Some Results in the Theory of Real Hardy Spaces and BMO

Shahaboddin Shaabani

A Thesis

in

The Department

of

Mathematics and Statistics

Presented in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy (Mathematics) at

Concordia University

Montreal, Quebec, Canada

May 2025

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Shahaboddin Shaabani

Entitled: Some Results in the Theory of Real Hardy Spaces and BMO

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Mathematics)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

____________________ Chair
Dr. Lata Narayanan

____________________ External Examiner
Dr. Loukas Grafakos

____________________ Arm's Length Examiner
Dr. Ronald Stern

____________________ Examiner
Dr. Maria Ntekoume

____________________ Examiner
Dr. Alexey Kokotov

____________________ Thesis Supervisor(s)
Dr. Galia Dafni

Approved by

____
Dr. Cody Hyndman, Graduate Program Director

June 25 2025
Date of Defence

____
Dr. Pascale Sicotte, Dean of Faculty of Arts and Science

ABSTRACT

Some Results in the Theory of Real Hardy Spaces and BMO

Shahaboddin Shaabani, Ph.D.

Concordia University, 2025

The results presented in this thesis concern several aspects of the theory of real Hardy spaces and BMO. The first problem we address is the extension problem for Hardy spaces, for which we provide a complete resolution. Specifically, we investigate: for which open subsets Ω can every element of $H^p(\Omega)$ be extended to an element of $H^p(\mathbb{R}^n)$ with comparable norms? We give a complete geometric characterization of such domains.

We then turn our attention to the behavior of maximal operators on the space BMO. In this context, we study four questions, including the discontinuity of the Hardy–Littlewood maximal operator, its boundedness on VMO, and the unboundedness of both the strong and directional maximal operators on BMO. This part concludes with a counterexample demonstrating the failure of the Fubini property for this class of functions.

The final part of the thesis focuses on paraproducts and their operator norms on Hardy spaces. We establish sharp lower bounds for the norms of these operators acting on various types of Hardy spaces, both in the one-parameter and multi-parameter settings. These results yield an alternative characterization of Hardy spaces as admissible symbol classes for such operators.

Acknowledgments

I would like to extend my sincerest gratitude to my supervisor, Professor Galia Dafni, for her unwavering support, guidance, countless hours of insightful discussions, and exceptional supervision—without which this work would not have been possible. She set me on the road and showed me the way. I am also thankful to the members of my examining committee—Prof. Lata Narayanan, Prof. Loukas Grafakos, Prof. Ronald Stern, Prof. Maria Ntekoume, and Prof. Alexey Kokotov—for generously giving their time to read my thesis, for their thoughtful comments, and for attending the defense.

Contents

1 Introduction **1**

1.1 Historical Background 1

1.2 Contributions of the Present Thesis 6

1.2.1 Extension Domains for Hardy Spaces 7

1.2.2 Maximal Operators on BMO 8

1.2.3 The Operator Norm of Paraproducts on Hardy Spaces 9

2 Extension Domains for Hardy Spaces **13**

2.1 Introduction 13

2.2 Notation and Definitions 14

2.3 Main result 19

2.4 Proof of Sufficiency 23

2.5 Proof of Necessity 32

2.6 Applications 38

3 Maximal Operators on BMO and Slices **40**

3.1 Introduction 40

3.2 Discontinuity of M on $BMO(\mathbb{R}^n)$ 44

3.3 The Uncentered Hardy-Littlewood Maximal Operator on $VMO(\mathbb{R}^n)$. 49

3.4 Slices of BMO functions and unboundedness of directional and strong maximal operators 53

4 The Operator Norm of Paraproducts on Hardy Spaces **61**

4.1 Introduction 61

4.2 preliminaries 64

4.2.1 Notation 64

4.2.2 Sparse Families 65

4.2.3 Dyadic Hardy Spaces 66

4.2.4 Dyadic Paraproducts 69

4.2.5 Littlewood-Paley Operators 72

4.2.6 Real Hardy Spaces 74

4.2.7 Fourier Paraproducts 76

4.3 The operator norm of Dyadic paraproducts on dyadic hardy spaces . 78

4.4 Fourier paraproducts 91

5 The Operator Norm of Paraproducts on Bi-parameter Hardy Spaces **120**

5.1 Introduction 120

5.2 preliminaries 123

5.2.1 Bi-parameter Dyadic Hardy Spaces 126

5.2.2 Bi-parameter Hardy spaces in the Continuous Setting 130

5.3 Bi-parameter Dyadic Paraproducts 133

5.4 Bi-Parameter Fourier Paraproducts 145

Bibliography 150

Chapter 1

Introduction

1.1 Historical Background

Hardy spaces have their roots in the study of the behavior of analytic functions near the boundary of a region. Let f be an analytic function in the open unit disk $\mathbb{D}$ of the complex plane. Then what can be said about the behavior of f near the boundary of $\mathbb{D}$? Does it converge to a value at all points? How about at almost every point, or in some specific norm such as the L^p-norm? If so, in which sense can the boundary be approached? With no restriction, radially, or something in between? Of course, without any restriction on the growth of f, one can see through very simple counterexamples that the answer to all these questions is negative. However, under a condition such as

$$\|f\|_{H^p(\mathbb{D})} := \sup_{0<r<1} \left(\int_0^{2\pi} |f(re^{i\theta})|^p d\theta \right)^{\frac{1}{p}} < \infty,$$

for some $0 < p \leq \infty$, one can build a fruitful theory with many applications to other fields of analysis. G. H. Hardy was the first to consider the above quantity [34], and the systematic study of the above class of functions was initiated by F. Riesz in 1923, who named them Hardy spaces [59]. The classical Hardy spaces $H^p(\mathbb{D})$ found their place very quickly in analysis, and many mathematicians addressed various types of questions regarding this scale of spaces in the first half of the twentieth century. Indeed, they became so popular that many were trying to find their analogue in several variables. One specific reason for this interest lies in the close relation between $H^p(\mathbb{D})$ and the theory of Fourier series, especially their conjugate functions. Indeed, it is not hard to show that in the reflexive range where $1 < p < \infty$, the class $H^p(\mathbb{D})$ is naturally isometric to $L^p(\partial\mathbb{D})$, and many questions about the latter can be understood as questions about the former. With more work, one can see that $H^p(\mathbb{D})$ behaves much better under the action of singular integral transformations such as the Hilbert transform, in the full range of exponents $0 < p < \infty$, as opposed to their ill behavior on $L^p(\partial\mathbb{D})$ for $0 < p \leq 1$.

For many years, attempts in finding a similar theory in higher dimensions were in vain. This was specifically due to the use of the "complex method," which was rather central to the $H^p(\mathbb{D})$ theory. A method based on analyzing the roots of holomorphic functions in the unit disk, factoring them out, and by using powers converting questions about $H^p(\mathbb{D})$ to ones about $H^2(\mathbb{D})$, where the rich geometry of Hilbert spaces is available. Unfortunately, such methods would fail in higher dimensions because of the complicated nature of zeros of holomorphic functions in several variables. The first major progress in this direction appeared in 1960, when E. M. Stein and G. Weiss came up with a brilliant theory of Hardy spaces in higher dimensions [67]. This new way of interpreting Hardy spaces was based on the theory of "generalized Cauchy-Riemann equations" in the upper half-space of $\mathbb{R}^{n+1}$, and was very successful, although in a limited range of exponents. Finally, in 1972, Stein and Fefferman found the correct formulation of the definition of Hardy spaces in all dimensions and for the full range of exponents [21]. Specifically, for $0 < p < \infty$, the space $H^p(\mathbb{R}^n)$ can be characterized as "boundary values" of harmonic functions, u, defined in the upper half-space of $\mathbb{R}^{n+1}$ and satisfying

$$\int_{\mathbb{R}^n} \sup_{t>0} |u(x,t)|^p dx < \infty.$$

As a byproduct, they showed that, similarly to the classical setting, $H^p(\mathbb{R}^n)$ is naturally isomorphic to $L^p(\mathbb{R}^n)$ in the reflexive range, stable under the action of smooth singular integrals, and can be used instead of $L^p(\mathbb{R}^n)$ in interpolation theorems. By applying this new theory, boundedness of certain Calderón-Zygmund operators was proved and sharper estimates for solutions to the wave equation were established. This new theory is referred to as the theory of real Hardy spaces due to the fact that $H^p(\mathbb{R})$ is the complexification of the real parts of functions in $H^p(\mathbb{R}^2_+)$, the Hardy space of holomorphic functions in the upper half-plane, $\mathbb{R}^2_+$. In the same paper [21], the celebrated duality theorem of Fefferman which was previously announced, was established. Fefferman's theorem states that the dual of $H^1(\mathbb{R}^n)$ is naturally isomorphic to $BMO(\mathbb{R}^n)$, the space of functions of bounded mean oscillation, which was introduced by F. John and L. Nirenberg earlier, in 1961 [41].

Similarly to the way that $H^1(\mathbb{R}^n)$ is a good replacement for $L^1(\mathbb{R}^n)$, the space $BMO(\mathbb{R}^n)$, in some respects, plays the role of $L^\infty(\mathbb{R}^n)$. For instance, singular integrals which are not bounded on $L^\infty(\mathbb{R}^n)$ are bounded on $BMO(\mathbb{R}^n)$, and this space can be used to interpolate between various Lebesgue spaces. Since the discovery of the "H^1–BMO" duality, many other relations have been revealed, including the atomic decomposition of Hardy spaces and the role of Carleson measures in harmonic analysis. In some sense, Carleson measures are geometric cousins of BMO functions, and this can be made precise via the use of smooth wavelets or almost-orthogonal expansions. This way, one gets a characterization of these spaces in terms of wavelets which is very geometric in nature. As is often said "harmonic analysis is infected by the Schwartz tails;" if one uses the non-smooth Haar wavelets, which are simpler to work with, one gets the dyadic counterparts of Hardy spaces and BMO. These dyadic model spaces are much "cleaner" and nowadays are an essential part of modern harmonic analysis. They have proved to be much more than a toy model, and through their analysis, one can see clearly the central role of martingales, and as a result, a satisfying theory of Hardy spaces and BMO can be developed in the general and abstract context of martingales with no difficulty.

In extremal cases, where the mass of a measure is concentrated at the centers of a collection of dyadic cubes, the Carleson measure condition is equivalent to the collection satisfying the so-called "Carleson packing condition." It turns out that this condition is equivalent to the "sparseness" of the family, which simply means that objects in the collection are allowed to overlap but they must have large portions which lie disjoint from each other. The innocent concept of spareness, which is a simple generalization of disjointness, has solidly found its place in harmonic analysis, and the method of "sparse domination" formulated by A. K. Lerner has proved to be very powerful. For instance, the A_2 conjecture, asking about the sharp dependence on the weight constant of the norm of singular operators on Lebesgue spaces equipped with a Muckenhoupt weight, was proved in a simple way by this method [49]. The previous proof of this conjecture was based on decomposing the operator in terms of the shift operators and is much more involved [39]. However, Lerner's proof is very simple, and ever since its discovery, it has been applied to a broad range of problems in a successful way. The sparse method has changed our understanding of the methods used previously, and almost all the previously found arguments can be derived from that, painlessly. It is also notable that although the sparse method has been formulated by Lerner, similar ideas have been used by P. W. Jones in the late 70s and early 80s in connection with his spectacular constructions in the study of BMO.

Contemplating similar questions to the ones raised before, but about boundary values of holomorphic functions in several variables defined on the product of some unit disks, brings us to the theory of multi-parameter or product Hardy spaces. The term "parameter" was coined by A. Zygmund in connection with differentiation properties of the basis of rectangles with sides parallel to fixed coordinate axes. In multi-parameter or product theory, one usually deals with operators with a tensor-product structure and commuting with a non-isotropic group of dilations. As a result, covering properties of collections of rectangles expanding or shrinking independently in several directions come into play. Due to the complicated nature of the way that these rectangles may overlap, the study of product Hardy spaces is more complicated than that of the usual or one-parameter ones. Nevertheless, the first step in the study of such problems was taken by R. Gundy and E. M. Stein in 1979 [33], and the theory was developed further by many including L. Carleson, A. Chang, R. Fefferman, J. L. Journé and G. Merryfield. The product theory is rather similar to the one described before, although the proofs are somewhat different. For instance, in the case of two parameters, again there is a notion of BMO called product BMO, denoted by $BMO(\mathbb{R} \otimes \mathbb{R})$, and as expected, it is the space dual to product $H^1(\mathbb{R} \otimes \mathbb{R})$. In addition, the notion of Carleson measure and its relation to product BMO extends with no difficulty. Also, one can prove the atomic decomposition for product Hardy spaces, and various characterizations of them extend to this new setting. However, as some counterexamples, specifically that of Carleson [14], have shown, the product theory can get very complicated, and in comparison to the classical theory, it is not well understood.

1.2 Contributions of the Present Thesis

After a quick introduction to the meandering theory of Hardy spaces, here we try to describe our contributions to the theory. Before doing so, we would like to recall a point of view from which one should see this manuscript. In the author's belief, it is rather a characteristic of harmonic analysis that the methods and ideas are more important than the results. Here, a non-expert can see an obsession with asking questions about the largest range of values of p for which a statement is true for L^p or H^p spaces, and it often happens that such statements are true for L^2. For instance, convergence of Fourier series in L^2 follows instantly from Plancherel, but a harmonic analyst is mostly interested in what happens in L^p. The reason here is that a proof which is true for the sharpest range of exponents uses almost the minimal structures required for the validity of that phenomenon, and in this way, ideas propagate through other theories. For instance, in the case of norm convergence of Fourier series, the proof idea handling the p case has proved to be very fruitful and is now called Calderón-Zygmund theory, with many applications to PDEs. There are many more examples, and we do not prolong the discussion here. For this reason, our emphasis in the present thesis is more on the methods than on the results, and below we will compare the previous methods to our new ones.

