-,e}(x)BMO\left(\mathbb{R}^{(1+2\epsilon)}\right)$, and moreover,

$$\|b\|_{BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)} \approx \left\|b_{+,e}\right\|_{BMO\left(\mathbb{R}^{(1+2\epsilon)}\right)}\left\|b_{-,e}\right\|_{BMO\left(\mathbb{R}^{(1+2\epsilon)}\right)}.$$

For every$f \in L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)$, we have

$$\left\|\left[b, \mathrm{R}_{N,(1+\epsilon)}\right](f)\right\|^2_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)} = \int_{\mathbb{R}_+^{(1+2\epsilon)}} \left[b, \mathrm{R}_{N,(1+\epsilon)}\right](f)(x)^2 dx + \int_{\mathbb{R}_-^{(1+2\epsilon)}} \left[b, \mathrm{R}_{N,(1+\epsilon)}\right](f)(x)^2 dx =: I + II.$$

For the term I, note that when $x \in \mathbb{R}_+^{(1+2\epsilon)}$, we have

$$\left[b, \mathrm{R}_{N,(1+\epsilon)}\right](f)(x) = b(x)b, \mathrm{R}_{N,(1+\epsilon)}(f)(x) - \mathrm{R}_{N,(1+\epsilon)}(bf)(x)$$
$$= b_{+,e}(x)R_{(1+\epsilon)}\left(f_{+,e}\right)(x) - R_{(1+\epsilon)}\left(b_{+,e}f_{+,e}\right)(x) = \left[b_{+,e}, \mathrm{R}_{(1+\epsilon)}\right]\left(f_{+,e}\right)(x),$$

which implies that

$$I = \int_{\mathbb{R}_+^{(1+2\epsilon)}} \left[b, \mathrm{R}_{N,(1+\epsilon)}\right](f)(x)^2 dx = \int_{\mathbb{R}_+^{(1+2\epsilon)}} \left[b_{+,e}, \mathrm{R}_{(1+\epsilon)}\right]\left(f_{+,e}\right)(x)^2 dx$$
$$\le \int_{\mathbb{R}^{(1+2\epsilon)}} \left[b_{+,e}, \mathrm{R}_{(1+\epsilon)}\right]\left(f_{+,e}\right)(x)^2 dx \le C\left\|b_{+,e}\right\|^2_{BMO\left(\mathbb{R}^{(1+2\epsilon)}\right)}\left\|f_{+,e}\right\|^2_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)},$$

Where $\mathrm{R}_{(1+\epsilon)}$is the classical $(1+\epsilon)$-th Riesz transform $\frac{\partial}{\partial \mathrm{x}_{(1+\epsilon)}}\Delta^{-\frac{1}{2}}$.

For the last estimate we use the result [3], which applies since we know from that $\mathrm{R}_{N,(1+\epsilon)}$ is a Calderón–Zygmund kernel. Similarly we can obtain that

$$II \le C\left\|b_{-,e}\right\|^2_{BMO\left(\mathbb{R}^{(1+2\epsilon)}\right)}\left\|f_{-,e}\right\|^2_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}.$$

Combining the estimates for I and II above, we obtain that

$$\left\|\left[b, \mathrm{R}_{N,(1+\epsilon)}\right](f)\right\|^2_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}$$
$$\le C\left\|b_{+,e}\right\|^2_{BMO\left(\mathbb{R}^{(1+2\epsilon)}\right)}\left\|f_{+,e}\right\|^2_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)} + C\left\|b_{+,e}\right\|^2_{BMO\left(\mathbb{R}^{(1+2\epsilon)}\right)}\left\|f_{-,e}\right\|^2_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}$$
$$\le C\|b\|^2_{BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)}\left(\left\|f_{+,e}\right\|^2_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)} + \left\|f_{-,e}\right\|^2_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}\right)$$
$$\le C\|b\|^2_{BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)}\|f\|^2_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)},$$

which yields that

$$\left\|\left[b, \mathrm{R}_{N,(1+\epsilon)}\right]: L^2\left(\mathbb{R}^{(1+2\epsilon)}\right) \to L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)\right\| \le C\|b\|_{BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)}.$$

Corollary (6.2.38)[183]: Let $g, h \in L^{\infty}\left(\mathbb{R}^{(1+2\epsilon)}\right)$with compact supports. Then for $0 \leq \epsilon \leq 2\epsilon$,

$$\left\|\Pi_{(1+\epsilon)}(h,g)\right\|_{H^1_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)} \leq C\|g\|_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}\|h\|_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}.$$

Proof. By the duality result of [65], stated in Proposition (6.2.12), we know that $H^1_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)^* = \mathrm{BMO}_{\Delta_N}(1+2\epsilon)$. A simple duality computation shows for $b \in BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)$ and for any $g, h \in L^{\infty}\left(\mathbb{R}^{(1+2\epsilon)}\right)$with compact supports:

$$\begin{aligned}\langle b, \Pi_{(1+\epsilon)}(g,h)\rangle_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)} &= \langle b, \mathrm{R}^*_{N,(1+\epsilon)}(g)h - \mathrm{R}_{N,(1+\epsilon)}(h)g\rangle_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)} \\ &= \langle g, \left[b, \mathrm{R}_{N,1+\epsilon}(h)\right]h\rangle_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}.\end{aligned}$$

Thus, from Corollary (6.2.37) , we obtain that

$$\left|\langle b, \Pi_{1+\epsilon}(g,h)\rangle_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}\right| \leq C\|b\|_{BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)}\|g\|_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}\|h\|_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}.$$

This, together with the duality of $H^1_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)$with $BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)$shows that $\Pi_{(1+\epsilon)}(g,h)$is in $H^1_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)$.

And then by testing $\Pi_{(1+\epsilon)}(g,h)$against $b \in BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)$ functions, we find:

$$\begin{aligned}\left\|\Pi_{(1+\epsilon)}(g,h)\right\|_{H^1_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)} &\approx \sup_{\|b\|_{BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)} \leq 1} \left|\langle \Pi_{(1+\epsilon)}(g,h), b\rangle_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}\right| \\ &\leq C\|g\|_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}\|h\|_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)} \sup_{\|b\|_{BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)} \leq 1} \|b\|_{BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)} \\ &\leq C\|g\|_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}\|h\|_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}.\end{aligned}$$

Corollary (6.2.39): Suppose f is a function satisfying:$\int_{\mathbb{R}^{(1+2\epsilon)}} f(x)dx = 0$, and $|f(x)| \leq \chi_{B(x_0,1)}(x) + \chi_{B((x+\epsilon)_0,1)}(x)$, where $|x_0 - (x+\epsilon)_0| := M > 10$. Then we have

$$\|f\|_{H^1_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)} \leq C_{(1+2\epsilon)} \log M. \tag{95}$$

Proof: First note that

$$\begin{aligned}f^+_{\Delta_N}(x) &= \sup_{\epsilon > -1}\left|e^{-(1+\epsilon)\Delta_N} f(x)\right| \\ &= \sup_{\epsilon > -1}\left|\int_{\mathbb{R}^{(1+2\epsilon)}} (1+\epsilon)_{(1+\epsilon),\Delta_N}(x,(x+\epsilon)) f((x+\epsilon)) d(x+\epsilon)\right| \\ &\leq \sup_{\epsilon > -1}\int_{\mathbb{R}^{(1+2\epsilon)}} \left|(1+\epsilon)_{(1+\epsilon),\Delta_N}(x,(x+\epsilon))\right| d(x+\epsilon) \leq C.\end{aligned}$$

Hence, we obtain that

$$\int_{B(x_0,5)} f^+_{\Delta_N}(x)dx + \int_{B((x+\epsilon)_0,5)} f^+_{\Delta_N}(x)dx \leq C_{(1+2\epsilon)}.$$

Now it suffices to estimate

$$\int_{\mathbb{R}^{(1+2\epsilon)} \backslash \left(B(x_0,5) \cup B((x+\epsilon)_0,5)\right)} f^+_{\Delta_N}(x)dx.$$

To see this, we write it as

$$\int_{\frac{\mathbb{R}^{(1+2\epsilon)}}{B(x_0,2M)}} f_{\Delta_N}^{+}(x)dx + \int_{B(x_0,2M)\backslash\left(B(x_0,5)\cup B((x+\epsilon)_0,5)\right)} f_{\Delta_N}^{+}(x)dx =: I + II.$$

We now estimate the term I. First note that from Hölder's regularity (85) of the heat kernel $(1+\epsilon)_{(1+\epsilon),\Delta_N}(x, x+\epsilon)$, we have

$$\left|(1+\epsilon)_{(1+\epsilon),\Delta_N}(x,(x+\epsilon)) - (1+\epsilon)_{(1+\epsilon),\Delta_N}(x,x_0)\right|$$
$$\leq C\left(\frac{|(x+\epsilon)-x_0|}{\sqrt{(1+\epsilon)}+|x-x_0|}\right)\frac{\sqrt{(1+\epsilon)}}{\left(\sqrt{(1+\epsilon)}+|x-x_0|\right)^{(2+2\epsilon)}}$$

for $|(x+\epsilon)-x_0| < \sqrt{(1+\epsilon)}$. Moreover, when $|(x+\epsilon)-x_0| \geq \sqrt{(1+\epsilon)}$, we have

$$\left|(1+\epsilon)_{(1+\epsilon),\Delta_N}(x,(x+\epsilon)) - (1+\epsilon)_{(1+\epsilon),\Delta_N}(x,x_0)\right|$$
$$\leq \left|(1+\epsilon)_{(1+\epsilon),\Delta_N}(x,(x+\epsilon))\right| + \left|(1+\epsilon)_{(1+\epsilon),\Delta_N}(x-x_0)\right|$$
$$\leq C\frac{e^{-|x-x_0|^2/c(1+\epsilon)}e^{-|\epsilon|^2/c(1+\epsilon)}}{(1+\epsilon)^{(1+2\epsilon)/2}} \leq C\left(\frac{|(x+\epsilon)-x_0|}{\sqrt{(1+\epsilon)}}\right)\frac{e^{-|x-x_0|^2/c(1+\epsilon)}}{(1+\epsilon)^{(1+2\epsilon)/2}}$$
$$\leq C\frac{|(x+\epsilon)-x_0|\sqrt{(1+\epsilon)}}{\left(\sqrt{(1+\epsilon)}+|x-x_0|\right)^{(3+2\epsilon)}}.$$