The thesis is based on four manuscripts by the author: Chapters 2-5 correspond to papers [62–65]. These concern different aspects of the theory, each of which is described below.

1.2.1 Extension Domains for Hardy Spaces

The Extension Problem

The study of Hardy spaces on open domains was initiated by Miyachi in [54], where, for a given proper open subset Ω of $\mathbb{R}^n$, a theory analogous to that of $H^p(\mathbb{R}^n)$ was developed for $H^p(\Omega)$. For a function space X that is well-defined over different open subsets of $\mathbb{R}^n$, a fundamental issue that arises in applications is the extension problem. That is, to ask: for exactly which open sets Ω does there exist a bounded extension operator from $X(\Omega)$ to $X(\mathbb{R}^n)$? This classical problem has a rich history and has been investigated for several function spaces, such as $W^{k,p}$, BMO, and others [42, 43]. For the case of H^p-extension domains, Miyachi provided a sufficient condition in [54], which, as we will see, is not a necessary one.

Theorem (Miyachi)**.** *Let Ω be a proper open subset of $\mathbb{R}^n$. Suppose there exists a constant $a > 1$ such that for every $x \in \Omega$ there is $x' \in \Omega^c$ with $d(x, x') < ad(x, \Omega^c)$ and $d(x', \Omega) > a^{-1}d(x, \Omega^c)$. Then there exists a bounded linear extension operator from $H^p(\Omega)$ to $H^p(\mathbb{R}^n)$.*

The Final Answer

Theorem 2.3.1, which is the main result of Chapter 2, gives a full description of these domains in terms of some sort of thickness condition for the complement of the domain near its boundary. In detail, except for the case where $n(\frac{1}{p} - 1)$ is not an integer and every open set is an extension domain for H^p, for $p = 1$, an open set Ω is an extension domain for H^1 precisely when there are two positive constants a and δ such that the density of Ω^c is at least δ on each ball of the form $B(x, ad(x, \Omega^c))$ with $x \in \Omega$, where $d(x, \Omega^c)$ is the distance from x to the complement of Ω. Also, for $p = \frac{n}{n+k}$, with $k = 1, 2, \ldots$, a similar characterization holds, although the word "density" must be replaced with "width." The width of a set is simply the smallest distance between all parallel hyperplanes in between which the set can be placed.

The reason for the sufficiency of such thickness conditions is that, as noticed by Miyachi, the main task in extending distributions in $H^p(\Omega)$ is to extend bounded functions whose support is located on Whitney cubes of Ω, the so-called (p, Ω)-atoms. Then, it is not difficult to see that when these conditions hold, the extension can be done by putting either a constant or a linear combination of derivatives of Dirac masses supported on that part of Ω^c which lies on the enlarged cube. On the other hand, the failure of these thickness conditions makes it possible to construct some functions in the dual of $H^p(\mathbb{R}^n)$, which are supported in Ω, have small norms, but large averages, showing that any bounded extension of (p, Ω)-atoms is impossible.

1.2.2 Maximal Operators on BMO

The Continuity Problem

Maximal operators are among the most important sublinear operators in harmonic analysis, and their boundedness properties on various function spaces are of great interest. Indeed, operators of this kind dominate the other ones, and many pointwise convergence results follow from their boundedness properties. Boundedness of the classical Hardy-Littlewood maximal operator M, on $BMO(\mathbb{R}^n)$, was already established in [3]. Another proof of this was given in [2], and a third one in [60]. However, in these works, the continuity problem was not addressed. The maximal operator is nonlinear, and for such operators, continuity is not equivalent to boundedness. Nevertheless, this operator is pointwise sublinear, and this makes it continuous on $L^p(\mathbb{R}^n)$ for $1 < p \leq \infty$. Moreover, as was shown in [52], the operator is continuous on the Sobolev space $W^{1,p}(\mathbb{R}^n)$ with $1 < p < \infty$.

Our New Results

In Chapter 3, we present three results regarding the action of maximal operators on $BMO(\mathbb{R}^n)$. First, we show that although the Hardy-Littlewood maximal operator is bounded on $BMO(\mathbb{R}^n)$, it is not continuous, which is the content of Theorem 3.2.1. Second, it is shown that the subspace of $BMO(\mathbb{R}^n)$ consisting of functions of vanishing mean oscillation, namely $VMO(\mathbb{R}^n)$, is preserved under the action of this operator, and this is done in Theorem 3.3.1. Third, in Theorem 3.4.1, we show that unlike the operator M, the strong and directional maximal operators are unbounded on $BMO(\mathbb{R}^n)$. Finally, as a by-product of the construction in this theorem, we show that $BMO(\mathbb{R}^n)$ fails to satisfy the Fubini property. More precisely, we give an example of a function in $BMO(\mathbb{R}^2)$ with the property that none of its slices belong to $BMO(\mathbb{R})$.

1.2.3 The Operator Norm of Paraproducts on Hardy Spaces

Significance of Paraproducts

Paraproducts are among the most important bilinear forms and arise naturally in many problems of harmonic analysis. Usually, they appear when working with the product of two functions and trying to understand the product from the Fourier point of view. The product consists of many intricate frequency interactions and must be divided into manageable pieces, each of which contains a certain type of interaction. Therefore, the product, which is a bilinear form itself, is divided into three different forms called paraproducts based on the low-high, high-low, or low-low frequency interactions. Some paraproducts behave much better than the product, and they satisfy a Hölder-type inequality in the full range of exponents, although with Lebesgue spaces replaced by Hardy spaces.

As F. Riesz taught us, one cannot do any better than Hölder's inequality when working with the product of functions on Lebesgue spaces, and we proved the same to be true when working with paraproducts on Hardy spaces. In detail, in our two recent works, we have complemented the existing results and proved a new characterization of BMO and Hardy spaces in the full range of exponents as admissible symbols of bounded paraproducts. This was first done in the one-parameter setting, where the tree structure of dyadic cubes is available, and later was generalized to the multi-parameter setting, in which case such a structure no longer exists. Below, we review the previously known results and compare them to our new ones.

Previous Results

The first known result on the operator norm of paraproducts appears in [7], where it is shown that for the dyadic paraproduct operator π_g (see (4.10)), it holds that

$$\|\pi_g\|_{L^p(\mathbb{R})\to L^p(\mathbb{R})} \simeq \|g\|_{BMO(\mathbb{R})}, \quad 1 < p < \infty.$$

In the above, the upper bound was already well known, but the lower bound was new. However, the idea of the proof is very simple: if the operator is tested against a suitable test function, the desired lower bound for the mean oscillation follows easily. Regarding the operator norm of π_g, the next result appeared in [36], where it was shown that

$$\|\pi_g\|_{L^p(\mathbb{R}^n)\to L^q(\mathbb{R}^n)} \simeq \|g\|_{\dot{L}^r(\mathbb{R}^n)}, \quad \frac{1}{q} = \frac{1}{p} + \frac{1}{r}, \quad 1 < q < p < \infty.$$

This time, the proof is more involved, and the idea in [36] is to use the local mean oscillation inequality [48, 50, 51] and obtain a sparse domination of the symbol g. Then, bounding the mean oscillation of g as in the previous case, along with duality, completes the proof.

New Results and Methods

Our main results in Chapter 4, namely Theorems 4.3.1 and 4.4.1, give a full description of the operator norm of dyadic and Fourier paraproducts for all values of the exponents. In this chapter, we study operators from H^p to H^q, and when $p \leq q$, our method is rather elementary and mostly the same as the one used in [7] and [36]. However, when $q < p$, and especially when $q < 1$, we cannot rely on duality arguments like the one used in [36], for the reason that the Hahn-Banach theorem does not hold any longer for H^q when $0 < q < 1$ [20], so the operator and its adjoint may not have the same norm. To deal with this, we use the sparse domination method in [50] and establish a suitable domination of the square function of the symbol g. Then, we will show that, with such domination, it is always possible to construct a random test function, which, after plugging into the operator and using boundedness, gives the desired result.

In addition, in Chapter 5, we extend the above-mentioned results for a type of paraproduct, with the same geometric structure, to the product setting. Theorems 5.3.1 and 5.4.1, which contain the main results of Chapter 5, are completely new, even though the ideas used in the proofs are rather similar to those used in Chapter 4. Here, the tree structure of the dyadic cubes, which is crucial to our previous arguments, is not present. Nevertheless, we employ a similar strategy and demonstrate how the previous one-parameter arguments can be modified to work in the multi-parameter setting. Here, we develop a sparse domination method which is based on contracting families of open sets and not families of cubes or rectangles. Another challenge to overcome here is to construct an analogue of characteristic functions of open sets belonging to the product Hardy spaces. Such functions must have two rather contradictory properties: they must behave like a constant on the set, while having a lot of cancellation in two perpendicular directions. However, as we will show, using a local "square-maximal" equivalence of the product Hardy spaces, captured by Brossard's inequality [10], makes such a construction always possible.

Chapter 2

Extension Domains for Hardy Spaces

2.1 Introduction

In general for a given function space X, the extension problem is to characterize open subsets $\Omega \subset \mathbb{R}^n$, such that there exists a bounded (linear or nonlinear) operator, extending every element of $X(\Omega)$ to an element of $X(\mathbb{R}^n)$. Such domains are called extension domains for X.

This problem has been studied for several classes of functions like Sobolev spaces $W^{k,p}$ [43], functions of bounded mean oscillation BMO [42], VMO [13], etc. Here, we're going to give an answer to the extension problem for Hardy spaces H^p when $0 < p \leq 1$. As it is well known, this scale of spaces serves as a good substitute for L^p when $0 < p \leq 1$ [66]. Unlike L^p, being in H^p is not just a matter of "size" but also a matter of "cancellation". Indeed, these distributions must have vanishing moments up to order $n(\frac{1}{p} - 1)$, and this makes the extension problem a bit nontrivial. In Theorem 2.3.1, which is the main result of this paper, we give a characterization of these domains, and below we explain the idea behind the proof.

By Miyachi's atomic decomposition theorem, the extension problem is reduced to extending certain elements of $H^p(\Omega)$, namely (p,Ω)-atoms. These are atoms without any cancellations and are supported on Whitney cubes of Ω. Then since our characterization requires Ω^c to be thick in some sense, in dilated Whitney cubes, we are able to create the required cancellation. When $p=1$, this can be done by putting a proper constant on Ω^c in that dilated cube, and for $p=\frac{n}{n+k}$, adding a suitable linear combination of derivatives of Dirac masses supported on Ω^c will do the job. The latter case involves interpolating data with polynomials of several variables, and we show that with the condition mentioned in Theorem 2.3.1, it is always possible to do that. Also in the absence of these thickness conditions, we are able to construct some functions in the dual of $H^p(\mathbb{R}^n)$, namely $BMO(\mathbb{R}^n)$ and $\dot{\Lambda}^k(\mathbb{R}^n)$, which are supported on Ω and have small norms yet large averages. This prevents any bounded extension of (p,Ω)-atoms. Finally, when $n(\frac{1}{p}-1)$ is not an integer, it can be shown that every open set is an extension domain.

2.2 Notation and Definitions

We begin by fixing some notation that we use throughout the paper. We use $d(x,H)$ for the distance of x from H, and $d(x)=d(x,\Omega^c)$. The space of smooth functions with compact support in Ω is denoted by $C_0^\infty(\Omega)$, distributions in Ω by $\mathcal{D}'(\Omega)$, and $\langle,\rangle$ is used for pairing distributions and test functions. $C^k(\mathbb{R}^n)$ is used for the space of functions with bounded continuous derivatives up to order k with the usual sup-norm, and $\|P\|_{C^k(E)}$ means that we take the sup-norms only on E. Also the space of polynomials of degree no more than k in $\mathbb{R}^n$ is denoted by $\mathbb{P}_k$, as usual ∂^α is used for partial derivatives, $|\alpha|$ for the total degree of α, and other famous notations is used as well. By a cube Q we mean a cube with sides parallel to the axis and we use $l(Q)$ for its side length. We use aQ or aB to denote the concentric dilation of a cube Q or a ball B. Also $\mathcal{W}(\Omega)$ is used for the family of Whitney cubes of Ω (for the definition of Whitney cubes see [29] App.J, or [54]). Moreover, the useful symbols $\lesssim$, $\gtrsim$ and $\approx$ are used to avoid writing unimportant constants usually depending on k and n.

Next, let us recall definitions of the dual of Hardy spaces $H^p(\mathbb{R}^n)$. The space of functions of bounded mean oscillation $BMO(\mathbb{R}^n)$ consists of all locally integrable functions f such that

$$\sup_Q ⨍_Q |f - ⨍_Q f| < \infty,$$

where in the above the sup is taken over all cubes Q. Modulo constants, $BMO(\mathbb{R}^n)$ is a Banach space with the above quantity as norm, and as it is well known it is the dual space of $H^1(\mathbb{R}^n)$ (see [25] Ch.3, [30] Ch.3, [66] Ch.4).

When $0 < p < 1$, the dual space of $H^p(\mathbb{R}^n)$ coincides with the homogeneous Lipschitz space $\dot{\Lambda}^{n(\frac{1}{p}-1)}(\mathbb{R}^n)$, which we now define. For a function f defined on $\mathbb{R}^n$ and $h \in \mathbb{R}^n$, the forward difference operator Δ_h is defined by $\Delta_h(f)(x) := f(x+h) - f(x)$, and for any positive integer k let $\Delta_h^k(f) := \Delta_h(\Delta_h ... (\Delta_h(f)))$ be the k-th iteration of Δ_h. Then for any positive real number γ, the space of homogeneous Lipschitz functions $\dot{\Lambda}^{\gamma}(\mathbb{R}^n)$ consists of all locally integrable functions f such that

$$\sup_{h \neq 0} \| \frac{\Delta_h^{[\gamma]+1}(f)}{|h|^{\gamma}} \|_{L^\infty(\mathbb{R}^n)} < \infty.$$

Modulo polynomials of degree no more than γ, $\dot{\Lambda}^{\gamma}(\mathbb{R}^n)$ is also a Banach space with the above quantity as norm, and as we already mentioned, for $0 < p < 1$ the dual space of $H^p(\mathbb{R}^n)$ is $\dot{\Lambda}^{n(\frac{1}{p}-1)}(\mathbb{R}^n)$ (see [25] Ch.3, [30] Ch.1). Here we want to make this clear that $\dot{\Lambda}^{\gamma}(\mathbb{R}^n)$ is referred to by different names. For instance, when $\gamma = 1$ it's called Zygmund space, or it is identical with the homogeneous Besov space $\dot{B}^{\gamma}_{\infty,\infty}(\mathbb{R}^n)$ (see [47] Ch.17).

Now we turn to the definition of Hardy spaces on domains. Let Ω be a proper open subset of $\mathbb{R}^n$ and f a distribution in $\mathcal{D}'(\Omega)$. Also, let φ be a smooth function supported in the unit ball of $\mathbb{R}^n$ such that $\int \varphi = 1$, and for $t > 0$, let $\varphi_t(x) = t^{-n}\varphi(\frac{x}{t})$. Then, the maximal function of f associated to Ω and φ, denoted by $M_{\Omega,\varphi}$, is defined by

$$M_{\Omega,\varphi}(f)(x) := \sup\{|\langle f, \varphi_t(x-.)\rangle| : 0 < t < \operatorname{dist}(x, \Omega^c)\} \quad \text{for} \quad x \in \Omega.$$

Definition 2.2.1. *For $0 < p \leq 1$, $H^p(\Omega)$ is the space of all distributions f on Ω such that $M_{\Omega,\varphi}(f)$ is in $L^p(\Omega)$. This defines a linear space, which is independent of φ and can be quasi normed (normed when $p = 1$) by setting $\|f\|_{H^p(\Omega)} = \|M_{\Omega,\varphi}(f)\|_{L^p(\Omega)}$. For different choices of φ, these quasi-norms are equivalent.*

The Hardy space $H^p(\Omega)$ was introduced by Miyachi in [54], and an atomic decomposition theorem was proved for it there. As a result, many fundamental properties of these spaces, including characterization of the dual space, density of $C_0^{\infty}(\Omega) \cap H^p(\Omega)$, complex interpolation and extension problem were studied although the latter was not fully resolved there. Meaning the following sufficient condition was obtained but as we will see it's not necessary.

Theorem 2.2.1 (Miyachi)**.** *Let Ω be a proper open subset of $\mathbb{R}^n$. Suppose there exists a constant $a > 1$ such that for every $x \in \Omega$ there is $x' \in \Omega^c$ with $d(x, x') < ad(x, \Omega^c)$ and $d(x', \Omega) > a^{-1}d(x, \Omega^c)$. Then there exists a bounded linear extension operator from $H^p(\Omega)$ to $H^p(\mathbb{R}^n)$.*

We continue this section by recalling definitions of two types of atoms. The first kind are the usual atoms appearing in the atomic decomposition of $H^p(\mathbb{R}^n)$.

Definition 2.2.2. *A bounded function f is called a p-atom, if there exists a ball B (the supporting ball of f) such that supp$(f) \subset B$, $\|f\|_{L^\infty} \leq |B|^{-\frac{1}{p}}$ (size condition), and $\langle f, P\rangle = 0$ for all polynomials with degree no more than $n(\frac{1}{p} - 1)$ (cancellation condition).*

However, there's another type of atom which appears in the atomic decomposition of $H^p(\Omega)$. These atoms do not need to have any kind of cancellations and are called (p, Ω)-atoms.

Definition 2.2.3. *Let Ω be a proper open subset of $\mathbb{R}^n$. We say that a bounded function g is a (p, Ω)-atom if it is supported on some Whitney cube Q of Ω, and $\|g\|_{L^\infty} \leq |Q|^{-\frac{1}{p}}$.*

In [54], it's shown that such atoms are elements of $H^p(\Omega)$ and are uniformly bounded in this space. After fixing these definitions let us recall the atomic decomposition theorem for Hardy spaces on domains.

Theorem 2.2.2 (Miyachi)**.** *Let Ω be a proper open subset of $\mathbb{R}^n$, $0 < p \leq 1$ and $f \in H^p(\Omega)$, then there exist sequences of nonnegative numbers $\{\lambda_i\}, \{\mu_j\}$, a sequence of p-atoms $\{f_i\}$, and a sequence of (p, Ω)-atoms $\{g_j\}$, such that the latter is depended linearly on f, each f_i is supported in Ω, and*

$$f = \sum \lambda_i f_i + \sum \mu_j g_j \quad \textit{with} \quad \|f\|_{H^p(\Omega)} \approx \left(\sum \lambda_i^p + \sum \mu_j^p\right)^{\frac{1}{p}}. \tag{2.1}$$

In the above, first a suitable Whitney decomposition of the domain is fixed and then roughly speaking, g_js are obtained by projecting f onto the space of polynomials localized on Whitney cubes of Ω. This is why g_j depends linearly on f, but this is not the case for f_js. Also, it follows from the construction that if f is compactly supported in Ω, then the support of all f_js and g_js are contained in a probably larger compact subset of Ω [54]. It should be noted here that even if f is a $(p,\Omega)-$atom, the above decomposition gives us some p-atoms and (p,Ω)-atoms which are probably different from f.

In [54], by using this powerful theorem, the extension problem for distributions in $H^p(\Omega)$ was reduced to the especial case of extending (p,Ω)-atoms. Although the following lemma is used there, but it's not stated, so we decided to state it here as a lemma.

Lemma 2.2.1. *For an open set Ω to be an H^p extension domain it is necessary and sufficient that for every (p,Ω)-atom g there exists an extension $\tilde{g} \in H^p(\mathbb{R}^n)$ such that $\|\tilde{g}\|_{H^p(\mathbb{R}^n)} \lesssim 1$. Moreover, if this extension is linear in terms of g then there is a bounded linear extension operator on $H^p(\Omega)$.*

Proof. The necessity of the above condition follows directly from the definition of extension domains and the fact that (p,Ω)-atoms have bounded norms.
For the sufficiency, take a distribution $f \in H^p(\Omega)$, decompose it into atoms as in Theorem 2.2.2 and extend it in the following way:

$$\tilde{f} = \sum \lambda_i \tilde{f}_i + \sum \mu_j \tilde{g}_j,$$

where $\tilde{f}_i$ is the extension of f_i by considering it to be zero outside of Ω. Since $\mathrm{supp}(f_i) \subset \Omega$, $\tilde{f}$ is a well-defined distribution which belong to $H^p(\mathbb{R}^n)$ and its norm is bounded by

$$\|\tilde{f}\|^p_{H^p(\mathbb{R}^n)} \leq \sum \lambda_i^p \|\tilde{f}_i\|^p_{H^p(\mathbb{R}^n)} + \sum \mu_j^p \|\tilde{g}_j\|^p_{H^p(\mathbb{R}^n)} \lesssim \sum \lambda_i^p + \sum \mu_j^p \lesssim \|f\|^p_{H^p(\Omega)}.$$

So this defines a bounded extension operator on $H^p(\Omega)$. Whenever the extension of (p,Ω)-atoms is linear, the above operator is also linear on compactly supported distributions in $H^p(\Omega)$. To see this, we note that $\sum \mu_j \tilde{g}_j$ is linear in terms of f and $\sum \lambda_i \tilde{f}_i = (f - \sum \mu_j g_j)\chi_\Omega$, which is true because the support of f, f_is and g_js all are contained in a compact subset of Ω. This implies that $\sum \lambda_i \tilde{f}_i$ is also linear in f. Now, the operator $f \to \tilde{f}$ is densely defined and bounded, so it has a unique extension to $H^p(\Omega)$, which is the desired linear extension operator.

□

2.3 Main result

Definition 2.3.1. *Let $E \subset \mathbb{R}^n$ and ν be a unit vector, then the width of E in direction ν, $w(E, \nu)$, is defined as the smallest number $\delta > 0$ such that E lies entirely between two parallel hyperplanes with normal ν and of distance δ from each other. Also we define the width of E, $w(E)$, to be*

$$w(E) := \inf_{|\nu|=1} w(E, \nu).$$

When $n = 1$, by $w(E)$ we simply mean the diameter of E.

Theorem 2.3.1. *Let $0 < p \leq 1$ then a proper open subset $\Omega \subset \mathbb{R}^n$ is an H^p extension domain if and only if*

- *For $p = 1$ there exist constants $a > 1$ and $\delta > 0$, such that for every $x \in \Omega$*

$$\frac{|\Omega^c \cap B(x, ad(x))|}{|B(x, ad(x))|} > \delta. \tag{2.2}$$

- *For $p = \frac{n}{k+n}$, $k = 1, 2, \ldots$, there exist positive constants $a > 1$ and $\delta > 0$ such that for every $x \in \Omega$,*

$$\frac{w(\Omega^c \cap B(x, ad(x)))}{d(x)} > \delta. \tag{2.3}$$

- *And if* $p \neq \frac{n}{k+n}$, $k = 0, 1, 2, 3...$ *then every proper open subset is an extension domain.*

Moreover in all cases the extension operator can be taken to be linear.

In the above, the condition (2.2) is stronger than (2.3) because if (2.2) holds and $\Omega^c \cap B(x, ad(x))$ lies between two parallel hyperplanes of distance $\delta' d(x)$, then

$$|\Omega^c \cap B(x, ad(x))| \leq \delta' d(x)(2ad(x))^{n-1},$$

which implies that $\delta' \geq 2^{1-n} c_n a \delta$, where c_n is the volume of the unit ball of $\mathbb{R}^n$. So (2.3) holds with the same a and δ replaced by $2^{1-n} c_n a \delta$. Also the example below shows that these two conditions are different.

For a closed unit cube $Q = I_1 \times ... \times I_n$ in $\mathbb{R}^n$, let C_j be the usual Cantor middle third subset of I_j and call $C_1 \times ... \times C_n$ the Cantor dust in Q. Now consider the grid of unit cubes in $\mathbb{R}^n$ with vertices on $\mathbb{Z}^n$, and from each cube in this grid remove its Cantor dust and call the remaining set Ω. For $n \geq 2$, Ω is connected and open with $|\Omega^c| = 0$ so it doesn't satisfy (2.2) for any choice of a and δ but it satisfies (2.3). To see this take $x = (x_1, ..., x_n) \in \Omega$ then it lies in some cube Q of the grid and without the loss of generality assume $Q = [0, 1]^n$. Now suppose $d(x, \Omega^c) \approx \max_{1 \leq j \leq n} |x_j - y_j| = |x_1 - y_1|$ for some $y = (y_1, ..., y_n)$ with all $y_j \in C$, where C is the Cantor set in $[0, 1]$. From the ternary expansion of x_j and y_j, it's clear that $|x_j - y_j| \approx 3^{-k_j}$ where k_j is the first place that their digits are different. So if in the expansion of y_j we change the k_1-th digit from 0 to 2 or from 2 to 0, we get another point y'_j with $|y_j - y'_j| \approx 3^{-k_1}$. Then note that the cube $Q' = [y_1, y'_1] \times ... \times [y_j, y'_j]$ has all its vertices on Ω^c and $d(x, Q') \approx d(x, \Omega^c) \approx l(Q')$. This shows that in every direction, the width of Ω^c in a dilation of $B(x, d(x, \Omega^c))$ is comparable to $d(x, \Omega^c)$ and (2.3) holds (see figure below).

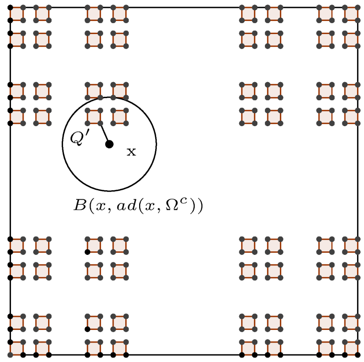

Figure 2.1: The white area is in $\Omega \cap Q$ and black dots belong to Ω^c

The condition (2.3) in Theorem 2.3.1 is equivalent to the geometric condition that characterizes "sets preserving the global Markov inequality". A closed set $F \subset \mathbb{R}^n$ is said to preserve the global Markov inequality if for every positive integer k and every ball $B_r(y)$ with $y \in F, r > 0$, we have

$$\|\nabla P\|_{L^\infty(F \cap B_r(y))} \leq c(F,k) r^{-1} \|P\|_{L^\infty(F \cap B_r(y))}, \quad P \in \mathbb{P}_k, \tag{2.4}$$

where $c(F,k)$ is a constant depending only on F and k. It turns out that such sets can be characterized in terms of the width. More specifically, F preserves the global Markov inequality if and only if there exists a constant $\varepsilon > 0$ such that for all $y \in F, r > 0$ we have

$$\frac{w(F \cap B_r(y))}{r} > \varepsilon \tag{2.5}$$

(see [44] Theorem 2 p. 38, and [11] Chapter 9 Section 9.2). In the next proposition we will show that Ω satisfies (2.3) exactly when Ω^c preserves the global Markov inequality.