Now note that from the cancellation condition of f and Hölder's regularity of the heat kernel $(1+\epsilon)_{(1+\epsilon)}(x, x+\epsilon)$ as above, we have

$$f_{\Delta_N}^{+}(x) = \sup_{\epsilon>-1}\left|\int_{\mathbb{R}^{(1+2\epsilon)}}\left[(1+\epsilon)_{(1+\epsilon),\Delta_N}(x,(x+\epsilon)) - (1+\epsilon)_{(1+\epsilon),\Delta_N}(x,x_0)\right]f((x\right.$$
$$\left. +\epsilon))d(x+\epsilon)\right|$$
$$\leq C\sup_{\epsilon>-1}\int_{B(x_0,1)\cup B((x+\epsilon)_0,1)}\frac{|(x+\epsilon)-x_0|\sqrt{(1+\epsilon)}}{\left(\sqrt{(1+\epsilon)}+|x-x_0|\right)^{(3+2\epsilon)}}d(x+\epsilon)$$
$$\leq C_{(1+2\epsilon)}\frac{|(x+\epsilon)_0-x_0|}{|x-x_0|^{(2+2\epsilon)}} = C_{(1+2\epsilon)}\frac{M}{|x-x_0|^{(2+2\epsilon)}}.$$

As a consequence, we obtain that

$$I \leq \int_{\mathbb{R}^{(1+2\epsilon)}/B(x_0,2M)} C_{(1+2\epsilon)}\frac{M}{|x-x_0|^{(2+2\epsilon)}}dx \leq C_{(1+2\epsilon)}.$$

We now turn to the termII. Note that when $x \in B(x_0,2M)\backslash\left(B(x_0,5)\cup B((x+\epsilon)_0,5)\right)$, we have

$$\left|\int_{\mathbb{R}^{(1+2\epsilon)}}(1+\epsilon)_{(1+\epsilon),\Delta_N}(x,(x+\epsilon))f((x+\epsilon))d(x+\epsilon)\right|$$
$$\leq \int_{B(x_0,1)}\left|(1+\epsilon)_{(1+\epsilon),\Delta_N}(x,(x+\epsilon))\right|d(x+\epsilon)$$
$$+\int_{B((x+\epsilon)_0,1)}\left|(1+\epsilon)_{(1+\epsilon),\Delta_N}(x,(x+\epsilon))\right|d(x+\epsilon).$$

When $\epsilon > 0$, from the size estimate of the heat kernel $(1+\epsilon)_{(1+\epsilon),\Delta_N}(x,(x+\epsilon))$, we have

$$\left|\int_{\mathbb{R}^{(1+2\epsilon)}}(1+\epsilon)_{(1+\epsilon),\Delta_N}(x,(x+\epsilon))f((x+\epsilon))d(x+\epsilon)\right|$$
$$\leq C\frac{1}{|x-x_0|^{(1+2\epsilon)}}+C\frac{1}{|x-(x+\epsilon)_0|^{(1+2\epsilon)}}.$$

When $\epsilon \leq 0$, similarly we obtain that

$$\left|\int_{\mathbb{R}^{(1+2\epsilon)}}(1+\epsilon)_{(1+\epsilon),\Delta_N}(x,(x+\epsilon))f((x+\epsilon))d(x+\epsilon)\right|$$
$$\leq C\frac{1}{|x-x_0|^{(2+2\epsilon)}}+C\frac{1}{|x-(x+\epsilon)_0|^{(2+2\epsilon)}}$$
$$\leq C\frac{1}{|x-x_0|^{(1+2\epsilon)}}+C\frac{1}{|x-(x+\epsilon)_0|^{(1+2\epsilon)}}.$$

Thus,

$$II \leq \int_{B(x_0,2M)\setminus\left(B(x_0,5)\cup B((x+\epsilon)_0,5)\right)} f_{\Delta_N}^{+}(x)dx$$
$$\leq C\int_{B(x_0,2M)\setminus\left(B(x_0,5)\cup B((x+\epsilon)_0,5)\right)}\frac{1}{|x-x_0|^{(1+2\epsilon)}}+\frac{1}{|-\epsilon)_0|^{(1+2\epsilon)}}dx$$
$$\leq C_{(1+2\epsilon)}\log M.$$

Combining all the estimates above, we obtain that

$$\|f\|_{H_{\Delta_N}^1\left(\mathbb{R}^{(1+2\epsilon)}\right)} = \left\|f_{\Delta_N}^{+}\right\|_{L^1\left(\mathbb{R}^{(1+2\epsilon)}\right)} \leq C_{(1+2\epsilon)}\log M.$$

Corollary (6.2.40)[183]: Suppose $0 \leq \epsilon \leq 2\epsilon$. For every $H_{\Delta_N}^1\left(\mathbb{R}^{(1+2\epsilon)}\right)$-atom $a(x)$ and for all $\varepsilon > 0$ there exist a large positive number M and $g,h \in L^\infty\left(\mathbb{R}^{(1+2\epsilon)}\right)$ with compact supports such that:

$$\left\|a-\Pi_{(1+\epsilon)}(h,g)\right\|_{H_{\Delta_N}^1\left(\mathbb{R}^{(1+2\epsilon)}\right)} < \varepsilon$$

and $\|g\|_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}\|h\|_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)} \leq CM^{(1+2\epsilon)}$.

Proof: Let $a(x)$ be an $H_{\Delta_N}^1\left(\mathbb{R}^{(1+2\epsilon)}\right)$-atom, supported in $B(x_0,r)$.We first consider the construction of the bilinear form $\Pi_{(1+\epsilon)}(h,g)$ for $\epsilon \geq 0$ and the approximation to $a(x)$. To

begin with, for the ball $B(x_0, r)$, we now consider the following cases: Case1:$x_{0,(1+2\epsilon)} \geq 0$; Case2:$x_{0,(1+2\epsilon)} < 0$.
We first consider Case1. To begin with, fix $\varepsilon > 0$. Choose $M \in [100, \infty)$ sufficiently large so that $\frac{\log M}{M} < \varepsilon$. Now select $(x+\epsilon)_0 \in \mathbb{R}_+^{(1+2\epsilon)}$in the following way: for $0 \leq i-1 \leq 2\epsilon$, choose $(x+\epsilon)_{0,i} > 0$ such that $(x+\epsilon)_{0,i} - x_{0,i} = \frac{Mr}{\sqrt{(1+2\epsilon)}}$, where $x_{0,i}$ (reps. $(x+\epsilon)_{0,i}$) is the ith coordinate of x_0(reps.$(x+\epsilon)_0$).
Note that for this $(x+\epsilon)_0$, it is clear that $B((x+\epsilon)_0, r) \subset \mathbb{R}_+^{(1+2\epsilon)}$and we have $|x_0 - (x+\epsilon)_0| = Mr$. Moreover, for any $(x+\epsilon) \in B((x+\epsilon)_0, r)$, we also have $|x_0 - (x+\epsilon)| > \frac{Mr}{2}$. We set

$$g(x) := \chi_{B((x+\epsilon)_0, r)}(x) \quad \text{and} \quad h(x) := \frac{a(x)}{R^*_{N,1+\epsilon} g(x_0)}. \tag{96}$$

We first claim that

$$\left|R^*_{N,(1+\epsilon)} g(x_0)\right| \geq CM^{-(1+2\epsilon)}, \qquad 0 \leq \epsilon \leq 2\epsilon \,. \tag{97}$$

In fact, for $(1+\epsilon) = 1, \dots, 2\epsilon$, from Corollary (6.2.30), we have

$$R^*_{N,(1+\epsilon)} g(x_0) = \left| \int_{B((x+\epsilon)_0, r)} R_{N,(1+\epsilon)}((x+\epsilon), x_0) d(x+\epsilon) \right|$$

$$= C_{(1+2\epsilon)} \left| \int_{B((x+\epsilon)_0, r)} \left(\frac{(x+\epsilon)_{(1+2\epsilon)} - x_{0,(1+2\epsilon)}}{|x_0 - (x+\epsilon)|^{(2+2\epsilon)}} \right.\right.$$

$$\left.\left. + \frac{(x+\epsilon)_{(1+2\epsilon)} - x_{0,(1+2\epsilon)}}{\left(\left|x_0' - (x+\epsilon)'\right|^2 + \left|x_{0,(1+2\epsilon)} + (x+\epsilon)_{(1+2\epsilon)}\right|^2\right)^{(1+\epsilon)}} \right) d(x+\epsilon) \right|$$

$$= C_{(1+2\epsilon)} \big|(x+\epsilon)_{(1+2\epsilon)}$$

$$- x_{0,(1+2\epsilon)}\big| \left| \int_{B((x+\epsilon)_0, r)} \left(\frac{1}{|x_0 - (x+\epsilon)|^{(2+2\epsilon)}} \right.\right.$$

$$\left.\left. + \frac{1}{\left(\left|x_0' - (x+\epsilon)'\right|^2 + \left|x_{0,(1+2\epsilon)} + (x+\epsilon)_{(1+2\epsilon)}\right|^2\right)^{(1+\epsilon)}} \right) d(x+\epsilon) \right|$$

$$\geq CMr \int_{B((x+\epsilon)_0, r)} \frac{1}{|x_0 - (x+\epsilon)|^{(2+2\epsilon)}} d(x+\epsilon) \geq CM^{-(1+2\epsilon)}.$$