Proposition 2.3.1. *A proper open set $\Omega \subset \mathbb{R}^n$ satisfies condition (2.3) if and only if Ω^c preserves the global Markov inequality.*

Proof. First suppose Ω^c preserves the global Markov inequality hence for all $y \in \Omega^c$ and $r > 0$, (2.5) holds with $F = \Omega^c$ and some $\varepsilon > 0$. Now for an arbitrary $x \in \Omega$ we consider $B(x, 2d(x))$ and choose $y \in \Omega^c \cap B(x, 2d(x))$ such that $|y - x| = d(x)$. Then we note that $B(y, d(x)) \subset B(x, 2d(x))$ and this together with (2.5) gives us

$$w(B(x, 2d(x)) \cap \Omega^c) \geq w(B(y, d(x)) \cap \Omega^c) > \varepsilon d(x),$$

which proves that (2.3) holds for Ω with $a = 2$ and $\delta = \varepsilon$.

To prove the other direction, suppose (2.3) holds for Ω with some $a > 1$ and $\delta > 0$. Now for an arbitrary $y \in \Omega^c$ and $r > 0$ we consider the sphere

$$S(y, \frac{r}{1+a}) = \left\{ z \in \mathbb{R}^n : |z - y| = \frac{r}{1+a} \right\}.$$

Then we note that if $S(y, \frac{r}{1+a}) \subset \Omega^c$ we have

$$w(B(y, r) \cap \Omega^c) \geq w(S(y, \frac{r}{1+a})) = \frac{2r}{1+a},$$

which shows that (2.5) holds with $\varepsilon = \frac{2}{1+a}$. So it remains to assume $S(y, \frac{r}{1+a}) \cap \Omega \neq \emptyset$, in which case we set

$$\varepsilon_0 = \sup \left\{ d(z) : z \in \Omega \cap S(y, \frac{r}{1+a}) \right\},$$

and consider two cases $\varepsilon_0 < \frac{r}{2(1+a)}$ and $\varepsilon_0 \geq \frac{r}{2(1+a)}$. In the first case, for each $z \in S(y, \frac{r}{1+a})$ there exists some $y_z \in B(y, r) \cap \Omega^c$ such that $|y_z - z| < \frac{r}{2(1+a)}$. Therefore, in the direction $\nu = \frac{z-y}{|z-y|}$ we have

$$|(y - y_z) \cdot \nu| \geq |y - z| - |y_z - z| > \frac{r}{2(1+a)}.$$

This shows that $w(\Omega^c \cap B(y, r)) \geq \frac{r}{2(1+a)}$ and hence (2.5) holds with $\varepsilon = \frac{1}{2(1+a)}$. Finally, when $\varepsilon_0 \geq \frac{r}{2(1+a)}$ there is $x \in S(y, \frac{r}{1+a}) \cap \Omega$ such that $\frac{r}{2(1+a)} \leq d(x) \leq \frac{r}{1+a}$.

Noting that $B(x, ad(x)) \subset B(y, r)$ together with (2.3) gives us

$$w(\Omega^c \cap B(y,r)) \geq w(\Omega^c \cap B(x, ad(x))) > \delta d(x) \geq \delta \frac{r}{2(1+a)},$$

which shows that (2.5) holds with $\varepsilon = \frac{\delta}{2(1+a)}$. So in any case the condition (2.5) holds with $\varepsilon = \frac{\min(1,\delta)}{2(1+a)}$, and this completes the proof.

□

We break the proof of Theorem 2.3.1 into two parts, the sufficiency and the necessity. In the next section, we will see it is easy to show that (2.2) is enough for extending each $(1, \Omega)$-atom to some 1-atom in $\mathbb{R}^n$. To prove that (2.3) is sufficient, we need to show that it allows us to choose an interpolation problem and solve it properly, which is the content of Lemma 2.4.4. This minor difficulty comes from the fact that when $n(\frac{1}{p} - 1) \in \mathbb{N}$, the order of distributions in H^p can not exceed than $n(\frac{1}{p} - 1) - 1$, while they must have vanishing moments up to order $n(\frac{1}{p} - 1)$. This becomes more clear when it is compared to the case $n(\frac{1}{p} - 1) \notin \mathbb{N} \cup \{0\}$.

2.4 Proof of Sufficiency

We begin with the following lemma, which looks like a reverse Markov's inequality (2.4) and is already known. Nevertheless, for the convenience of the reader we give a proof of this here (see [44] Remark 1 p. 35, and [11] Theorem 9.21).

Lemma 2.4.1. *Suppose E is a subset of the unit ball B of $\mathbb{R}^n$. Then for any $k = 1, 2, ..,$ we have*

$$\|P\|_{C^{k-1}(E)} > c(n,k) w(E) \|P\|_{L^\infty(B)}, \quad P \in \mathbb{P}_k. \tag{2.6}$$

Proof. Suppose $P \in \mathbb{P}_k$ with $\|P\|_{L^\infty(B)} = 1$ but $\|P\|_{C^{k-1}(E)} \leq cw(E)$. We will show that if $c = c(n,k)$ is chosen sufficiently small, the set E lies entirely between two parallel hyperplanes at distance less than $\frac{1}{2}w(E)$, which is a contradiction. Take $x_0 \in E$ and write the Taylor expansion of P around this point then we have

$$P(x) = \sum_{|\alpha|\le k-1} \frac{\partial^\alpha P(x_0)}{\alpha!}(x-x_0)^\alpha + \sum_{|\alpha|=k} \frac{\partial^\alpha P(x_0)}{\alpha!}(x-x_0)^\alpha,$$

which implies the following inequality:

$$\|P\|_{L^\infty(B)} \le c'(n,k)cw(E) + c'(n,k)\max_{|\alpha|=k}|\partial^\alpha P(x_0)|.$$

Noting that $w(E) \le 2$, and for the moment choosing $c \le \frac{1}{4c'}$, we get $|\partial^\alpha P(x_0)| \ge \frac{1}{2c'}$ for some α with $|\alpha| = k$. Now suppose $\alpha = (\alpha_1, ..., \alpha_n)$ with $\alpha_i = \max_{1\le j\le n} \alpha_j$ and let $\alpha' = (\alpha_1, .., \alpha_i - 1.., \alpha_n)$. Then we note that in the derivative $\partial^{\alpha'} P$, x_i appears in one and only one term with the coefficient $\partial^\alpha P(x_0)$. So the polynomial $\partial^{\alpha'} P(x)$ has the form

$$\partial^{\alpha'} P(x) = a \cdot x + b, \quad |a| \ge \frac{1}{2c'},$$

and if we normalize it by setting $\nu = \frac{a}{|a|}$ and $b' = \frac{b}{|a|}$, for all $x \in E$ we get

$$|\nu \cdot x + b'| \le 2c'cw(E),$$

which means that $w(E,\nu) \le 4c'cw(E)$. So if we take $c(n,k) = \frac{1}{8c'}$ we get the desired contradiction and this finishes the proof. □

In the next two lemmas we show that how (2.6) can be relaxed so that we can replace $\|P\|_{C^{k-1}(E)}$ by a smaller quantity $\max_{(x,\beta)\in F} |\partial^\beta P(x)|$, where F is a set of pairs of points and indices with $\#F = \dim(\mathbb{P}_k)$. The first one is a general geometric result bounding the width of a compact set in $\mathbb{R}^d$ with the width of its finite subsets with $d+1$ points.

Lemma 2.4.2. *Let E be a bounded subset of $\mathbb{R}^d$ and $\varepsilon > 0$ such that for each d points*

$x_1, x_2, ..., x_d$ in E, there exists a unit vector ν satisfying $|\nu \cdot x_i| \leq \varepsilon$ for $1 \leq i \leq d$. Then there is a unit vector ν_0 such that $|\nu_0 \cdot x| \leq 2d\varepsilon$ for every $x \in E$.

Proof. First we note that the assumption on E holds for $\bar{E}$ too, which allows us to assume E is compact. This implies that there are $x_1, x_2, ..., x_{d-1} \in E$ such that the $(d-1)$-dimensional area of the simplex $\Delta(0, x_1, ..x_{d-1})$ is maximal. Let $A = |\Delta(0, x_1, ..x_{d-1})|$ and note that if $A = 0$, E entirely lies in a $d-2$ dimensional subspace and ν_0 can be taken as a unit vector perpendicular to that. Then suppose $A > 0$ and let ν_0 be the normal vector to $H = span(x_1, ..., x_{d-1})$. We prove this vector has the desired property, which means that for any $x \in E$ we have to show that $|\nu_0 \cdot x| \leq 2d\varepsilon$ or equivalently $d(x, H) \leq 2d\varepsilon$.

Now from the assumption there is a unit vector ν such that $|\nu \cdot x_i| \leq \varepsilon$, $|\nu \cdot x| \leq \varepsilon$, which implies that the d dimensional simplex $\Delta(0, x, x_1, ..., x_{d-1})$ lies between two parallel hyperplane of distance 2ε from each other. Consequently, the length of one of the altitudes of this simplex must be no more than 2ε. Let us call the length of this altitude h and the area of its corresponding face A'. There are two cases, either 0 is one of the vertices of this face or not. In the former case $A' \leq A$, by maximality of A, and in the latter, $A' \leq dA$. The reason for this is the fact that the area of each face of a simplex is no more than the sum of areas of other faces, and all the other faces have 0 as a vertex. Now that we know $A' \leq dA$, we calculate the d dimensional volume of the simplex $\Delta(0, x, x_1, ..., x_{d-1})$ in two ways with different altitudes and faces, and we get $d(x, H)A = hA' \leq 2\varepsilon.dA$. Hence $d(x, H) \leq 2d\varepsilon$ and for each $x \in E$, $|\nu_0 \cdot x| \leq 2d\varepsilon$. □

Lemma 2.4.3. *Let k be a positive integer, B the unit ball of $\mathbb{R}^n$ and $E \subset B$ satisfying*

$$\|P\|_{C^{k-1}(E)} > \delta \|P\|_{L^\infty(B)}, \quad P \in \mathbb{P}_k. \tag{2.7}$$

Then there exists a set $F \subset \{(x,\beta)|x \in E, |\beta| \leq k-1\}$ *with* $\#F = dim(\mathbb{P}_k)$ *such that*

$$\max_{(x,\beta)\in F} |\partial^\beta P(x)| > c(n,k)\delta \|P\|_{L^\infty(B)}, \quad P \in \mathbb{P}_k. \tag{2.8}$$

Proof. For $P \in \mathbb{P}_k$ with $P(y) = \sum_{|\alpha|\leq k} c_\alpha y^\alpha$ let $\hat{P} = (c_\alpha)_{|\alpha|\leq k}$. Here, $\hat{P}$ is a d-dimensional vector with $d = \dim\mathbb{P}_k$, and we note that this defines an isomorphism between $\mathbb{P}_k$ and $\mathbb{R}^d$. Then from the finiteness of the dimension we have $\|\hat{P}\| \approx \|P\|_{L^\infty(B)}$, where $\|\hat{P}\|$ is the Euclidean norm in $\mathbb{R}^d$. So in (2.7) we may replace $\|P\|_{L^\infty(B)}$ by $\|\hat{P}\|$, and δ by $c'(n,k)\delta$. Then for each $x \in E$ and $|\beta| \leq k-1$ we set $(x,\hat{\beta}) = (b_\alpha)_{|\alpha|\leq k}$, where this vector is defined by

$$b_\alpha = \begin{cases} \frac{\alpha!}{(\alpha-\beta)!} x^{\alpha-\beta} & \beta \leq \alpha \\ 0 & \text{otherwise} \end{cases},$$

and it satisfies

$$\partial^\beta P(x) = \left\langle \hat{P}, (x,\hat{\beta}) \right\rangle, \quad P \in \mathbb{P}_k.$$

Now consider the set

$$\hat{E} = \left\{ (x,\hat{\beta}) | x \in E, |\beta| \leq k-1 \right\},$$

which is bounded in $\mathbb{R}^d$, and note that if for every d points like $Q_1, .., Q_d$ in $\hat{E}$ there exists a unit vector $\hat{P}$ with $|\left\langle \hat{P}, Q_i \right\rangle| \leq \frac{c'(k,n)}{2d}\delta$, then according to Lemma 2.4.2 there is a unit vector $\hat{P}$ such that $|\left\langle \hat{P}, Q \right\rangle| \leq c'(k,n)\delta$ for each $Q \in \hat{E}$. This means that $\|P\|_{C^{k-1}(E)} \leq c'(k,n)\delta$, which is a contradiction to (2.7). So there exists d points $Q_1, .., Q_d$ in $\hat{E}$ such that

$$\inf_{\|\hat{P}\|=1} \max_{1\leq i\leq d} |\left\langle \hat{P}, Q_i \right\rangle| > \frac{c'(k,n)}{2d}\delta,$$

and this means that there is a set $F \subset \{(x,\beta)|x \in E, |\beta| \leq k-1\}$ with $\#F = \dim(\mathbb{P}_k)$

such that

$$\max_{(x,\beta)\in F} |\partial^\beta P(x)| > \frac{c'(k,n)}{2d}\delta\|\hat{P}\|, \quad P \in \mathbb{P}_k.$$

Now it's enough to replace $\|\hat{P}\|$ by $\|P\|_{L^\infty(B)}$ and get (2.8). □

Lemma 2.4.4. *Let E be a subset of B, the unit ball of $\mathbb{R}^n$, with $w(E) > 0$. Then there exists a set $F \subset \{(x,\beta)|x \in E, |\beta| \le k-1\}$ with $\#F = dim(\mathbb{P}_k)$, and a basis of polynomials $\{P_{(x,\beta)}\}$ for $\mathbb{P}_k$ such that for all $(x',\beta') \in F$ we have*

$$\partial^{\beta'}(P_{(x,\beta)})(x') = \begin{cases} 1 & (x',\beta') = (x,\beta) \\ 0 & (x',\beta') \ne (x,\beta) \end{cases} \tag{2.9}$$

Moreover,

$$\|P_{(x,\beta)}\|_{L^\infty(B)} \lesssim w(E)^{-1} \quad (x,\beta) \in F. \tag{2.10}$$

Proof. From Lemmas 2.4.1 and 2.4.3 it follows that there exists a set F with $\#F = \dim(\mathbb{P}_k)$ such that for every $P \in \mathbb{P}_k$ we have

$$\|P\|_{L^\infty(B)} \lesssim w(E)^{-1} \max_{(x,\beta)\in F} |\partial^\beta P(x)|.$$

This implies that the mapping $P \longrightarrow (\partial^\beta P(x))_{(x,\beta)\in F}$ is an isomorphism from $\mathbb{P}_k$ to $\mathbb{R}^{\dim(\mathbb{P}_k)}$, therefore for each $(x,\beta) \in F$, (2.9) has a unique solution $P_{x,\beta}$. These polynomials satisfy (2.10), and since $\#F = \dim(\mathbb{P}_k)$, they form a basis for $\mathbb{P}_k$, and this finishes the proof. □

Our last lemma in this section is a simple generalization of boundedness of norms of atoms in $H^p(\mathbb{R}^n)$ (see [66] Ch.3 Further results 5.5).