As a consequence, we get that the claim(97) holds.
As for Case2, we handle it in a symmetric way as follows. Fix $\varepsilon > 0$. Choose$M \in [100, \infty)$ sufficiently large so that $\frac{\log M}{M} < \varepsilon$. Now select $(x+\epsilon)_0 \in \mathbb{R}_+^{(1+2\epsilon)}$in the following way: for $0 \leq i-1 \leq 2\epsilon$, choose $(x+\epsilon)_{0,i} > 0$ such that $(x+\epsilon)_{0,i} - x_{0,i} = -\frac{Mr}{\sqrt{(1+2\epsilon)}}$. Note that for this $(x+\epsilon)_0$, it is clear that $B((x+\epsilon)_0, r) \subset \mathbb{R}_-^{(1+2\epsilon)}$and we

have $|x_0-(x+\epsilon)_0| = Mr$. Moreover, for any $(x+\epsilon)\in B((1+2\epsilon)_0,r)$, we also have $|x_0-(x+\epsilon)| > \frac{Mr}{2}$. We now define the functions g and h as in(56), and the following the same estimates, we can obtain that the claim(97) holds.
From the definitions of the functions g and$h$, we obtain that supp $g(x) = B((x+\epsilon)_0,r)$ and supp $h(x) = B(x_0,r)$. Moreover, from(97) we obtain that

$$\|g\|_{L^2(\mathbb{R}^{(1+2\epsilon)})} \approx r^{\frac{(1+2\epsilon)}{2}} \text{ and } \|h\|_{L^2(\mathbb{R}^{(1+2\epsilon)})} = \frac{1}{|R_{N,(1+\epsilon)}g(x_0)|}\|a\|_{L^2(\mathbb{R}^{(1+2\epsilon)})}$$

$$\leq CM^{(1+2\epsilon)}r^{-\frac{(1+2\epsilon)}{2}}.$$

Hence $\|g\|_{L^2(\mathbb{R}^{(1+2\epsilon)})}\|h\|_{L^2(\mathbb{R}^{(1+2\epsilon)})} \leq CM^{(1+2\epsilon)}$. Now write

$$a(x) - \left(h(x)R^*_{N,(1+\epsilon)}g(x) - g(x)R_{N,(1+\epsilon)}h(x)\right)$$

$$= a(x)\frac{R^*_{N,(1+\epsilon)}g(x_0) - R^*_{N,(1+\epsilon)}g(x)}{R^*_{N,(1+\epsilon)}g(x_0)} - g(x)R_{N,(1+\epsilon)}h(x)$$

$$=: W_1(x) + W_2(x).$$

By definition, it is obvious that $W_1(x)$is supported on $B(x_0,r)$ and $W_2(x)$ is supported on $B((x+\epsilon)_0,r)$.
We first turn to $W_1(x)$. For $x\in B(x_0,r)$,

$$|W_1(x)| = |a(x)|\frac{|R^*_{N,(1+\epsilon)}g(x_0) - R^*_{N,(1+\epsilon)}g(x)|}{R^*_{N,(1+\epsilon)}g(x_0)}$$

$$\leq CM^{(1+2\epsilon)}\|a\|_{L^\infty(\mathbb{R}^{(1+2\epsilon)})}\int_{B((x+\epsilon)_0,r)}\Big|R_{N,(1+\epsilon)}((x+\epsilon),x_0)$$

$$- R_{N,(1+\epsilon)}((x+\epsilon),x)\Big|d(x+\epsilon) \leq C\frac{M^{(1+2\epsilon)}}{r^{(1+2\epsilon)}}\int_{B((x+\epsilon)_0,r)}\frac{|x-x_0|}{|\epsilon|^{(2+2\epsilon)}}d(x+\epsilon)$$

$$\leq C\frac{1}{Mr^{(1+2\epsilon)}}.$$

Hence $|W_1(x)| \leq C\frac{1}{Mr^{(1+2\epsilon)}}\chi_{B(x_0,r)}(x)$.
We next estimate $W_2(x)$. From the definition of $g(x)$, we have

$$|W_2(x)| = \chi_{B((x+\epsilon)_0,r)}(x)\left|R_{N,(x+\epsilon)}h(x)\right|$$

$$= \chi_{B((x+\epsilon)_0,r)}(x)\frac{1}{|R^*_{N,(x+\epsilon)}g(x_0)|}\left|\int_{B((x+\epsilon)_0,r)} R_{N,(x+\epsilon)}h(x,(x+\epsilon))a((x+\epsilon))d(x+\epsilon)\right|$$

$$= \chi_{B((x+\epsilon)_0,r)}(x)\frac{1}{|R^*_{N,(x+\epsilon)}g(x_0)|}\left|\int_{B((x+\epsilon)_0,r)} R_{N,(1+\epsilon)}\big(x,(x+\epsilon)\big)a_+\big((x+\epsilon)\big)d(x+\epsilon)\right|,$$

where the last equality follows from the fact that $x\in B((x+\epsilon)_0,r)\subset \mathbb{R}_+^{(1+2\epsilon)}$and from the definition of the Riesz kernel $R_N(x,(x+\epsilon))$ as in (86). Hence, from the cancellation property of $a_+((x+\epsilon))$, we get

$$|W_2(x)| = \chi_{B((x+\epsilon)_0,r)}(x)\frac{1}{\left|R^*_{N,(1+2\epsilon)}g(x_0)\right|}\left|\int\limits_{B(x_0,r)}\Big(R_{N,(1+2\epsilon)}(x,(x+\epsilon))\right.$$

$$\left. - R_{N,(1+2\epsilon)}(x,x_0)\Big)a_+((x+\epsilon))d(x+\epsilon)\right|$$

$$\leq C\chi_{B((x+\epsilon)_0,r)}(x)M^{(1+2\epsilon)}\int\limits_{B(x_0,r)}\|a\|_{L^\infty(\mathbb{R}^{(1+2\epsilon)})}\frac{|(x+\epsilon)-x_0|}{|x-x_0|^{(2+2\epsilon)}}d(x+\epsilon)$$

$$\leq \frac{C}{Mr^{(1+2\epsilon)}}\chi_{B((x+\epsilon)_0,r)}(x).$$

Combining the estimates of $W_1$and W_2, we obtain that

$$\left|a(x) - \Big(h(x)R^*_{N,(1+\epsilon)}g(x) - g(x)R_{N,(1+\epsilon)}h(x)\Big)\right|$$

$$\leq \frac{C}{M^{(1+2\epsilon)}}\Big(\chi_{B(x_0,r)}(x) + \chi_{B((x+\epsilon)_0,r)}(x)\Big). \tag{98}$$

Next we point out that

$$\int\Big[a(x) - \Big(h(x)R^*_{N,(1+\epsilon)}g(x) - g(x)R_{N,(1+2\epsilon)}h(x)\Big)\Big]dx$$

$$= \int a(x)dx - \int\Big(h(x)R^*_{N,(1+2\epsilon)}g(x) - g(x)R_{N,(1+2\epsilon)}h(x)\Big)dx = 0, \tag{99}$$

since $a(x)$ has cancellation (Corollary (6.2.36)) and the second integral equals 0 just by the definitions of g and h.

Then the size estimate(98) and the cancellation(99), together with Corollary (6.2.39), imply that

$$\left\|a(x) - \Big(h(x)R^*_{N,(1+\epsilon)}g(x) - g(x)R_{N,(1+\epsilon)}h(x)\Big)\right\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})} \leq C\frac{\log M}{M} \leq C\epsilon.$$

This proves the result for $0 \leq \epsilon \leq 2\epsilon - 1$.

We now consider the bilinear form $\Pi_{(1+\epsilon)}(g,h)$and its approximation to $a(x)$. Again, for the ball $B(x_0,r)$, we now consider the following cases: Case1: $x_{0,(1+2\epsilon)} \geq 0$; Case2: $x_{0,(1+2\epsilon)} < 0$.

It suffices to consider the Case 1 since the other can be handled symmetrically. In this case, for $x_0$with $x_{0,(1+2\epsilon)} \geq 0$, choose $(x+\epsilon)_0$ such that $(x+\epsilon)_{0,i} - x_{0,i} = \frac{Mr}{\sqrt{(1+2\epsilon)}}$ for $i = 1,\ldots,1+2\epsilon$. We now define the functions g and h as in (96). This, together with Corollary (6.2.30), yields

$$R^*_{N,(1+\epsilon)}g(x)\left|\int\limits_{B((x+\epsilon)_0,r)}R_{N,(1+\epsilon)}(x,x_0)d(x+\epsilon)\right|$$

$$= C_{(1+2\epsilon)}\left|\int\limits_{B((x+\epsilon)_0,r)}\left(\frac{(x+\epsilon)_{(1+2\epsilon)} - x_{0,(1+2\epsilon)}}{|x_0-(x+\epsilon)|^{(2+2\epsilon)}}\right.\right.$$

$$\left.\left. + \frac{x_{0,(1+2\epsilon)} + (x+\epsilon)_{(1+2\epsilon)}}{\Big(|x_0' - (x+\epsilon)'|^2 + |x_{0,(1+2\epsilon)} + (x+\epsilon)_{(1+2\epsilon)}|^2\Big)^{(1+\epsilon)}}\right)d(x+\epsilon)\right|$$

$$\geq C_{(1+2\epsilon)}\left|\int_{B((x+\epsilon)_0,r)} \frac{(x+\epsilon)_{(1+2\epsilon)}-x_{0,(1+2\epsilon)}}{|x_0-(x+\epsilon)|^{(2+2\epsilon)}}d(x+\epsilon)\right|$$

$$= C_{(1+2\epsilon)}\left|(x+\epsilon)_{(1+2\epsilon)}-x_{0,(1+2\epsilon)}\right|\left|\int_{B((x+\epsilon)_0,r)} \frac{1}{|x_0-(x+\epsilon)|^{2+2\epsilon}}d(x+\epsilon)\right| \geq CM^{(1+2\epsilon)}.$$

Here, we obtain that the claim(97) holds for these g and h.

Now following the approximation as that for $R_{N,(1+\epsilon)}$with $0 \leq \epsilon \leq -1+2\epsilon$, we obtain that

$$\left\|a(x)-\left(h(x)R^*_{N,(1+\epsilon)}g(x)-g(x)R_{N,(1+\epsilon)}h(x)\right)\right\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})} \leq C\frac{\log M}{M} \leq C\epsilon. \quad (100)$$

With this approximation result, we can now prove the main Theorem (6.2.2), restated below (see [178]).

Corollary (6.2.41)[183]: Suppose $0 \leq \epsilon \leq 2\epsilon$ For any $f \in H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$ there exists sequences $\{\lambda^k_j\} \in \ell^1$and functions $g^k_j, h^k_j \in L^\infty(\mathbb{R}^{(1+2\epsilon)})$with compact supports such that

$$f=\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\lambda^k_{(1+\epsilon)}\Pi_{(1+\epsilon)}(g^k_j,h^k_j).$$

Moreover, we have that:

$$\|f\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})} \approx \inf\left\{\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}|\lambda^k_j|\|g^k_j\|_{L^2(\mathbb{R}^{(1+2\epsilon)})}\|h^k_j\|_{L^2(\mathbb{R}^{(1+2\epsilon)})}: f = \sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\lambda^k_j\Pi_{(1+\epsilon)}(g_j,h_j)\right\}.$$

Proof: By Corollary (6.2.38) we have that $\left\|\Pi_{(1+\epsilon)}(g,h)\right\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})} \leq C\|g\|_{L^2(\mathbb{R}^{(1+2\epsilon)})}\|h\|_{L^2(\mathbb{R}^{(1+\epsilon)})}$, it is immediate that we have for any representation of $f=\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\lambda^k_j\Pi_{(1+\epsilon)}(g_j,h_j)$ that

$$\|f\|_{H^1_{\Delta_N}}(\mathbb{R}^{(1+2\epsilon)}) \leq C\inf\left\{\begin{array}{c}\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}|\lambda^k_j|\|g^k_j\|_{L^2(\mathbb{R}^{(1+2\epsilon)})}\|h^k_j\|_{L^2(\mathbb{R}^{(1+2\epsilon)})}\\ : f=\sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\lambda^k_j\Pi_{(1+\epsilon)}(g^k_j,h^k_j)\end{array}\right\}.$$

We turn to show that the other inequality hold and that it is possible to obtain such a decomposition for any $f \in H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$. By the atomic decomposition for $H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$, Corollary (6.2.33) for any $f \in H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$we can find a sequence $\{\lambda^1_j\} \in \ell^1$and sequence of $H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$-atoms a^1_j so that $f=\sum_{j=1}^{\infty}\lambda^1_j a^1_j$ and $\sum_{j=1}^{\infty}|\lambda^1_j| \leq C_0\|f\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})}$.

We explicitly track the implied absolute constant C_0appearing from the atomic decomposition since it will play a role in the convergence of the approach. Fix $\varepsilon > 0$ so that $\varepsilon C_0 < 1$. Then we also have a large positive number M with $\frac{\log M}{M} < \varepsilon$. We apply Corollary (6.2.40) to each atom a_j^1. So there exists $g_j^1, h_j^1 \in L^\infty(\mathbb{R}^{(1+2\epsilon)})$with compact supports and satisfying $\|g_j^1\|_{L^2(\mathbb{R}^{(1+2\epsilon)})}\|h_j^1\|_{L^2(\mathbb{R}^{(1+2\epsilon)})} \le CM^{(1+2\epsilon)}$and

$$\left\|a_j^1 - \Pi_{(1+\epsilon)}(g_j^1, h_j^1)\right\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})} < \varepsilon \forall_j.$$

Now note that we have

$$f = \sum_{j=1}^{\infty} \lambda_j^1 a_j^1 = \sum_{j=1}^{\infty} \lambda_j^1 \Pi_{(1+\epsilon)}(g_j^1, h_j^1) + \sum_{j=1}^{\infty} \lambda_j^1 \left(a_j^1 - \Pi_{(1+\epsilon)}(g_j^1, h_j^1)\right) \coloneqq M_1 + E_1.$$

Observe that we have

$$\begin{aligned}\|\mathrm{E}_1\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})} &\le \sum_{j=1}^{\infty} |\lambda_j^1| \left\|a_j^1 - \Pi_{(1+\epsilon)}(g_j^1, h_j^1)\right\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})} \\ &\le \varepsilon \sum_{j=1}^{\infty} |\lambda_j^1| \le \varepsilon C_0 \|f\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})}.\end{aligned}$$

We now iterate the construction on the function E_1. Since $\mathrm{E}_1 \in H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$, we can apply the atomic decomposition in $H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$, Corollary (6.2.33), to find a sequence$\{\lambda_j^k\} \in \ell^1$and a sequence of $H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})$-atoms $\{a_j^2\}$ so that $E_1 = \sum_{j=1}^{\infty} \lambda_j^k a_j^2$ and

$$\sum_{j=1}^{\infty} |\lambda_j^2| \le C_0 \|E_1\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})} \le \varepsilon C_0^2 \|f\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})}.$$

Again, we will apply Corollary (6.2.40) to each atom a_j^2. So there exist $g_j^2, h_j^2 \in L^\infty(\mathbb{R}^{(1+2\epsilon)})$ with compact supports and satisfying

$$\|g_j^2\|_{L^2(\mathbb{R}^{(1+2\epsilon)})}\|h_j^2\|_{L^2(\mathbb{R}^{(1+2\epsilon)})} \le CM^{(1+2\epsilon)}$$

and

$$\left\|a_j^2 - \Pi_{(1+\epsilon)}(g_j^2, h_j^2)\right\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})} < \varepsilon, \forall_j.$$

We then have that:

$$E_1 = \sum_{j=1}^{\infty} \lambda_j^2 a_j^2 = \sum_{j=1}^{\infty} \lambda_j^2 \Pi_{(1+\epsilon)}(g_j^2, h_j^2) + \sum_{j=1}^{\infty} \lambda_j^2 \left(a_j^2 - \Pi_{(1+\epsilon)}(g_j^2, h_j^2)\right) \coloneqq M_2 + E_2.$$

But, as before observe that

$$\begin{aligned}\|\mathrm{E}_2\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})} &\le \sum_{j=1}^{\infty} |\lambda_j^2| \left\|a_j^2 - \Pi_{(1+\epsilon)}(g_j^2, h_j^2)\right\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})} \\ &\le \varepsilon \sum_{j=1}^{\infty} |\lambda_j^2| \le (\varepsilon C_0)^2 \|f\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})}.\end{aligned}$$

And, this implies for f that we have:

$$f = \sum_{j=1}^{\infty} \lambda_j^1 a_j^1 = \sum_{j=1}^{\infty} \lambda_j^1 \Pi_{(1+\epsilon)}(g_j^1, h_j^1) + \sum_{j=1}^{\infty} \lambda_j^1 \left(a_j^1 - \Pi_{(1+\epsilon)}(g_j^1, h_j^1)\right) = M_1 + E_1$$

$$= M_1 + M_2 + E_2 = \sum_{k=1}^{2} \sum_{j=1}^{\infty} \lambda_j^k \Pi_{(1+\epsilon)}(g_j^k, h_j^k) + E_2.$$

Repeating this construction for each $1 \le k \le K$ produces functions

$$g_j^k, h_j^k \in L^\infty\left(\mathbb{R}^{(1+2\epsilon)}\right)$$

with compact supports and satisfying

$$\|g_j^k\|_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)} \|h_j^k\|_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)} \le CM^{(1+2\epsilon)}$$

for all j, sequences

$$\{\lambda_j^k\} \in \ell^1 \text{ with } \|\{\lambda_j^k\}\|_{\ell^1} \le \varepsilon^{k-1} C_0^2 \|f\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})},$$

and a function $E_K \in H^1_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)$ with

$$\|E_K\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})} \le (\varepsilon C_0)^K \|f\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})}$$

so that

$$f = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \lambda_j^k \Pi_{(1+\epsilon)}(g_j^k, h_j^k) + E_K.$$

Passing $K \to \infty$ gives the desired decomposition of $f = \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \lambda_j^k \Pi_{(1+\epsilon)}(g_j^k, h_j^k)$. We also have that:

$$\sum_{k=1}^{\infty} \sum_{j=1}^{\infty} |\lambda_j^k| \le \sum_{k=1}^{\infty} \varepsilon^{-1} (\varepsilon C_0)^k \|f\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})} = \frac{C_0}{1 - \varepsilon C_0} \|f\|_{H^1_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})}.$$

Finally, we dispense with the proof of Corollary (6.2.42) (see [178]).

Corollary (6.2.42)[183]: Suppose $b \in \cup_{\epsilon \ge 0} L^{(1+\epsilon)}_{loc}\left(\mathbb{R}^{(1+2\epsilon)}\right)$.

If b is in $BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)$, then for $\epsilon \ge 0$, the commutator

$$[b, R_{N,(1+\epsilon)}](f)(x) = b(x) R_{N,(1+\epsilon)}(f)(x) - R_{N,(1+\epsilon)}(bf)(x)$$

is a bounded map on $L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)$, with operator norm

$$\left\|[b, R_{N,1+\epsilon}]: L^2\left(\mathbb{R}^{(1+2\epsilon)}\right) \to L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)\right\| \le C \|b\|_{\mathrm{BMO}_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})}.$$

Conversely, for $\epsilon \ge 0$, if $[b, R_{N,1+\epsilon}]$ are bounded on $L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)$ then

$$b \text{ is in } \mathrm{BMO}_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right) \text{and } \|b\|_{\mathrm{BMO}_{\Delta_N}(\mathbb{R}^{(1+2\epsilon)})} \le C \left\| \begin{matrix} [b, R_{N,1+\epsilon}]: L^2\left(\mathbb{R}^{(1+2\epsilon)}\right) \\ \to L^2\left(\mathbb{R}^{(1+2\epsilon)}\right) \end{matrix} \right\|.$$

We point out that Theorem (6.2.2) and Corollary (6.2.42) can be extended to work for $L^{(1+\epsilon)}\left(\mathbb{R}^{(1+2\epsilon)}\right)$ when $0 < \epsilon < \infty$.

For $\epsilon \ge 0$, the fractional operator $\Delta_N^{-(1+\epsilon)/2}$ of the operator Δ_N is defined by

$$\Delta_N^{-(1+\epsilon)/2} f(x) = \frac{1}{\Gamma((1+\epsilon)/2)} \int_0^\infty e^{-(1+\epsilon)\Delta_N}(f)(x) \frac{d(1+\epsilon)}{(1+\epsilon)^{-\epsilon/2}}.$$

Proof. The upper bound in this theorem is contained in Corollary (6.2.38).

For the lower bound, we first note that from Corollary (6.2.33), $H^1_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)$ has equivalent characterizations via atoms, which shows that

$$H^1_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right) \cap L^\infty_c\left(\mathbb{R}^{(1+2\epsilon)}\right)$$

is dense in $H^1_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)$ with respect to the $H^1_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)$ norm, where we use $L^\infty_c\left(\mathbb{R}^{(1+2\epsilon)}\right)$ to denote the L^∞ function with compact supports.