Lemma 2.4.5. *Suppose B is the unit ball of $\mathbb{R}^n$, f a distribution supported in B and of the form $f = g + \sum_{(x,\beta)\in F} c_{x,\beta}\partial^\beta\delta_x$ with $\|g\|_{L^\infty} \le 1$, $|c_{x,\beta}| \le 1$, and F a nonempty* *finite set of pairs* (x, β) *such that* $x \in B$ *and* $|\beta| < n(1/p-1)$*. Moreover, suppose all moments of* f *vanishes up to order* $N_p = [n(1/p-1)]$*. Then* $f \in H^p(\mathbb{R}^n)$ *and* $\|f\|_{H^p(\mathbb{R}^n)} \lesssim \#F$*.*

Proof. Let $\varphi \in C_0^\infty(B)$, $0 \leq \varphi \leq 1$ with $\int \varphi = 1$. We estimate $M_\varphi(f)$ by following the usual local-nonlocal argument. For $x_0 \in 3B$ we have

$$\varphi_t * f(x_0) = \varphi_t * g(x_0) + \sum_{(x,\beta)\in F} (-1)^\beta c_{x,\beta} t^{-n-|\beta|} (\partial^\beta \varphi)(\frac{x_0 - x}{t}) = \mathrm{I} + \mathrm{II}.$$

Now for $t > 0$, $|\mathrm{I}| \leq 1$, and we note that in the second term $(\partial^\beta \varphi)(\frac{x_0-x}{t}) \neq 0$ only if $t > |x - x_0|$, so this term is bounded by

$$|\mathrm{II}| \lesssim \sum_{(x,\beta)\in F} |x - x_0|^{-n-|\beta|}.$$

This gives us an estimate for the local part and it remains to estimate $M_\varphi(f)$ on $(3B)^c$, where we have to use cancellation properties of f. So this time assume $x_0 \notin 3B$ and note that since $\varphi_t(x_0 - \cdot)$ is supported in $B_t(x_0)$, $f * \varphi_t(x_0) \neq 0$ only if $t > |x_0|/2$. To use the cancellation properties of f, let P be the Taylor polynomial of $\varphi_t(x_0 - \cdot)$ around 0 and of order N_p. Then we have

$$\begin{aligned} &f * \varphi_t(x_0) = \langle f, \varphi_t(x_0 - \cdot)\rangle = \langle f, \varphi_t(x_0 - \cdot) - P\rangle \\ &= \int g(y)(\varphi_t(x_0 - y) - P(y))dy + \sum_{(x,\beta)\in F} (-1)^\beta c_{x,\beta} \partial^\beta (\varphi_t(x_0 - y) - P(y))|_{y=x} \\ &= \mathrm{I} + \mathrm{II}. \end{aligned}$$

Now from the Taylor remainder theorem it follows that for $y \in B$ we have

$$|\varphi_t(x_0 - y) - P(y)| \leq \|\nabla^{N_p+1}(\varphi_t)\|_{L^\infty} \lesssim t^{-(N_p+1+n)},$$

therefore $|\mathrm{I}| \lesssim |x_0|^{-(N_p+1+n)}$. To estimate the second term we note that $\partial^\beta P$ is the $(N_p - |\beta|)$-th order Taylor polynomial of $\partial^\beta(\varphi_t(x_0 - \cdot))$ around 0, and this gives us

$$|\partial^\beta(\varphi_t(x_0 - y) - P(y))| \lesssim \|\nabla^{N_p - |\beta| + 1}(\partial^\beta(\varphi_t))\|_{L^\infty} \lesssim t^{-(N_p+1+n)} \lesssim |x_0|^{-(N_p+1+n)},$$

so we have $|\mathrm{II}| \lesssim |x_0|^{-(N_p+1+n)}$. Next, by bringing the estimates for the local and nonlocal parts together we obtain

$$M_\varphi(f)(x_0) \lesssim \left(1 + \sum_{(x,\beta)\in F} |x - x_0|^{-n-|\beta|}\right) \mathbb{1}_{3B}(x_0) + |x_0|^{-(N_p+1+n)} \mathbb{1}_{(3B)^c}(x_0).$$

Now using the above estimate and our assumptions on β and N_p gives us

$$\|f\|_{H^p(\mathbb{R}^n)} = \|M_\varphi(f)\|_{L^p(\mathbb{R}^n)} \lesssim \#F,$$

which completes the proof. □

With the above lemmas in hand we are ready to prove the sufficiency part of the Theorem 2.3.1.

Proof of sufficiency. According to Lemma 2.2.1 it is enough to extend each (p, Ω)-atom to an element of $H^p(\mathbb{R}^n)$ with a uniform bound on its norm.

Case $p = 1$. Take a $(1, \Omega)-$ atom g supported on a ball $B_r(x_0)$, with $r \approx d(x_0)$. Now, from the assumption there are positive constants a and δ such that (2.2) holds. Then we extend g by putting a constant on $\Omega^c \cap B_{ar}(x_0)$, to create the required cancellation. So let

$$\tilde{g} = g - c.\mathbb{1}_{\Omega^c \cap B_{ar}(x_0)}, \quad c = \frac{\int g}{|\Omega^c \cap B_{ar}(x_0)|}.$$

Here, we note that $\text{supp}(\tilde{g}) \subset B_{ar}(x_0)$, $\int \tilde{g} = 0$ and this function is bounded by

$$\|\tilde{g}\|_{L^\infty} \leq \|g\|_{L^\infty} \left(1 + a^{-n} \frac{|B_{ar}(x_0)|}{|\Omega^c \cap B_{ar}(x_0)|}\right) \lesssim \|g\|_{L^\infty} \left(1 + a^{-n}\delta^{-1}\right) \lesssim_{a,\delta} |B_{ar}(x_0)|^{-1}.$$

Hence $\tilde{g}$ is a constant multiple of a 1-atom so $\|\tilde{g}\|_{H^1(\mathbb{R}^n)} \leq c(a, \delta, n)$, and this finishes the proof of this case.

Case $p = \frac{n}{k+n}$, $k = 1, 2, \ldots$. Let g be a (p, Ω)-atom similar to the previous case, and let $a > 1$, $\delta > 0$ be the constants that satisfy (2.3). By a translation and dilation, we bring everything to the unit ball and set

$$E = \frac{\left(\Omega^c \cap \bar{B}_{ar}\right) - x_0}{ar}, \quad g' = g(x_0 + arx).$$

Now E is a subset of the unit ball B with $w(E) > a^{-1}\delta$, so let F be a set provided by Lemma 2.4.4. Then we extend g' by

$$\widetilde{g'} = g' + \sum_{(x,\beta)\in F} c_{x,\beta} \partial^\beta \delta_x,$$

where the coefficient $c_{x,\beta}$ are chosen in way that $\widetilde{g'}$ has vanishing moments up to order k. To see this can be done, recall that from Lemma 2.4.4 we may choose a basis of polynomials for $\mathbb{P}_k$ $\{P_{x,\beta}\}$, such that $\partial^\beta(P_{x,\beta})(x) = 1$ and $\partial^{\beta'}(P_{x,\beta})(x') = 0$ when $(x, \beta) \neq (x', \beta')$. Then the cancellation condition for $\widetilde{g'}$ is equivalent to

$$\left\langle \widetilde{g'}, P_{x,\beta} \right\rangle = 0 \quad (x, \beta) \in F,$$

which gives us the following equation for the coefficients $c_{x,\beta}$:

$$\int g'(y)P_{x,\beta}(y)dy + \sum_{(x',\beta')\in F} c_{x',\beta'} \left\langle \partial^{\beta'}\delta_{x'}, P_{x,\beta}\right\rangle = 0 \quad (x,\beta)\in F. \tag{2.11}$$

Now we note that

$$\left\langle \partial^{\beta'}\delta_{x'}, P_{x,\beta}\right\rangle = (-1)^{|\beta'|}\partial^{\beta'}(P_{x,\beta})(x'), \tag{2.12}$$

so from (2.11), (2.12) and Lemma 2.4.4 we have

$$c_{x,\beta} = (-1)^{|\beta|+1}\int g'(y)P_{x,\beta}(y)dy. \tag{2.13}$$

This shows that $\widetilde{g'}$ depends linearly on g, $\left\langle \tilde{g}', P\right\rangle = 0$ for every $P\in\mathbb{P}_k$, and moreover

$$|c_{x,\beta}| \leq |B|\|g'\|_{L^\infty(B)}\|P_{x,\beta}\|_{L^\infty(B)} \lesssim \|g'\|_{L^\infty(B)}w(E)^{-1} \lesssim a\delta^{-1}\|g'\|_{L^\infty(B)}.$$

Now $\widetilde{g'}/\|g'\|_{L^\infty(B)}$ satisfies all the required conditions of Lemma 2.4.5 and by applying that we get

$$\|\widetilde{g'}\|_{H^p(\mathbb{R}^n)} \lesssim \|g'\|_{L^\infty(B)} \lesssim_{a,\delta} |B_{ar}(x_0)|^{\frac{-1}{p}}.$$

Finally, let $\tilde{g} = \widetilde{g'}(\frac{x-x_0}{ar})$ which is an extension of g with $\|\tilde{g}\|_{H^p(\mathbb{R}^n)} \leq c(n,k,a,\delta)$, and this completes the proof of the second case.

Case $p \neq \frac{n}{n+k}$, $k = 0,1,2,\ldots$. This case is similar the previous one and even simpler. This time we can take a to be any number such that $E = \Omega^c \cap B_{ar} \neq \emptyset$ and after a dilation and translation, we pick a point $x\in E$ and extend g' by

$$\widetilde{g'} = g' + \sum_{|\beta|\leq[n(1/p-1)]} c_\beta \partial^\beta\delta_x, \quad c_\beta = (-1)^{|\beta|+1}\int g'(y)\frac{(y-x)^\beta}{\beta!}dy.$$

The rest is just like the previous case.

In all the above cases, the method of extension is linear in terms of g so from Lemma 2.2.1, there is a linear extension operator if the conditions of Theorem 2.3.1 hold. $\square$

2.5 Proof of Necessity

Now that we've proved the conditions of Theorem 2.3.1 are sufficient, we have to show they are necessary as well. To do this, we need to construct some functions in the dual of $H^p(\mathbb{R}^n)$ to prevent any bounded extension when those conditions are not satisfied. We need some lemmas and our first one is a simple consequence of a theorem proved by Uchiyama. We give it below and its proof is found in [27,69].

Theorem (Uchiyama)**.** *Let $\lambda > 0$ and E , B be two subsets of $\mathbb{R}^n$ such that for any cube Q we have*

$$\min\left\{\frac{|Q \cap E|}{|Q|}, \frac{|Q \cap B|}{|Q|}\right\} \leq 2^{-2n\lambda}. \tag{2.14}$$

Then there exists a function $f \in BMO(\mathbb{R}^n)$ such that $f = 1$ on B and $f = 0$ on E with $\|f\|_{BMO(\mathbb{R}^n)} \lesssim \frac{1}{\lambda}$.

Lemma 2.5.1. *Let $0 < \varepsilon \leq \frac{1}{2}$, $a > 3$ and B the unit ball of $\mathbb{R}^n$. Then for every set $E \subset (2B)^c$ such that $|E \cap aB| < \varepsilon$, there exists a function $f \in BMO(\mathbb{R}^n)$ with the following properties:*

- $f = 0$ *on* E
- $f = 1$ *on* B
- $\|f\|_{BMO(\mathbb{R}^n)} \lesssim \frac{1}{\min\{\log a, \log \varepsilon^{-1}\}}$

Proof. Let us check the condition of the above theorem for B and E. If a cube Q doesn't intersect one of the sets (2.14) holds for any $\lambda > 0$, so assume it intersects both of them. Then there are two possibilities either $Q \cap (aB)^c \neq \emptyset$ or $Q \subset aB$. In the first case $l(Q) \geq \frac{(a-2)}{\sqrt{n}}$ which implies $\frac{|Q \cap B|}{|Q|} \lesssim a^{-n}$, and in the second one we have $l(Q) \geq 1$, which gives us $\frac{|Q \cap A|}{|Q|} = \frac{|Q \cap E \cap aB|}{|Q|} \leq \varepsilon$. So for $\lambda \approx \min\{\log a, \log \varepsilon^{-1}\}$ (2.14) holds and the claim follows from the theorem. ☐

To construct similar functions of the above lemma in $\dot{\Lambda}^k(\mathbb{R}^n)$, we need a couple of facts which are well known and we bring here as lemmas. Here it's convenient to write $\mathrm{BMO}(\mathbb{R}) = \dot{\Lambda}^0(\mathbb{R})$.