Then using the weak factorization in Theorem (6.2.2) we have that for $f \in H^1_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right) \cap L^\infty_c\left(\mathbb{R}^{(1+2\epsilon)}\right)$,

$$\left|\langle b, f\rangle_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}\right| \leq \sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\left|\lambda_j^k\right| \left|\langle b, \Pi_{(1+\epsilon)}\left(g_j^k, h_j^k\right)\rangle_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}\right|$$

$$= \sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\left|\lambda_j^k\right| \left|\langle g_j^k, \left[b, R_{N,(1+\epsilon)}\right] h_j^k\rangle_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}\right|.$$

Hence we have that

$$\left|\langle b, f\rangle_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}\right| \leq \sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\left|\lambda_j^k\right| \left\|\left[b, R_{N,(1+\epsilon)}\right]\left(h_j^k\right)\right\|_{L^2\left(\mathbb{R}^{(1+2\epsilon}\right)} \left\|g_j^k\right\|_{L^2\left(\mathbb{R}^{(1+2\epsilon}\right)}$$

$$\leq \left\|\left[b, R_{N,(1+\epsilon)}\right]: L^2\left(\mathbb{R}^{(1+2\epsilon)}\right) \to L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)\right\| \sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\left|\lambda_j^k\right| \left\|g_j^k\right\|_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)} \left\|h_j^k\right\|_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}$$

$$\leq C\left\|\left[b, R_{N,(1+\epsilon)}\right]: L^2\left(\mathbb{R}^{(1+2\epsilon)}\right) \to L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)\right\| \|f\|_{H^1_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)}.$$

By the duality between $\mathrm{BMO}_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)$ and $H^1_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)$ we have that:

$$\|b\|_{BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)} \approx \sup_{\|f\|_{H^1_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right) \leq 1}} \left|\langle b, f\rangle_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}\right|$$

$$\leq C\left\|\left[b, R_{N,(1+\epsilon)}\right]: L^2\left(\mathbb{R}^{(1+2\epsilon)}\right) \to L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)\right\|$$

Corollary (6.2.43)[183]: If b is in $\mathrm{BMO}_{\Delta_\mathrm{N}}\left(\mathbb{R}^{(1+2\epsilon)}\right)$, then for $\epsilon > 0$, the commutator

$$\left[b, \Delta_N^{-(1+\epsilon)/2}\right](f)(x) = b(x)\Delta_N^{-(1+\epsilon)/2}(f)(x) - \Delta_N^{-(1+\epsilon)/2}(bf)(x)$$

is a bounded map from $L^{1+\epsilon}(\mathbb{R}^{1+2\epsilon})$ to $L^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)}\left(\mathbb{R}^{(1+2\epsilon)}\right)$ with operator norm

$$\left\|\left[b, \Delta_N^{-(1+\epsilon)/2}\right]: L^{1+\epsilon}\left(\mathbb{R}^{(1+2\epsilon)}\right) \to L^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)}\left(\mathbb{R}^{(1+2\epsilon)}\right)\right\| \leq C\|b\|_{BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)},$$

where $\epsilon \geq 0$.

Proof. Suppose b is in $\mathrm{BMO}_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)$. Then according to [65], we have that $b_{+,e} \in BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)$ and $b_{-,e} \in BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)$, and moreover,

$$\|b\|_{BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)} \approx \left\|b_{+,e}\right\|_{BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)} \left\|b_{-,e}\right\|_{BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)}.$$

For every $f \in L^{(1+\epsilon)}\left(\mathbb{R}^{(1+2\epsilon)}\right)$, we have

$$\left\|\left[b, \Delta_N^{-(1+\epsilon)/2}\right](f)\right\|^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)}_{L^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)}\left(\mathbb{R}^{(1+2\epsilon)}\right)}$$

$$= \int_{\mathbb{R}_+^{(1+2\epsilon)}} \left[b, \Delta_N^{-(1+\epsilon)/2}\right](f)(x)^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)} dx$$

$$+ \int_{\mathbb{R}_-^{(1+2\epsilon)}} \left[b, \Delta_N^{-(1+\epsilon)/2}\right](f)(x)^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)} dx =: I + II.$$

For the term$I$, note that when $x \in \mathbb{R}_+^{(1+2\epsilon)}$, we have

$$\left[b, \Delta_N^{-(1+\epsilon)/2}\right](f)(x) = b(x)\Delta_N^{-(1+\epsilon)/2}(f)(x) - \Delta_N^{-(1+\epsilon)/2}(bf)(x)$$
$$= b_{+,e}(x)\Delta^{-(1+\epsilon)/2}\left(f_{+,e}\right)(x) - \Delta^{-(1+\epsilon)/2}\left(b_{+,e}f_{+,e}\right)(x)$$
$$= \left[b_{+,e}, \Delta^{-(1+\epsilon)/2}\right]\left(f_{+,e}\right)(x),$$

which implies that

$$I = \int_{\mathbb{R}_+^{(1+2\epsilon)}} \left[b, \Delta^{-(1+\epsilon)/2}\right](f)(x)^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)} dx$$
$$= \int_{\mathbb{R}_+^{(1+2\epsilon)}} \left[b_{+,e}, \Delta^{-(1+\epsilon)/2}\right]\left(f_{+,e}\right)(x)^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)} dx$$
$$\leq \int_{\mathbb{R}^{(1+2\epsilon)}} \left[b_{+,e}, \Delta^{-(1+\epsilon)/2}\right]\left(f_{+,e}\right)(x)^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)} dx$$
$$\leq C\left\|b_{+,e}\right\|_{BMO\left(\mathbb{R}^{(1+2\epsilon)}\right)}^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)}\left\|f_{+,e}\right\|_{L^{(1+\epsilon)}\left(\mathbb{R}^{(1+2\epsilon)}\right)}^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)}.$$

For the last estimate we use the result [3], which applies since we know from that $R_{N,(1+\epsilon)}$ is a Calderón–Zygmund kernel. Similarly we can obtain that

$$II \leq C\left\|b_{-,e}\right\|_{BMO\left(\mathbb{R}^{(1+2\epsilon)}\right)}^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)}\left\|f_{-,e}\right\|_{L^2\left(\mathbb{R}^{(1+2\epsilon)}\right)}^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)}.$$

Combining the estimates for I and II above, we obtain that

$$\left\|\left[b, \Delta_N^{-(1+\epsilon)/2}\right](f)\right\|_{L^{(1+\epsilon)}\left(\mathbb{R}^{(1+2\epsilon)}\right)}^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)}$$
$$\leq C\left\|b_{+,e}\right\|_{BMO\left(\mathbb{R}^{(1+2\epsilon)}\right)}^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)}\left\|f_{+,e}\right\|_{L^{(1+\epsilon)}\left(\mathbb{R}^{(1+2\epsilon)}\right)}^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)}$$
$$+ C\left\|b_{-,e}\right\|_{BMO\left(\mathbb{R}^{(1+\epsilon)}\right)}^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)}\left\|f_{-,e}\right\|_{L^{(1+\epsilon)}\left(\mathbb{R}^{(1+2\epsilon)}\right)}^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)}$$
$$\leq C\|b\|_{BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)}^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)}\left(\left\|f_{+,e}\right\|_{L^{(1+\epsilon)}\left(\mathbb{R}^{(1+2\epsilon)}\right)}^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)} + \left\|f_{-,e}\right\|_{L^{(1+\epsilon)}\left(\mathbb{R}^{(1+2\epsilon)}\right)}^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)}\right)$$
$$\leq C\|b\|_{BMO_{\Delta_N}\left(\mathbb{R}^{(1+2\epsilon)}\right)}^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)}\|f\|_{L^{(1+\epsilon)}\left(\mathbb{R}^{(1+2\epsilon)}\right)}^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)},$$

which yields that

$$\left\|\left[b, R_{N,(1+\epsilon)}\right]: L^{(1+\epsilon)}\left(\mathbb{R}^{(1+2\epsilon)}\right) \to L^{\left(\frac{1+3\epsilon+2\epsilon^2}{\epsilon}\right)}\left(\mathbb{R}^{(1+\epsilon)}\right)\right\| \leq C\|b\|_{BMO_{\Delta_N}\left(\mathbb{R}^{(1+\epsilon)}\right)}.$$

List of Symbols

References

[1] SARAH H. FERGUSON, MICHAEL T. LACEY, A characterization of product BMO by commutators, Acta Math., 189 (2002), 143-160.

[2] CHANG, S.-Y. A. ~: FEEFERMAN, R., Some recent developments in Fourier analysis and H^p-theory on product domains. Bull. Amer. Math. Soc. (N.S.), 12 (1985), 1-43.

[3] COIFMAN, R.R., ROCHBERG, R. ~% WEISS, G., Factorization theorems for Hardy spaces in several variables. Ann. of Math. (2), 103 (1976), 611-635.

[4] FEFFERMAN, R., Bounded mean oscillation on the polydisk. Ann. of Math. (2), 110 (1979), 395-406.

[5] FEI~CUSON, S.H. &: SADOSKY, C., Characterizations of bounded mean oscillation on the polydisk in terms of Hankel operators and Carleson measures. J. Anal. Math., 81 (2000), 239-267.

[6] GUNDY, R. F. ~ STEIN, E. M., H^p theory for the poly-disc. Proc. Nat. Acad. Sci. U.S.A., 76 (1979), 1026-1029.

[7] JOURNE, J.-L., A covering lemma for product spaces. Proe. Amer. Math. Soc., 96 (1986), 593-598.

[8] MEYER, Y., Wavelets and Operators. Translated from the 1990 French original by D. H. Salinger. Cambridge Stud. Adv. Math., 37. Cambridge Univ. Press, Cambridge, 1992.

[9] NEHARI, Z., On bounded bilinear forms. Ann. of Math. (2), 65 (1957), 153-162.

[10] PIPHER, J., Journ~'s covering lemma and its extension to higher dimensions. Duke Math. J., 53 (1986), 683-690.

[11] STEIN, E. M. & WEISS, G., Introduction to Fourier Analysis on Euclidean Spaces. Princeton Math. Ser., 32. Princeton Univ. Press, Princeton, N J, 1971.

[12] Camil Muscalu, Jill Pipher, Terence Tao and Christoph Thiele, Multi-parameter paraproducts, Rev. Mat. Iberoamericana 22 (2006), no. 3, 963–976.

[13] Coifman, R. and Meyer, Y.: Ondelettes et operateurs III. Operateurs multilin´eaires. Actualit´es Math´ematiques. Hermann, Paris, 1991.

[14] Fefferman, C. and Stein, E. M.: Some maximal inequalities. Amer. J. Math. 93 (1971), 107–115.

[15] Grafakos, L. and Torres, R.: Multilinear Calder´on-Zygmund theory. Adv. Math. 165 (2002), 124–164.

[16] Journ´ e, J.: Calder´on-Zygmund operators on product spaces. Rev. Mat. Iberoamericana 1 (1985), 55–91.

[17] Kenig, C. E. and Stein, E. M.: Multilinear estimates and fractional integration. Math. Res. Lett. 6 (1999), 1–15.

[18] Muscalu, C., Pipher, J, Tao, T. and Thiele, C.: Bi-parameter paraproducts. Acta Math. 193 (2004), 269–296.

[19] Stein, E. M.: Harmonic analysis: real variable methods, orthogonality, and oscillatory integrals. Princeton University Press, Princeton, NJ, 1993.

[20] Yu. Brudnyi, An Algorithm of Nonlinear Approximation by Piecewise Polynomials,

[21] Yu. Brudnyi, Nonlinear approximation by refinable functions and piecewise polynomials. In: Approximation Theory X, Vanderbild Univ. Press (2002), p. 91-112.

[22] Yu. Brudnyi, Nonlinear piecewise polynomial approximation of functions from Besove spaces. (Submitted)
[23] Yu. Brudnyi and I. Kozlov, An Algorithm of Nonlinear Approximation by refinable functions. (Submitted)
[24] YU. BRUDNY˘I, NONLINEAR N-TERM APPROXIMATION BY REFINABLE FUNCTIONS, Vol. 16 (2005), No. 1, Pages 143–179.
[25] C. Bandt, Self-similar sets. 5. Integer matrices and fractal tilings of $\mathbb{R}^n$, Proc. Amer. Math. Soc. 112 (1991), 549–562. MR1036982 (92d:58093)
[26] Yu. Brudny˘ı, Rational approximation and imbedding theorems, Dokl. Akad. Nauk SSSR 247 (1979), no. 2, 269–272; English transl., Soviet Math. Dokl. 20 (1979), no. 4, 681–684. MR0545347 (81d:41016)
[27] Adaptive approximation of functions with singularities, Trudy Moskov. Mat. Obshch. 55 (1994), 149–242; English transl., Trans. Moscow Math. Soc. 1994, 123–186 (1995). MR1468457 (98f:41004)
[28] Yu. Brudny˘ı and I. Irodova, Nonlinear spline approximation of functions of several variables and B-spaces, Algebra i Analiz 4 (1992), no. 4, 45–79; English transl., St. Petersburg Math. J. 4 (1993), no. 4, 667–694. MR1190782 (93i:41023)
[29] Yu. Brudny˘ı and N. Kruglyak, A family of approximation spaces, Studies in the Theory of Functions of Several Real Variables, No. 2, Yaroslav. Gos. Univ., Yaroslavl , 1978, pp. 15–42. (Russian) MR0559913 (81m:46101)
[30] Interpolation functors and interpolation spaces. Vol. 1, North-Holland Math. Library, vol. 47, North-Holland, Amsterdam, 1991. MR1107298 (93b:46141)
[31] Yu. Brudny˘ı and B. Kotlyar, A certain problem of combinatorial geometry, Sibirsk. Mat. Zh. 11 (1970), no. 5, 1171–1173; English transl., Siberian Math. J. 11 (1970), 870–871. MR0270270 (42:5160)
[32] J. Bergh and J. Löfström, Interpolation spaces. An introduction, Grundlehren Math. Wiss., vol. 223, Springer-Verlag, Berlin–New York, 1976. MR0482275 (58:2349)
[33] M. Birman and M. Solomyak, Piecewise polynomial approximations of functions of classes Wα p , Mat. Sb. (N. S.) 73 (1967), no. 3, 331–355; English transl., Math. USSR–Sb. 2 (1967), 295–317. MR0217487 (36:576)
[34] Estimates for the singular numbers of integral operators, Uspekhi Mat. Nauk 32 (1977), no. 1, 17–84; English transl., Russian Math. Surveys 32 (1977), no. 1, 15–89. MR0438186 (55:11104)
[35] Quantitative analysis in Sobolev imbedding theorems and applications to spectral theory, Tenth Mathematical School (Summer School, Kaciveli/Nalchik, 1972), Inst. Mat. Akad. Nauk Ukrain. SSR, Kiev, 1974, pp. 5–189. (Russian) MR0482138 (58:2224)
[36] A. S. Cavaretta, W. Dahmen, and C. A. Miccelli, Stationary subdivisions, Mem. Amer. Math. Soc. 93 (1991), no. 453, 186 pp. MR1079033 (92h:65017)
[37] A. Cohen, K. Gröchenig, and L. F. Villemoes, Regularity of multivariate refinable functions, Constr. Approx. 15 (1999), 241–255. MR1668921 (2000b:42028)
[38] R. A. DeVore, Nonlinear approximation, Acta Numer., vol. 7, Cambridge Univ. Press, Cambridge, 1998, pp. 51–150. MR1689432 (2001a:41034)
[39] R. A. DeVore, B. Jawerth, and V. Popov, Compression of wavelet decompositions, Amer. J. Math. 114 (1992), 737–785. MR1175690 (94a:42045)
[40] R. A. DeVore, P. Petrushev, and X. M. Yu, Nonlinear wavelet approximation in the space C(Rd), Progress in Approximation Theory (Tampa, FL, 1990), Springer Ser.

Comput. Math., vol 19, Springer, New York, 1992, pp. 261–283. MR1240786 (94h:41070)
[41] J. E. Hutchinson, Fractals and self-similarity, Indiana Univ. Math. J. 30 (1981), 713–747. MR0625600 (82h:49026)
[42] X.-G. He, K.-S. Lau, and H. Rao, Self-affine sets and graph-directed systems, Constr. Approx. 19 (2003), 373–397; Preprint, City Univ. of Hong-Kong, 2002, 27 pp.
[43] R. Q. Jia, Approximation properties of multivariate wavelets, Preprint, Univ. of Alberta, Edmonton, 1996.
[44] R. Q. Jia and C. A. Micchelli, On linear independence for integer translates of a finite number of functions, Proc. Edinburgh Math. Soc. (2) 36 (1993), 69–85. MR1200188 (94e:41044)
[45] P. G. Lemari´e-Rieusset, Projecteurs invariants, matrices de dilatation, ondelettes et analyses multi-r´esolutions, Rev. Mat. Iberoamericana 10 (1994), no. 2, 283–347. MR1286477 (95e:42039)
[46] J. C. Lagarias and Y. Wang, Self-affine tiles in $\mathbb{R}^n$, Adv. Math. 121 (1996), 21–49. MR1399601 (97d:52034)
[47] P. Petrushev and V. Popov, Rational approximation of real functions, Encyclopedia Math. Appl., vol. 28, Cambridge Univ. Press, Cambridge, 1987. MR0940242 (89i:41022)
[48] K. H. Rosen (ed.), Handbook of discrete and combinatorial mathematics, CRC Press, Boca Raton, FL, 2000. MR1725200 (2000g:05001)
[49] E. Schmidt, Zur Theorie der linearen und nonlinearen Integralgleichungen. I, Math. Ann. 63 (1907), 433–476.
[50] V. N. Temlyakov, Nonlinear methods of approximation, IMI–Preprint Ser., Univ. of South Caroline, 2001, 1–57.
[51] P. Wojtaszczyk, A mathematical introduction to wavelets, London Math. Soc. Student Text, vol. 37, Cambridge Univ. Press, Cambridge, 1997. MR1436437 (98j:42025)
[52] D. -X. Zhou, Self-similar lattice tilings and subdivision schemes, SIAM J. Math. Anal. 33 (2001), no. 1, 1–15 (electronic). MR1857987 (2002g:41031)
[53] LIXIN YAN, CLASSES OF HARDY SPACES ASSOCIATED WITH OPERATORS, DUALITY THEOREM AND APPLICATIONS, Volume 360, Number 8, August 2008, Pages 4383–4408.
[54] D. Albrecht, X.T. Duong and A. McIntosh, Operator theory and harmonic analysis, Instructional Workshop on Analysis and Geometry (Proc. Centre Math. Analysis, 34, A.N.U., Canberra, 1996) pp 77-136. MR1394696 (97e:47001)
[55] P. Auscher and E. Russ, Hardy spaces and divergence operators on strongly Lipschitz domain of Rn, J. Funct. Anal., 201 (2003), 148-184. MR1986158 (2004c:42049)
[56] P. Auscher and P. Tchamitchian, Calcul fonctionnel pr´ecis´e pour des op´erateurs elliptiques complexes en dimension un (et applications `a certaines ´equations elliptiques complexes en dimension deux), Ann. Institut Fourier (Grenoble), 45 (1995), 721-778. MR1340951 (96f:35036)
[57] P. Auscher and P. Tchamitchian, Square root problem for divergence operators and related topics, Asterisque, 249, Soc. Math. France, 1998. MR1651262 (2000c:47092)

[58] P. Auscher, X.T. Duong and A. McIntosh, Boundedness of Banach space valued singular integral operators and Hardy spaces, unpublished preprint, (2005).
[59] S. Blunck and P.C. Kunstmann, Weak type (p, p) estimates for Riesz transforms, Math. Z., 247 (2004), 137-148. MR2054523 (2005f:35071)
[60] T. Coulhon and X.T. Duong, Maximal regularity and kernel bounds: observations on a theorem by Hieber and Pr¨uss, Adv. Differential Equations, 5 (2000), 343-368. MR1734546 (2001d:34087)
[61] D-C. Chang, S.G. Krantz and E.M. Stein, H^p theory on a smooth domain in $\mathbb{R}^n$ and elliptic boundary value problems, J. Funct. Anal., 114 (1993), 286–347. MR1223705 (94j:46032)
[62] R.R. Coifman, Y. Meyer and E.M. Stein, Un nouvel espace adapt´e a l’´etude des op´erateurs d´efinis par des int´egrales singuli´eres, in “Proc. Conf. Harmonic Analysis, Cortona”, Lecture Notes in Math. Vol. 992, pp, 1-15, Springer-Verlag, Berlin/New York, 1983. MR729344 (85j:42032)
[63] R.R. Coifman, Y. Meyer and E.M. Stein, Some new functions and their applications to harmonic analysis, J. Funct. Analysis, 62 (1985), 304-315. MR791851 (86i:46029)
[64] E.B. Davies, Heat kernels and spectral theory, Cambridge Univ. Press, 1989. MR990239 (90e:35123)
[65] D.G. Deng, X.T. Duong, A. Sikora and L.X. Yan, Comparison of the classical BMO with the BMO spaces associated with operators and applications, to appear, Rev. Mat. Iberoamericana (2008).
[66] D.G. Deng, X.T. Duong and L.X. Yan, A characterization of the Morrey-Campanato spaces, Math. Z., 250 (2005), 641-655. MR2179615 (2006g:42039)
[67] X.T. Duong and A. McIntosh, Singular integral operators with non-smooth kernels on irregular domains, Rev. Mat. Iberoamericana, 15 (1999), 233-265. MR1715407 (2001e:42017a)
[68] X.T. Duong, E.M. Ouhabaz and L.X. Yan, Endpoint estimates for Riesz transforms of magnetic Schrödinger operators, Arkiv för Matematik 44 (2006), 261-275. MR2292721 (2008a:35224)
[69] P.L. Duren, B.W. Romberg and A.L. Shields, Linear functionals on H^p spaces with $0 <p< 1$, J. Reine Angew. Math., 238 (1969), 32-60. MR0259579 (41:4217)
[70] X.T. Duong and L.X. Yan, New function spaces of BMO type, the John-Nirenberg inequality, interpolation and applications, Comm. Pure Appl. Math., 58 (2005), 1375-1420. MR2162784 (2006i:26012)
[71] X.T. Duong and L.X. Yan, Duality of Hardy and BMO spaces associated with operators with heat kernel bounds, J. Amer. Math. Soc., 18 (2005), 943-973. MR2163867 (2006d:42037)
[72] X.T. Duong and L.X. Yan, New Morrey-Campanato spaces associated with operators and applications, preprint, 2005.
[73] J. Dziuba´nski and J. Zienkiewicz, H^p spaces associated with Schrödinger operators with potentials from reverse Hölder classes, Colloq. Math., 98 (2003), 5-38. MR2032068 (2004k:42038)
[74] C. Fefferman and E.M. Stein, H^p spaces of several variables, Acta Math., 129 (1972), 137– 195. MR0447953 (56:6263)
[75] F. John and L. Nirenberg, On functions of bounded mean oscillation. Comm. Pure Appl. Math., 14 (1961), 415–426. MR0131498 (24:A1348)