Lemma 2.5.2. *Let $f \in \dot{\Lambda}^k(\mathbb{R})$ and $F(x) = \int_0^x f(t)dt$, Then $F \in \dot{\Lambda}^{k+1}(\mathbb{R})$ with $\|F\|_{\dot{\Lambda}^{k+1}(\mathbb{R})} \lesssim \|f\|_{\dot{\Lambda}^k(\mathbb{R})}$ for $k = 0, 1, 2, \ldots$.*

Proof. For $k = 0$ we have

$$|\Delta_h^2 F(x)/h| = \left|1/h \int_{x+h}^{x+2h} f - 1/h \int_x^{x+h} f\right| \lesssim \|f\|_{\mathrm{BMO}(\mathbb{R})},$$

and for $k \geq 1$ we have

$$\Delta_h^{k+2} F(x) = \Delta_h^{k+1}(\Delta_h^1 F)(x) = \Delta_h^{k+1} \int_0^h f(x+t)dt = \int_0^h (\Delta_h^{k+1} f)(x+t)dt,$$

which implies

$$|\frac{\Delta_h^{k+2} F(x)}{h^{k+1}}| \leq \frac{1}{|h|} \int_0^h |\frac{\Delta_h^{k+1}}{h^k}|(x+t)dt \leq \|f\|_{\dot{\Lambda}^k(\mathbb{R})}.$$

So in any case we get $\|F\|_{\dot{\Lambda}^{k+1}(\mathbb{R})} \lesssim \|f\|_{\dot{\Lambda}^k(\mathbb{R})}$. ☐

Lemma 2.5.3. *Let k be a positive integer, $f \in \dot{\Lambda}^k(\mathbb{R}^n)$ and $g \in C^k(\mathbb{R}^n)$. Then $fg \in \dot{\Lambda}^k(\mathbb{R}^n)$ and $\|fg\|_{\dot{\Lambda}^k(\mathbb{R}^n)} \lesssim \sum_{l=1}^k \|f\|_{\dot{\Lambda}^l(\mathbb{R}^n)} \|\nabla^{k-l}(g)\|_{L^\infty} + \|f\|_{L^\infty} \|\nabla^k(g)\|_{L^\infty}$.*

Proof. Let $\tau^h g(x) = g(x+h)$, so for $\Delta_h^1(fg)$ we have

$$\Delta_h^1(fg) = \Delta_h^1(f)\tau^h g + f\Delta_h^1(g). \tag{2.15}$$

By taking the difference operator $k+1$ times and applying (2.15) each time we can see that $\Delta_h^{k+1}(fg)$ has the form

$$\Delta_h^{k+1}(fg) = \sum_{-1\le l\le k}\ \sum_{0\le m,s\le k+1} c_{l,k,m,s}\tau^{mh}\Delta_h^{l+1}(f)\tau^{sh}\Delta_h^{k-l}(g), \tag{2.16}$$

where in this formula $\Delta_h^0(f) = f$ and $c_{l,k,m,s}$ are some integers. Now for $1 \le l \le k$ we have

$$|\Delta_h^{l+1}(f)| \le \|f\|_{\dot\Lambda^l(\mathbb{R})}|h|^l, \quad |\Delta_h^{k-l}(g)| \lesssim \|\nabla^{k-l}(g)\|_{L^\infty}|h|^{k-l},$$

where $\|\nabla^0(g)\|_{L^\infty} = \|g\|_{L^\infty}$, and for $l = -1, 0$ we have

$$|\Delta_h^{l+1}(f)| \lesssim \|f\|_{L^\infty}, \quad |\Delta_h^{k-l}(g)| \lesssim \|\nabla^k(g)\|_{L^\infty}|h|^k.$$

Putting these estimates in (2.16) completes the proof.

□

One more property of $\dot\Lambda^k(\mathbb{R}^n)$ that is needed here is the inequality

$$\|f\|_{\dot\Lambda^l(\mathbb{R}^n)} \lesssim \|f\|_{L^\infty}^{1-\frac{l}{k}}\|f\|_{\dot\Lambda^k(\mathbb{R}^n)}^{\frac{l}{k}} \quad 1 \le l \le k, \tag{2.17}$$

which is a consequence of the fact that $\dot\Lambda^l(\mathbb{R}^n)$ is the interpolation space between $L^\infty(\mathbb{R}^n)$ and $\dot\Lambda^k(\mathbb{R}^n)$ (see [47] Theorem 17.30).

Lemma 2.5.4. *Suppose $0 < \varepsilon < 1/4$, $a > 4$ and let $I \subset [-a, a]$ be an interval with $|I| < 2\varepsilon$. Then for each $k = 1, 2, 3, ...$ there exists a function $f_k \in \dot\Lambda^k(\mathbb{R}^1)$ with the following properties:*

- $f_k = 0$ *on* $I \cup [-a,a]^c$
- $|f_k| \gtrsim \min\{\log a, \log \varepsilon^{-1}\}$ *on* $J = [x_0 - 1/2, x_0 + 1/2]^c \cap [-1,1]$
- $\|f_k\|_{L^\infty} \lesssim a^k$
- $\|f_k\|_{\dot{\Lambda}^k(\mathbb{R}^1)} \lesssim 1$

Proof. Suppose x_0 is the mid point of I and consider the function

$$g(t) = \min\left\{\log^+ |\frac{t - x_0}{\varepsilon}|, \log^+ |\frac{a}{t}|\right\},$$

which is nonnegative, vanishes on $I \cup [-a,a]^c$, $g \geq \min\{\log a, \log \frac{\varepsilon^{-1}}{4}\}$ on the interval $[x_0 - 1/4, x_0 + 1/4]^c \cap [-1,1]$ and also $\|g\|_{\text{BMO}} \lesssim 1$. Now let $F_1(x) = \int_{x_0}^x g(t)dt$ which vanishes on I and $|F_1| \gtrsim \min\{\log a, \log \varepsilon^{-1}\}$ on $J = [x_0 - 1/2, x_0 + 1/2]^c \cap [-1,1]$. Also, according to Lemma 2.5.2 we have $\|F_1\|_{\dot{\Lambda}^1(\mathbb{R}^1)} \lesssim 1$ and we note that

$$|F_1(x)| \leq \int_{-a}^a g(t)dt \leq \int_{-a}^a \log^+ |\frac{a}{t}|dt \lesssim a.$$

Now let $\varphi \in C_0^\infty([-1,1])$ such that $0 \leq \varphi \leq 1$, $\varphi = 1$ on $[-1/2, 1/2]$ and set $\varphi_a(x) = \varphi(x/a)$. Then the function $f_1(x) = F_1(x)\varphi_a(x)$ has the following properties:

- $f_1 = 0$ on $I \cup [-a,a]^c$
- f_1 is positive on the right hand side of x_0 and negative on the other side with $|f_1| \gtrsim \min\{\log a, \log \varepsilon^{-1}\}$ on J
- $\|f_1\|_{L^\infty} \lesssim a$
- $\|f_1\|_{\dot{\Lambda}^1(\mathbb{R}^1)} \lesssim 1$

For the last property it's enough to note that

$$\|F_1\|_{\dot{\Lambda}^1(\mathbb{R}^1)} \lesssim 1, \quad \|F_1\|_{L^\infty} \lesssim a, \quad \|\varphi_a\|_{L^\infty} \leq 1, \quad \|\varphi_a'\|_{L^\infty} \lesssim a^{-1},$$

and therefore Lemma 2.5.3 implies $\|f_1\|_{\dot\Lambda^1(\mathbb{R}^1)} \lesssim 1$. We continue this process, meaning each time we integrate the function of the previous step from x_0 and with φ_a cut it off the interval $[-a, a]$ then at k-th step we have $F_k(x) = \int_{x_0}^x f_{k-1}(t)dt$ and $f_k = F_k\varphi_a$.

The function F_k vanishes on I, is positive on the right hand side of x_0 and negative or positive (depending on k) on the other side. Also, it follows from Lemma 2.5.2 that

$$\|F_k\|_{\dot\Lambda^k(\mathbb{R}^1)} \lesssim 1, \tag{2.18}$$

and moreover

$$\|F_k\|_{L^\infty} \leq \int_{-a}^{a} |f_{k-1}| \lesssim a^k. \tag{2.19}$$

Now from (2.17-2.19) we get $\|F_k\|_{\dot\Lambda^l(\mathbb{R}^1)} \lesssim a^{k-l}$ for $1 \leq l \leq k$. We also have the bound $\|\varphi_a^{(k-l)}\|_{L^\infty} \lesssim a^{l-k}$ for $0 \leq l \leq k$. Now these facts combined with Lemma 2.5.3 shows that $\|f_k\|_{\dot\Lambda^k(\mathbb{R}^1)} \lesssim 1$. Finally we observe that at the previous step, f_{k-1} changes sign only at x_0, which implies that by integrating it from x_0 no cancellation happens and then $|f_k| \gtrsim \min\{\log a, \log \varepsilon^{-1}\}$ on J. The proof is now complete. □

The following is an analog of Lemma 2.5.1 in the $\dot\Lambda^k(\mathbb{R}^n)$ setting.

Lemma 2.5.5. *Let $0 < \varepsilon < 1/4$, $a > 4$ and B the unit ball of $\mathbb{R}^n$. Then for every set $E \subset \mathbb{R}^n$ with $w(E \cap aB) < \varepsilon$, there exists a function $f \in \dot\Lambda^k(\mathbb{R}^n)$ such that*

- $f = 0$ *on* E
- $\int_B |f| \approx 1$
- $\|f\|_{\dot\Lambda^k(\mathbb{R}^n)} \lesssim \frac{1}{\min\{\log a, \log \varepsilon^{-1}\}}$

Proof. Let ν be a direction for which $w(E \cap aB, \nu) \leq \varepsilon$ and choose c such that $|\nu \cdot x + c| \leq \varepsilon$ for $x \in E \cap aB$. Also let $\varphi \in C_0^\infty(B)$ with $0 \leq \varphi \leq 1$ and $\varphi = 1$ on $\frac{1}{2}B$. Now take the function $f_k \in \dot\Lambda^k(\mathbb{R}^1)$ constructed in the previous lemma with

$I = [-\varepsilon, \varepsilon]$ and set

$$f(x) = \frac{1}{\min\{\log a, \log \varepsilon^{-1}\}} f_k(\nu \cdot x + c)\varphi(x/a).$$

This function has the desired properties. In fact the first two properties of f follows directly from properties of f_k, and for the third one we note that $\|f_k(\nu \cdot + c)\|_{\dot{\Lambda}^k(\mathbb{R}^n)} = \|f_k\|_{\dot{\Lambda}^k(\mathbb{R}^1)} \lesssim 1$, so an application of Lemma 2.5.3 gives us $\|f\|_{\dot{\Lambda}^k(\mathbb{R}^n)} \lesssim \frac{1}{\min\{\log a, \log \varepsilon^{-1}\}}$. □

Now that we have constructed our functions in the dual of Hardy spaces we can prove the necessity part of Theorem 2.3.1.

Proof of necessity. We have two cases:

Case $p = 1$. Suppose Ω doesn't satisfy (2.2) then for $\delta_j = 2^{-j}$ and $a_j = j$ there's a point x_j such that (2.2) doesn't hold. Let us denote $B_j = B(x_j, \frac{1}{2}d(x_j))$ and $r_j = \frac{1}{2}d(x_j)$.

Now suppose Ω still is an extension domain and take a sequence of $(1, \Omega)$-atoms g_j with supporting ball B_j and denote their extensions by $\tilde{g_j}$. Then we must have $\|\tilde{g_j}\|_{H^1(\mathbb{R}^n)} \lesssim 1$. Now by a translation and dilation we move every thing to the unit ball and set

$$E_j = \frac{\Omega^c - x_j}{r_j}, \quad g_j'(x) = r_j^n g_j(r_j x + x_j), \quad \tilde{g_j}'(x) = r_j^n \tilde{g_j}(r_j x + x_j).$$

Here we note that E_j satisfies the condition in Lemma 2.5.1 with $\varepsilon = \varepsilon_j \approx \delta_j a_j^n$ and $a = a_j$. So there exists a function $f_j \in BMO(\mathbb{R}^n)$ with properties of Lemma 2.5.1. Now from duality and the fact that $\|\tilde{g_j}'\|_{H^1(\mathbb{R}^n)} \lesssim 1$ we get

$$| \langle \tilde{g_j}', f_j \rangle | \lesssim \|\tilde{g_j}'\|_{H^1(\mathbb{R}^n)} \|f_j\|_{BMO(\mathbb{R}^n)} \lesssim \frac{1}{\min \left\{ \log a_j, \log \varepsilon_j^{-1} \right\}}$$

and on the other hand since $f_j = 0$ on E_j we have

$$\langle \tilde{g_j}', f_j \rangle = \int_{B_1} g_j' f_j, \tag{2.20}$$

which implies the following contradiction if we choose $g_j = \operatorname{sgn}(f_j).|B_j|^{-1}$

$$1 \approx \int_{B_1} |f_j| \approx | \langle \tilde{g_j}', f_j \rangle | \lesssim \frac{1}{\min \left\{ \log a_j, \log \varepsilon_j^{-1} \right\}} \longrightarrow 0.$$

The proof of this case is now complete.

Case $p \neq 1$. The above argument holds word by word if instead of Lemma 2.5.1 we use Lemma 2.5.5 and note that the functions f_j vanishes in a neighborhood of E_j so (2.20) holds.