[76] S. Janson, M.H. Taibleson, G. Weiss, Elementary characterizations of the Morrey-Campanato spaces, Lecture Notes in Math., 992 (1983), 101-114. MR729349 (85k:46033)
[77] S. Hofmann and J.M. Martell, L^p bounds for Riesz transforms and square roots associated to second order elliptic operators, Publ. Mat., 47 (2003), 497-515. MR2006497 (2004i:35067)
[78] J.M. Martell, Sharp maximal functions associated with approximations of the identity in spaces of homogeneous type and applications, Studia Math., 161(2004), 113-145. MR2033231 (2005b:42016)
[79] A. McIntosh, Operators which have an H^∞-calculus, Miniconference on operator theory and partial differential equations (Proc. Centre Math. Analysis, ANU, Canberra 14, 1986), 210- 231. MR912940 (88k:47019)
[80] E.M. Ouhabaz, Analysis of heat equations on domains. London Math. Soc. Mono. 31, Princeton Univ. Press, (2004). MR2124040 (2005m:35001)
[81] S. Semmes, Square function estimates and the T(b) theorem, Proc. Amer. Math. Soc., 110 (1990), 721-726. MR1028049 (91h:42018)
[82] E.M. Stein, Singular integral and differentiability properties of functions, Princeton Univ. Press, 30, (1970). MR0290095 (44:7280)
[83] W.A. Strauss, Partial differential equations: An introduction. John Wiley & Sons, Inc., New York, 1992. MR1159712 (92m:35001)
[84] E.M. Stein and G. Weiss, On the theory of harmonic functions of several variables I, The theory of H^p spaces, Acta Math., 103 (1960), 25-62. MR0121579 (22:12315)
[85] A. Torchinsky, Real-variable methods in harmonic analysis, Pure and Applied Math., Vol 123, Academic Press, (1986). MR869816 (88e:42001)
[86] M.W. Taibleson and G. Weiss, The molecular characterization of certain Hardy spaces, Aste´risque. 77, 68-149 (1980). MR0604370 (83g:42012)
[87] A. Uchiyama and J.M. Wilson, Approximate identities and $H^1(\mathbb{R})$, Proc. Amer. Math. Soc., 88 (1983), 53-58. MR691278 (84c:42032)
[88] N. Varopoulos, L. Saloff-Coste and T. Coulhon, Analysis and geometry on groups. Cambridge Univ. Press, London, 1993. MR1218884 (95f:43008)
[89] G. Weiss, Some problems in the theory of Hardy spaces, Proc. Sympos. Pure Math. Vol. 35, Amer. Math. Soc., Providence, R.I., (1979), 189-200. MR545258 (80j:30051)
[90] L.X. Yan, Littlewood-Paley functions associated to second order elliptic operators, Math. Z., 246, (2004), 655-666. MR2045834 (2005a:42015)
[91] K. Yosida, Functional Analysis (Fifth edition), Springer-Verlag, Berlin, 1978. MR0500055 (58:17765)
[92] Alexopoulos, G.: Spectral multipliers for Markov chains. J. Math. Soc. Japan 56 (2004), no. 3, 833–852.
[93] Auscher, P., Russ, E. and Tchamitchian, P.: Hardy Sobolev spaces on strongly Lipschitz domain of Rn. J. Funct. Anal. 218 (2005), 54–109.
[94] Christ, M.: L^p bounds for spectral multipliers on nilpotent groups. Trans. Amer. Math. Soc. 328 (1991), 73–81.
[95] Coifman, R. and Weiss. G.: Extensions of Hardy spaces and their use in analysis. Bull. Amer. Math. Soc. 83 (1977), 569–645.
[96] Cowling, M. and Meda, S.: Harmonic analysis and ultracontractivity. Trans. Amer. Math. Soc. 340 (1993), 733–752.

[97] Duong, X. T., Ouhabaz, E. M. and Sikora, A.: Plancherel-type estimates and sharp spectral multipliers. J. Funct. Anal. 196 (2002), 443–485.
[98] Duong, X. T. and Yan, L.: On commutators of fractional integrals. Proc. Amer. Math. Soc. 132 (2004), 3549–3557.
[99] Dziubanski, J., Garrig ´ os, G., Mart ´ ´ınez, T., Torrea, J. L. and Zienkiewicz, J.: BMO spaces related to Schrödinger operators with potentials satisfying a reverse Hölder inequality. Math. Z. 249 (2005), 329–356.
[100] Gunawan, H.: On weighted estimates for Stein's maximal function. Bull. Austral. Math. Soc. 54 (1996), 35–39.
[101] Hebisch, W.: A multiplier theorem for Schrödinger operators. Colloq. Math. 60/61 (1990), no. 2, 659–664.
[102] Mauceri, G. and Meda, S.: Vector-valued multipliers on stratified groups. Rev. Mat. Iberoamericana 6 (1990), 141–154.
[103] Muller, D. and Stein, E. M. ¨ : On spectral multipliers for Heisenberg and related groups. J. Math. Pures Appl. (9) 73 (1994), 413–440.
[104] Sikora, A. and Wright, J.: Imaginary powers of Laplace operators. Proc. Amer. Math. Soc. 129 (2001), 1745–1754.
[105] Triebel, H.: Theory of function spaces. Monographs in Mathematics 78. Birkhauser verlag, Basel, Boston, Stuggart, 1983.
[106] Laurent Dalenc and Yumeng Ou, UPPER BOUND FOR MULTI-PARAMETER ITERATED COMMUTATORS, Publ. Mat. 60 (2016), 191–220.
[107] X.-Y. A. Chang and R. Fefferman, A continuous version of duality of H1 with BMO on the bidisc, Ann. of Math. (2) 112(1) (1980), 179–201. DOI: 10.2307/1971324.
[108] L. Dalenc and S. Petermichl, A lower bound criterion for iterated commutators, J. Funct. Anal. 266(8) (2014), 5300–5320. DOI: 10.1016/j.jfa.2014.02.015.
[109] T. P. Hytonen ¨ , Representation of singular integrals by dyadic operators, and the A2 theorem, Preprint (2011). arXiv:1108.5119.
[110] T. P. Hytonen ¨ , The sharp weighted bound for general Calder´on– Zygmund operators, Ann. of Math. (2) 175(3) (2012), 1473–1506. DOI: 10.4007/annals.2012.175.3.9.
[111] M. T. Lacey, S. Petermichl, J. C. Pipher, and B. D. Wick, Multiparameter Riesz commutators, Amer. J. Math. 131(3) (2009), 731–769. DOI: 10.1353/ajm.0.0059.
[112] M. T. Lacey, S. Petermichl, J. C. Pipher, and B. D. Wick, Iterated Riesz commutators: a simple proof of boundedness, in: "Harmonic analysis and partial differential equations", Contemp. Math. 505, Amer. Math. Soc., Providence, RI, 2010, pp. 171–178. DOI: 10.1090/conm/505/09922.
[113] M. Lacey and E. Terwilleger, Hankel operators in several complex variables and product BMO, Houston J. Math. 35(1) (2009), 159–183.
[114] H. Martikainen and T. Orponen, Some obstacles in characterising the boundedness of bi-parameter singular integrals, Preprint (2014). arXiv:1404.2216.
[115] Yumeng Ou, Stefanie Petermichl, Elizabeth Strouse, Higher order Journé commutators and characterizations of multi-parameter BMO, Advances in Mathematics 291 (2016) 24–58.
[116] A. Brown, P. Halmos, Algebraic properties of Toeplitz operators, J. Reine Angew. Math. 213 (1964) 89–102.

[117] L. Carleson, A counterexample for measures bounded on Hpspaces for the bidisk, Mittag-Leffler Rep. No. 7, Inst. Mittag-Leffler, 1974.
[118] R. Coifman, P.-L. Lions, Y. Meyer, S. Semmes, Compensated compactness and Hardy space, J.Math. Pures Appl. (9) 72 (1993) 247–286.
[119] M. Cotlar, C. Sadosky, The Helson–Szegö theorem in L^p of the bidimensional torus, Contemp. Math. 107 (1990) 19–37.
[120] R. Fefferman, Singular integrals on product domains, Bull. Amer. Math. Soc. (N.S.) 4(2) (1981) 195–201.
[121] R. Fefferman, Harmonic analysis on product spaces, Ann. of Math. (2) 126(1) (1987) 109–130.
[122] M. Lacey, S. Petermichl, J. Pipher, B. Wick, Multi-parameter Div–Curl lemmas, Bull. Lond. Math. Soc. 10 (2012) 1123–1131.
[123] H. Martikainen, Representation of bi-parameter singular integrals by dyadic operators, Adv. Math. 229(3) (2012) 1734–1761.
[124] Y. Ou, A Tpbq theorem on product spaces, Trans. Amer. Math. Soc. 367 (2015) 6159–6197, http://dx.doi.org/10.1090/S0002-9947-2015-06246-1.
[125] Y. Ou, Multi-parameter singular integral operators and representation theorem, arXiv:1410.8055, pp. 1–28.
[126] S. Petermichl, Dyadic shifts and a logarithmic estimate for Hankel operators with matrix symbol, C. R. Math. Acad. Sci. Paris 1(1) (2000) 455–460.
[127] S. Petermichl, S. Treil, A. Volberg, Why the Riesz transforms are averages of the dyadic shifts?, Publ. Mat. (ISSN0214-1493)46(2) (2002) 209–228.
[128] A. Uchiyama, A constructive proof of the Fefferman Stein decomposition of BMOpRnq, Acta Math. 148 (1982) 215–241.
[129] M. S. BIRMAN AND M. Z. SOLOMJAK, PIECEWISE-POLYNOMIAL APPROXIMATIONS OF FUNCTIONS OF THE CLASSES W_p^α
[130] M. & Birman and M. Z. Solomjak, On double Stieltjes operator integrals, Dokl. Akad. Nauk SSSR 165 (1965), 1223-1226 = Soviet Math. Dokl. 6 (1965), 1567-1571. MR 36 #6234.
[131] , Double Stieltjes operator integrals. I, Probl. Math. Phys. No. I. Spectral Theory and Wave Processes, Izdat. Leningrad. Univ., Leningrad, 1966, pp. 33-67. MR 35 #767b.
[132] Double Stieltjes operator integrals and multiplier problems on multipliers, Dokl. Akad. Nauk SSSR 171 (1966), 1251-1254 = Soviet Math. Dokl. 7 (1966), 1618-1621. MR 35 #767a.
[133] Double Stieltjes operator integrals. II, Probl. Math. Phys. No. 2, Izdat. Leningrad. Univ. Leningrad, 1967, pp. 26-60.
[134] A. N. Kolmogorov and V. M. Tihomirov, e-entropy and (-capacity of sets in functional spaces, Uspehi Mat. Nauk 14 (1959), no. 2 (86), 3-86; English transl., Amer. Math. Soc. Transl. (2) 17 (1961), 277-364. MR 22 #2890; MR 23 #A2031.
[135] S. M. Nikol skiT, Imbedding, continuation and approximation theorems for differentiable functions of several variables, Uspehi Mat. Nauk 16 (1961), no. 5 (101), 63-114. (Russian) MR 26 #6757.
[136] M. S. Birman and M. Z. Solomjak, a) Estimates of singular numbers of integral operators. I, Vestnik Leningrad Univ. 22 (1967), no. 7, 43-53 MR 35 #7173 b) II, Vestnik Leningrad. Univ. 22 (1967), no. 13, 21-28. MR 36 #739

[137] Approximation of functions of the Wa classes by piecewise polynomial functions, Dokl. Akad. Nauk SSSR 171 (1966), 1015-1018 = Soviet Math. Dokl. 7 (1966), 1573-1577. MR 35 #630.
[138] V. M. Tihomirov, Diameters of sets in functional spaces and the theory of best approximations, Uspehi Mat. Nauk 15 (I960), no. 3 (93), 81-120 = Russian Math. Surveys 15 (I960), no. % 75-111. MR 22 #8268.
[139] G. Polya and G. Szego, Aufgaben und Lehrsatze aus der Analysis. I, 2nd ed., Die Grundlehren der math. Wissenschaften, Band 19, Springer-Verlag, Berlin, 1954; Russian transl., IL, Moscow, 1956. MR 15, 512; MR 30 #1219a.
[140] S. L. Sobolev, Applications of functional analysis in mathematical physics, Izdat. Leningrad. Gos. Univ., Leningrad, 1950; English transl., Transl. Math. Monographs, vol. 7, Amer. Math. Soc , Providence, R. I., 1963· MR 14, 565; MR 29 #2624.
[141] V. M. Tihomirov, A remark on η-dimensional diameters of sets in Banach spaces, Uspehi Mat. Nauk 20 (1965), no. 1 (121), 227-230. MR 30 #2278.
[142] Yu. Brudny, Nonlinear piecewise polynomial approximation and multivariate BV spaces of a Wiener–L. Young type. I., http://dx.doi.org/10.1016/j.jat.2017.03.002.
[143] A. Brudnyi and Yu. Brudnyi, Methods of Geometric Analysis in Extension and Trace Problems. Vol. I. Birkh¨auser. 2011.
[144] P. Bechler, R. Devore, A. Kamont, G. Petrova and P. Wojtaszczyk, Greedy wavelet projections are bounded on BV , Trans. AMS, 359 (2007), 619–635.
[145] Yu. Brudnyi and I. Irodova, Nonlinear spline–approximation and B-spaces. Proc. Intern. Conf. in Approx. Theory (Kiev, 1987). Nauka. M. 71–75.
[146] Yu. Brudnyi and N. Kruglyak, Interpolation Functions and Interpolation Methods. I. North–Holland. 1991.
[147] M. Birman and M. Solomyak, Piecewise polynomial approximation of functions of classes W_p^α , Math. Sb. 73 (1967), 331–355; English transl. in Math. USSR – Sb., 2 (1967), 295–317.
[148] Yu. Brudnyi, A multidimensional analog of a theorem of Whitney. Mat. Sb. 82 (1970), 169–191; English transl. in Math. USSR – Sb. 11 (1970).
[149] Yu. Brudnyi, Spaces defined by local approximation, Trudy Moscov. Mat. Obshch. 24 (1971), 69–132; English transl. in Trans. Moscow Math. Soc. 24 (1971).
[150] Yu. Brudnyi, Nonlinear approximation of refinable functions and piecewise polynomials. Appr. Theory X: Splines, Wavelets and Applications, Vanderbilt Univ. Press, Nashville. 2002.
[151] A. Cohen, R. DeVore, P. Petrushev and H. Xu, Nonlinear approximation and the space BV ($\mathbb{R}^2$), Amer. J. Math. 121 (1999), 587–628.
[152] R. DeVore and G.G. Lorentz, Constructive Approximation. Springer. 1993.
[153] F. Harary, J. Hayers and H.-J. Wu, A survey of the theory of hypercube graph, Comp. and Math. with Applications, 15 (4), (1988), 277–289.
[154] S. Nikolskii, Approximation of Functions of Several variables and Embedding Theorems. Springer. 1975.
[155] J. Peetre, New Thoughts on Besov spaces. Duke Univ. Math. Ser. I. 1976.
[156] P. Wojtaszczyk, Projections and nonlinear approximation in the space BV ($\mathbb{R}^d$), Proc. London Math. Soc. 87 (2003). 471–497.
[157] W. Ziemer, Weakly Differentiable Functions. Springer–Verlag. 1989.

[158] Liang Song, Lixin Yan, A maximal function characterization for Hardy spaces associated to nonnegative self-adjoint operators satisfying Gaussianestimates, Advances in Mathematics 287 (2016) 463–484.
[159] P. Auscher, On necessary and sufficient conditions for L^p-estimates of Riesz transforms associated to elliptic operators on $\mathbb{R}^n$ and related estimates, Mem. Amer. Math. Soc. 186(871) (2007).
[160] P. Auscher, A. McIntosh, E. Russ, Hardy spaces of differential forms on Riemannian manifolds, J.Geom. Anal. 18 (2008) 192–248.
[161] F. Bernicot, J. Zhao, New abstract Hardy spaces, J. Funct. Anal. 255 (2008) 1761–1796.
[162] F. Cacciafesta, P. D'Ancona, Weighted L^p estimates for powers of selfadjoint operators, Adv. Math. 229 (2012) 501–530.
[163] A. Calderón, An atomic decomposition of distributions in parabolic H^p spaces, Adv. Math. 25(3) (1977) 216–225.
[164] A. Calderón, A. Torchinsky, Parabolic maximal functions associated with a distribution, Adv. Math. 16 (1975) 1–64.
[165] R. Coifman, A real variable characterization of H^p, Studia Math. 51 (1974) 269–274.
[166] T. Coulhon, A. Sikora, Gaussian heat kernel upper bounds via Phragmén–Lindelöf theorem, Proc. Lond. Math. Soc. 96 (2008) 507–544.
[167] D.G. Deng, Y.S. Han, Theory of H^p Spaces, Peking Univ. Press, Beijing, 1992.
[168] X.T. Duong, J. Li, Hardy spaces associated to operators satisfying Davies–Gaffney estimates and bounded holomorphic functional calculus, J. Funct. Anal. 264 (2013) 1409–1437.
[169] J. Dziubański, J. Zienkiewicz, Hardy spaces H^1 associated with Schrödinger operator with potential satisfying reverse Hölder inequality, Rev. Mat. Iberoam. 15 (1999) 279–296.
[170] R.M. Gong, L.X. Yan, Weighted L^p estimates for the area integral associated to self-adjoint opera-tors, Manuscripta Math. 144 (2014) 25–49.
[171] S. Hofmann, G.Z. Lu, D. Mitrea, M. Mitrea, L.X. Yan, Hardy spaces associated to nonnegative self-adjoint operators satisfying Davies–Gaffneyestimates, Mem. Amer. Math. Soc. 214(1007) (2011).
[172] S. Hofmann, S. Mayboroda, Hardy and BMO spaces associated to divergence form elliptic operators, Math. Ann. 344 (2009) 37–116.
[173] S. Hofmann, S. Mayboroda, A. McIntosh, Second order elliptic operators with complex bounded measurable coefficients in L^p, Sobolev and Hardy spaces, Ann. Sci. Éc. Norm. Supér. 44 (2011) 723–800.
[174] R. Jiang, D. Yang, New Orlicz–Hardy spaces associated with divergence form elliptic operators, J.Funct. Anal. 258 (2010) 1167–1224.
[175] R. Jiang, D. Yang, Orlicz–Hardy spaces associated with operators satisfying Davies–Gaffney esti-mates, Commun. Contemp. Math. 13 (2011) 331–373.
[176] R.H. Latter, A decomposition of $H^p(\mathbb{R}^n)$ in terms of atoms, Studia Math. 62 (1977) 92–102.
[177] S.Z. Lu, Four Lectures on Real H^p Spaces, World Scientific Publishing Co., Inc., River Edge, NJ, 1995.

[178] Ji Li, Brett D. Wick, Characterizations of $H^1_{\Delta_N}(\mathbb{R}^n)$ and $\mathrm{BMO}_{\Delta_N}(\mathbb{R}^n)$ via weak factorizations and commutators
[179] S. Chanillo, A note on commutators, Indiana Univ. Math. J. 31 (1982) 7–16.
[180] L. Grafakos, Modern Fourier Analysis, 3, Grad. Texts in Math., vol.250, Springer, New York, 2014.
[181] W.A. Strauss, Partial Differential Equation: An Introduction, John Wiley & Sons, Inc., New York, 2008.
[182] A. Uchiyama, The factorization of H^p on the space of homogeneous type, Pacific J. Math. 92(2) (1981) 453–468.
[183] Shawgy Hussein and Nissreen Alzobir Osman, Journé Commutators and Polynomial Approximations with Characterizations of Hardy Space and BMO, Ph.D. Thesis Sudan University of Science and Technology, Sudan (2019).