This finishes the proof of necessity. □

2.6 Applications

As an application of Theorem 2.3.1 we are able to characterize the dual of $H^1(\Omega)$ as a complemented subspace of $BMO(\mathbb{R}^n)$ (see [54] the final remark). Let us recall that the dual space of $H^1(\Omega)$ which is denoted by $BMO(\Omega)$ in [54], consists of those functions f such that

$$\|f\|_{BMO(\Omega)} = \sup_{2Q \subset \Omega} \fint_Q |f - \fint_Q f| + \sup_{Q \in \mathcal{W}(\Omega)} \fint_Q |f| < \infty,$$

where $\mathcal{W}(\Omega)$ is the set of all Whitney cubes of Ω. By extending $f \in BMO(\Omega)$ to be zero outside of Ω, we get $\tilde{f} \in BMO(\mathbb{R}^n)$ with $\|\tilde{f}\|_{BMO(\mathbb{R}^n)} \lesssim \|f\|_{BMO(\Omega)}$. The converse of this inequality is also true when Ω is an H^1-extension domain. Indeed by using duality we have

$$\|f\|_{BMO(\Omega)} \approx \sup_{\|g\|_{H^1(\Omega)}=1} \langle f, g\rangle \lesssim_{a,\delta} \sup_{\|\tilde{g}\|_{H^1(\mathbb{R}^n)}=1} \left\langle \tilde{f}, \tilde{g} \right\rangle \approx \|\tilde{f}\|_{BMO(\mathbb{R}^n)}.$$

The above is true because every $g \in H^1(\Omega)$ has a bounded extension $\tilde{g} \in H^1(\mathbb{R}^n)$. So we have the following:

Corollary 2.6.1. *When Ω is an extension domain for H^1, the dual of $H^1(\Omega)$ is isomorphic to the closed subspace of $BMO(\mathbb{R}^n)$, consisting of all functions with support on Ω.*

Next, let $E : H^1(\Omega) \longrightarrow H^1(\mathbb{R}^n)$ be the extension operator, then by duality the adjoint operator $E' : BMO(\mathbb{R}^n) \longrightarrow BMO(\Omega)$ given by

$$\langle E(f), g\rangle = \langle f, E'(g)\rangle,$$

is bounded. Now if $g \in BMO(\mathbb{R}^n)$ with support on Ω and f is any $(1,\Omega)$-atom supported on a Whitney cube Q, we have

$$\int fg = \int E(f)g = \int fE'(g).$$

So on each Whitney cube Q, $E'(g) = g$, and therefore $E'(g) = g$ on the whole Ω. This allows us to conclude the following:

Corollary 2.6.2. *Let Ω be an open subset of $\mathbb{R}^n$ satisfying (2.2). Then the subspace of all functions in $BMO(\mathbb{R}^n)$, vanishing outside of Ω is a complemented subspace of $BMO(\mathbb{R}^n)$. Moreover, the operator E' defined above is a bounded projection onto this subspace.*

Chapter 3

Maximal Operators on BMO and Slices

3.1 Introduction

Let $A \subset \mathbb{R}^n$ be a measurable set with positive finite measure and $f \in L^1_{loc}(\mathbb{R}^n)$. By the mean oscillation of f on A we mean the quantity

$$\operatorname{osc}(f, A) := \fint_A |f - f_A|,$$

where $\fint_A f$ and f_A mean the average of f over A, i.e. $\frac{1}{|A|}\int_A f$. Then it is said that f is of bounded mean oscillation if $\operatorname{osc}(f, Q)$ is uniformly bounded on all cubes Q (by a cube we mean a closed cube with sides parallel to the axes). The space of such functions is denoted by $BMO(\mathbb{R}^n)$, and modulo constants the following quantity defines a norm on this space:

$$\|f\|_{BMO(\mathbb{R}^n)} := \sup_Q O(f, Q).$$

Sometimes we use $\|f\|_{BMO(Q_0)}$ which means that we take the above supremum over all cubes contained in Q_0. $BMO(\mathbb{R}^n)$ is a Banach space, and since its introduction has played an important role in harmonic analysis. It is the dual of the Hardy space $H^1(\mathbb{R}^n)$, it contains $L^\infty(\mathbb{R}^n)$ and somehow serves as a substitute for it. For instance, Calderón-Zygmund singular integral operators map $BMO(\mathbb{R}^n)$ to itself and consequently these operators map $L^\infty(\mathbb{R}^n)$ to $BMO(\mathbb{R}^n)$ but not to itself [25].

Another important class of operators is the class of maximal operators, and the first objective of the present paper is to investigate the action of some of these operators on $BMO(\mathbb{R}^n)$. Let us recall that the uncentered Hardy-Littlewood maximal operator is defined by

$$Mf(x) := \sup_{x\in Q} \fint_Q |f|, \quad f \in L^1_{loc}(\mathbb{R}^n),$$

where the above supremum is taken over all cubes containing x. As it is well known M is of weak-type $(1,1)$ and bounded on $L^p(\mathbb{R}^n)$ for $1 < p \leq \infty$ [25]. For a function $f \in BMO(\mathbb{R}^n)$, it might be the case that Mf is identically equal to infinity. For instance, this is the case when $f(x) = \log|x|$. However, in [3] the authors proved that if this is not the case then Mf belongs to $BMO(\mathbb{R}^n)$ and for a dimensional constant $c(n)$ we have

$$\|Mf\|_{BMO(\mathbb{R}^n)} \leq c(n)\|f\|_{BMO(\mathbb{R}^n)}.$$

Another proof of this was given in [2], and a third one in [60], where the author proved that M preserves Poincaré inequalities. Regarding this, we ask the following question about the continuity of the uncentered Hardy-Littlewood maximal operator on $BMO(\mathbb{R}^n)$:

Question 3.1.1. *Let $f \in L^\infty(\mathbb{R}^n)$ and $\{f_k\}$ be a sequence of bounded functions converging to f in $BMO(\mathbb{R}^n)$. Is it true that $\{Mf_k\}$ converges to Mf in $BMO(\mathbb{R}^n)$?*

The operator M is nonlinear and for such operators continuity does not follow from boundedness. However, it is pointwise sublinear and this makes it continuous on $L^p(\mathbb{R}^n)$ for $1 < p \leq \infty$. In [52], a similar question has been studied for Sobolev spaces, where the author proved that M is continuous on $W^{1,p}(\mathbb{R}^n)$ for $1 < p < \infty$. However, in Section 2 we give a negative answer to the above question.

$BMO(\mathbb{R}^n)$ has an important subspace, namely $VMO(\mathbb{R}^n)$ or functions of vanishing mean oscillation. $VMO(\mathbb{R}^n)$ is the closure of the uniformly continuous functions in $BMO(\mathbb{R}^n)$. Another characterization of $VMO(\mathbb{R}^n)$ is given in terms of the modulus of mean oscillation which is defined by

$$\omega(f,\delta) := \sup_{l(Q)\leq\delta} O(f,Q), \tag{3.1}$$

and $f \in VMO(\mathbb{R}^n)$ exactly when $\lim\limits_{\delta\to 0} \omega(f,\delta) = 0$ (in the above by $l(Q)$ we mean the side length of Q) [61]. Regarding this subspace we ask:

Question 3.1.2. *Let $f \in VMO(\mathbb{R}^n)$ such that Mf is not identically equal to infinity. Is it true that $Mf \in VMO(\mathbb{R}^n)$?*

In Section 3 we provide a positive answer to this question.

In the last section, we consider the action of some other maximal operators on $BMO(\mathbb{R}^n)$. More specifically, the directional maximal operator in the direction $e_1 = (1, 0, ..., 0)$, M_{e_1}, and the strong maximal operator, M_s, which are defined as the following:

$$M_{e_1} f(x_1, x') := \sup_{x_1\in I} \fint_I |f_{x'}|, \qquad M_s f(x) := \sup_{x\in R} \fint_R |f|.$$

In the above, $f_{x'}(t) := f(t, x')$, where $(t, x') \in \mathbb{R} \times \mathbb{R}^{n-1}$, and the left supremum is taken over all closed intervals containing x_1. In a similar way one can define the directional maximal operator M_e, which is taken in the direction $e \in \mathbb{S}^{n-1}$, simply by taking the one dimensional uncentered Hardy-Littlewood maximal operator on every line in direction e. However, since $BMO(\mathbb{R}^n)$ is invariant under rotations it is enough to study M_{e_1}. In the above the right supremum is taken over all rectangles containing x, and by a rectangle we mean a closed rectangle with sides parallel to the axes. These are the most important maximal operators in multi-parameter harmonic analysis and are bounded and continuous on $L^p(\mathbb{R}^n)$ for $1 < p \le \infty$ [16]. Regarding these operators we ask:

Question 3.1.3. *Are there constants $C, C' \ge 1$ such that at least for every bounded function f the following inequalities hold?*

$$\|M_{e_1} f\|_{BMO(\mathbb{R}^n)} \le C\|f\|_{BMO(\mathbb{R}^n)}, \quad \|M_s f\|_{BMO(\mathbb{R}^n)} \le C'\|f\|_{BMO(\mathbb{R}^n)}.$$

To answer this question we have to study the properties of slices of functions in $BMO(\mathbb{R}^n)$, which is the second objective of this paper. Many function spaces have the property that their slices lie in the same scale of spaces. For example, almost every slice of a function in $L^p(\mathbb{R}^n)$ or $W^{1,p}(\mathbb{R}^n)$ lies in $L^p(\mathbb{R}^{n-1})$ or $W^{1,p}(\mathbb{R}^{n-1})$, respectively [47]. The same is true for $BMO_s(\mathbb{R}^n)$, strong BMO, which is the subspace of $BMO(\mathbb{R}^n)$ consisting of all functions with bounded mean oscillation on rectangles [18]. This property is also satisfied by the scale of homogeneous Lipschitz spaces $\dot{\Lambda}^{n(\frac{1}{p}-1)}(\mathbb{R}^n)$, the duals of $H^p(\mathbb{R}^n)$ for $0 < p < 1$ [25]. Regarding this, we ask our last question:

Question 3.1.4. *Is it true that almost every horizontal or vertical slice of a function in $BMO(\mathbb{R}^2)$ belongs to $BMO(\mathbb{R})$?*

In Section 4 we answer both questions negatively and in the last theorem of this paper we prove a property of the slices of functions in $BMO(\mathbb{R}^2)$.

Before we proceed further, let us fix some notation. By $A \lesssim B$, $A \gtrsim B$ and $A \approx B$, we mean $A \leq CB$, $A \geq CB$ and $C^{-1}B \leq A \leq CB$ respectively, where C is a constant independent of the important parameters.

3.2 Discontinuity of M on $BMO(\mathbb{R}^n)$

Our theorem in this section is the following:

Theorem 3.2.1. *Let f be a nonnegative function supported in $[0,1]$, $\|f\|_{L^\infty} \leq 1$ and $\|f\|_{L^1} > \log 2$. Then, there exists a sequence of bounded functions $\{f_n\}$ converging to f in $BMO(\mathbb{R})$ such that $\{Mf_n\}$ does not converge to Mf in $BMO(\mathbb{R})$.*

To prove this we need a couple of simple lemmas which we give below.

Lemma 3.2.1. *Let $T > 0$ and $h \in BMO[0, \frac{T}{2}]$. Then the even periodic extension of h, which is defined by*

$$H(x) := h(x), \quad x \in [0, \frac{T}{2}], \qquad H(-x) = H(x), \quad H(x+T) = H(x), \quad x \in \mathbb{R},$$

is in $BMO(\mathbb{R})$ and $\|H\|_{BMO(\mathbb{R})} \leq 10\|h\|_{BMO[0,\frac{T}{2}]}$.

Proof. For an arbitrary interval I, there are two possibilities:

(i) $|I| \leq \frac{T}{2}$.

In this case by a translation by an integer multiple of T and using periodicity of H, we may assume either $I \subset [-\frac{T}{2}, \frac{T}{2}]$ or $I \subset [0, T]$. Suppose $I \subset [-\frac{T}{2}, \frac{T}{2}]$ and note that if $0 \notin I$, we have either $I \subset [0, \frac{T}{2}]$ or $I \subset [-\frac{T}{2}, 0]$ and from the symmetry $O(H, I) \leq \|h\|_{BMO[0,\frac{T}{2}]}$. If $0 \in I$, then take the interval J centered at zero with the right half J^+, which contains I and $|J| \leq 2|I|$. Again from symmetry we get

$$O(H,I) \leq 2\frac{|J|}{|I|}O(H,J) \leq 4O(H,J^+) \leq 4\|h\|_{BMO[0,\frac{T}{2}]}.$$

The same argument works for $I \subset [0,T]$. This time we use the symmetry of H around $\frac{T}{2}$.

(ii) $|I| \geq \frac{T}{2}$.

This time take $J = [nT, mT]$ with $n, m \in \mathbb{Z}$ which contains I and $|J| \leq |I| + 2T \leq 5|I|$. And again like the previous cases, from the symmetry and periodicity of H, we get

$$O(H,I) \leq 2\frac{|J|}{|I|}O(H,J) \leq 10O(h,[0,\frac{T}{2}]) \leq 10\|h\|_{BMO[0,\frac{T}{2}]}.$$

The proof is now complete. □

In the above, the norm of the extension operator is independent of T, and we will use this in the proof of the next lemma.

Remark 3.2.1. *There are much more general ways to extend BMO functions to the outside of domains, but for the purpose of our paper the above simple lemma is enough. See [42] for more on extensions.*

Lemma 3.2.2. *For $c < -1$, there exists a sequence of functions $\{g_n\}$, $n \geq 1$ with the following properties:*

a. $g_n \geq 0$

b. $g_n = 0$ *on* $[c, 1]$

c. $\|g_n\|_{L^\infty} \leq 1$

d. $\lim_{n\to\infty} ⨍_{[0,n]} g_n = 1$

e. $\lim_{n\to\infty} \|g_n\|_{BMO(\mathbb{R})} = 0$.

Proof. Let $\log^+ |x| = \max\{0, \log |x|\}$ be the positive part of the logarithm, and consider the function $h_n(x) = \log^+ x$ on the interval $[0, n]$, which belongs to $BMO[0, n]$ with $\|h_n\|_{BMO[0,n]} \leq \| \log^+ | \cdot | \|_{BMO(\mathbb{R})}$. Then an application of Lemma 3.2.1 with $T = 2n$ gives us a sequence of nonnegative functions H_n with $\|H_n\|_{BMO(\mathbb{R})} \lesssim 1$ (here our bounds are independent of n). Now, let $g_n = \frac{1}{1+\log n} H_n(x) \left(1 - \chi_{[c,0]}(x)\right)$. Then, the first three properties are immediate from the definition, the forth one follows from integration, and the last one from

$$\|g_n\|_{BMO(\mathbb{R})} \leq \frac{1}{1 + \log n} \left(\|H_n\|_{BMO(\mathbb{R})} + \|H_n \chi_{[c,0]}\|_{L^\infty}\right) \lesssim \frac{\log |c|}{1 + \log n} \quad n \geq 1.$$

This finishes the proof. □

Now, we turn to the proof of the above theorem.

Proof of Theorem 3.2.1. Let f be as in the theorem, $a = \int_0^1 f$, $c < 0$ a constant with large magnitude to be determined later, and let g_n be the sequence constructed in Lemma 3.2.2.

We will show that

$$\varliminf_{n\to\infty} \|Mf_n - Mf\|_{BMO(\mathbb{R})} > 0, \quad f_n := f + \frac{a}{1-c} g_n, \quad n \geq 1. \tag{3.2}$$

This proves the theorem once we note that since f and g_n are bounded functions, f_n is bounded too. Also, from the fifth property of $\{g_n\}$ in the above lemma, $\{f_n\}$ converges to f in $BMO(\mathbb{R})$.

To begin with, we claim that $Mf_n = Mf$ on $[c, 0]$. To see this, note that from the positivity of f and g_n, $Mf_n(x) \geq Mf(x)$ for all values of x, and it remains to show that the reverse inequality holds also. For $x \in [c, 0]$, $Mf(x) \geq \frac{\int_c^1 f}{1-c} = \frac{a}{1-c}$, and for any interval I which contains x, we have two possibilities:

(i) either $I \subset (-\infty, 0)$, in which case from the third property of g_n we have

$$\fint_I f_n = \fint_I \left(f + \frac{a}{1-c} g_n \right) = \frac{a}{1-c} \fint_I g_n \leq \frac{a}{1-c} \|g_n\|_{L^\infty} \leq \frac{a}{1-c} \leq Mf(x),$$

(ii) or $I \cap [0, 1] \neq \emptyset$, in which case the second and the third property of g_n gives us

$$\begin{aligned} \fint_I f_n &= \fint_I \left(f + \frac{a}{1-c} g_n \right) = \frac{|I \cap [x, 1]|}{|I|} \fint_{I \cap [x,1]} f + \frac{|I \backslash [x, 1]|}{|I|} \frac{a}{1-c} \fint_{I \backslash [x,1]} g_n \\ &\leq \frac{|I \cap [x, 1]|}{|I|} Mf(x) + \frac{|I \backslash [x, 1]|}{|I|} \frac{a}{1-c} \leq Mf(x). \end{aligned}$$

This proves our claim.

Next, we look at the mean oscillation of $Mf_n - Mf$ on $[2c, 0]$. Because this function vanishes on $[c, 0]$, we have

$$O(Mf_n - Mf, [2c, 0]) \geq \frac{1}{4} \fint_{[2c,c]} (Mf_n - Mf). \tag{3.3}$$

To bound the right hand side of the above inequality from below, we note that $0 \leq f \leq \chi_{[0,1]}$ so $Mf(x) \leq M(\chi_{[0,1]})(x) = \frac{1}{1-x}$ for $x \leq 0$. Also, for $x \leq 0$ we have

$$Mf_n(x) = M \left(f + \frac{a}{1-c} g_n \right) (x) \geq \frac{a}{1-c} \fint_{[x,n]} g_n \geq \frac{a}{1-c} . \frac{n}{n-x} \fint_{[0,n]} g_n.$$

So we get the following estimate for the right hand side in (3.3):

$$\fint_{[2c,c]} (Mf_n - Mf) \geq \frac{a}{1-c} \fint_{[0,n]} g_n \fint_{[2c,c]} \frac{n}{n-x} dx - \fint_{[2c,c]} \frac{1}{1-x} dx. \tag{3.4}$$

Combining (3.3) and (3.4), gives us

$$\|Mf_n - Mf\|_{BMO(\mathbb{R})} \geq \frac{1}{4}\left(\frac{a}{1-c}\fint_{[0,n]} g_n \fint_{[2c,c]} \frac{n}{n-x}dx + \frac{1}{c}\log\left(1+\frac{c}{c-1}\right)\right).$$

Now, taking the limit inferior as $n \to \infty$ and using the forth property of g_n gives us

$$\varliminf_{n\to\infty} \|Mf_n - Mf\|_{BMO(\mathbb{R})} \geq \frac{1}{4}\left(\frac{a}{1-c} + \frac{1}{c}\log\left(1+\frac{c}{c-1}\right)\right).$$

This shows that if we have

$$a > \frac{c-1}{c}\log\left(1+\frac{c}{c-1}\right) \tag{3.5}$$

then (3.2) holds. Here we note that the function on the right hand side of (3.5), attains its minimum, which is $\log 2$, at infinity. Also, from the assumption $a > \log 2$, so if we choose $|c|$ sufficiently large (3.5) holds, and this completes the proof. □

By lifting the above functions to higher dimensions with

$$f(x_1, .., x_n) = f(x_1), \quad g_m(x_1, .., x_n) = g_m(x_1), \tag{3.6}$$

we obtain a counterexample for continuity of the n-dimensional uncentered Hardy-Littlewood maximal operator on $BMO(\mathbb{R}^n)$, simply because the $BMO(\mathbb{R}^n)$ norms and the maximal operator become one dimensional.

Corollary 3.2.1. *The uncentered Hardy-Littlewood maximal operator is bounded on* $L^\infty(\mathbb{R}^n)$ *equipped with the BMO norm, but it is not continuous.*

3.3 The Uncentered Hardy-Littlewood Maximal Operator on $VMO(\mathbb{R}^n)$

As it was mentioned before, $VMO(\mathbb{R}^n)$ is the $BMO(\mathbb{R}^n)$-closure of the uniformly continuous functions which belong to $BMO(\mathbb{R}^n)$. The operator M reduces modulus of continuity, because it is pointwise sublinear, so it preserves uniformly continuous functions. But from our previous result, one cannot deduce boundedness of M on $VMO(\mathbb{R}^n)$ by a limiting argument. Nevertheless, we have the following theorem:

Theorem 3.3.1. *Let $f \in VMO(\mathbb{R}^n)$ and suppose Mf is not identically equal to infinity. Then Mf belongs to $VMO(\mathbb{R}^n)$.*

Before we prove this, we bring the following lemma which is needed later.

Lemma 3.3.1. *Let A be a measurable subset of a cube Q of positive measure and $f \in BMO(\mathbb{R}^n)$ with $\|f\|_{BMO(\mathbb{R}^n)} = 1$; then we have*

$$⨍_A |f - f_Q| \lesssim 1 + \log \frac{|Q|}{|A|}. \tag{3.7}$$

Proof. From the John-Nirenberg inequality [41], there is a dimensional constant $c > 0$ such that

$$⨍_A e^{c|f-f_Q|} \leq \frac{|Q|}{|A|} ⨍_Q e^{c|f-f_Q|} \lesssim \frac{|Q|}{|A|}.$$

Now, Jensen's inequality gives us (3.7), as follows:

$$⨍_A |f - f_Q| = \frac{1}{c} ⨍_A \log e^{c|f-f_Q|} \leq \frac{1}{c} \log ⨍_A e^{c|f-f_Q|} \lesssim 1 + \log \frac{|Q|}{|A|}.$$

□

Remark 3.3.1. *In the above lemma, let A be a rectangle and take Q to be the smallest* *cube which contains it. Then*

$$O(f, A) \lesssim 1 + \log e(A),$$

where $e(A)$ is the eccentricity of A, or the ratio of the largest side to the smallest one.

We now turn to the proof of Theorem 3.3.1.

Proof of Theorm 3.3.1. Let f be as in the theorem then we have to show that $\overline{\lim}_{\delta \to 0} \omega(Mf, \delta) = 0$. Now, for every cube Q, we have $O(|f|, Q) \leq 2O(f, Q)$, which means that $|f| \in VMO(\mathbb{R}^n)$ too. From this together with $M(|f|) = Mf$, it is enough to prove the theorem for nonnegative functions. Also, from the homogeneity of M we may assume $\|f\|_{BMO(\mathbb{R}^n)} = 1$.

Let Q_0 be a cube and c a constant with $c > e$. We decompose M into the local part, M_1, and the nonlocal part, M_2, as follows:

$$M_1 f(x) := \sup_{\substack{x \in Q \\ l(Q) \leq cl(Q_0)}} f_Q \ , \qquad M_2 f(x) := \sup_{\substack{x \in Q \\ l(Q) \geq cl(Q_0)}} f_Q \ .$$

We have $Mf(x) = \max\{M_1 f(x), M_2 f(x)\}$ and so

$$O(Mf, Q_0) \lesssim O(M_1 f, Q_0) + O(M_2 f, Q_0). \tag{3.8}$$

To estimate the first term in the right hand side of (3.8), let Q_0^* be the concentric dilation of Q_0 with $l(Q_0^*) = 2cl(Q_0)$. Then for the local part we have

$$\begin{aligned} O(M_1 f, Q_0) &\leq 2 \fint_{Q_0} \left|M_1 f - f_{Q_0^*}\right| \leq 2 \fint_{Q_0} M_1 \left|f - f_{Q_0^*}\right| \\ &\leq 2 \left(\fint_{Q_0} \left(M_1 \left|f - f_{Q_0^*}\right|\right)^2 \right)^{\frac{1}{2}} \leq 2 \left(\frac{1}{|Q_0|} \int \left(M \left|f - f_{Q_0^*}\right| \chi_{Q_0^*}\right)^2 \right)^{\frac{1}{2}}. \end{aligned}$$

By using the boundedness of M on $L^2(\mathbb{R}^n)$ we get

$$O(M_1 f, Q_0) \lesssim c^{\frac{n}{2}} \left(\fint_{Q_0^*} \left| f - f_{Q_0^*} \right|^2 \right)^{\frac{1}{2}},$$

and an application of the John-Nirenberg inequality gives us

$$O(M_1 f, Q_0) \lesssim c^{\frac{n}{2}} \|f\|_{BMO(Q_0^*)}. \tag{3.9}$$

To estimate the mean oscillation of the nonlocal part, suppose $x, y \in Q_0$, $M_2 f(x) > M_2 f(y)$ and let Q be a cube with $l(Q) \geq cl(Q_0)$, which contains x and such that $M_2 f(y) < f_Q$. Now, let Q' be a cube such that $Q_0 \cup Q \subset Q'$, $l(Q') = l(Q) + l(Q_0)$, and let $A = Q' \backslash Q$. Then $M_2 f(y) \geq f_{Q'}$ and we have

$$\begin{aligned} f_Q - M_2 f(y) &\leq f_Q - f_{Q'} = f_Q - \left(\frac{|A|}{|Q'|} f_A + \frac{|Q|}{|Q'|} f_Q \right) = \frac{|A|}{|Q'|} \left(f_Q - f_A \right) \\ &\leq \frac{|A|}{|Q'|} \left(|f_Q - f_{Q'}| + |f_{Q'} - f_A| \right) \lesssim \frac{|A|}{|Q|} O(f, Q') + \frac{|A|}{|Q'|} \fint_A |f - f_{Q'}|. \end{aligned}$$

Here we note that $|A| = |Q'| - |Q| \approx l(Q_0) l(Q)^{n-1}$, and $l(Q') \approx l(Q)$. So from the above inequality and Lemma 3.3.1 we get

$$f_Q - M_2 f(y) \lesssim \frac{l(Q_0)}{l(Q)} \left(1 + \log \frac{l(Q)}{l(Q_0)} \right) \lesssim c^{-1} \log c.$$

The reason for the last inequality is that $\frac{l(Q_0)}{l(Q)} \leq c^{-1}$ and the function $-t \log t$ is increasing when $t < e^{-1}$. Finally, by taking the supremum over all such cubes Q, we obtain

$$|M_2 f(x) - M_2 f(y)| \lesssim c^{-1} \log c, \quad x, y \in Q_0.$$

So for the nonlocal part we have

$$O(M_2 f, Q_0) \leq \fint_{Q_0} \fint_{Q_0} |M_2 f(x) - M_2 f(y)| \, dx dy \lesssim c^{-1} \log c. \tag{3.10}$$

By putting (3.8), (3.9) and (3.10) together we get

$$O(Mf, Q_0) \lesssim c^{\frac{n}{2}} \|f\|_{BMO(Q_0^*)} + c^{-1} \log c,$$

and taking the supremum over all cubes Q_0 with $l(Q_0) \leq \delta$ gives us

$$\omega(Mf, \delta) \lesssim c^{\frac{n}{2}} \omega(f, 2c\delta) + c^{-1} \log c.$$

To finish the proof, it is enough to take the limit superior as $\delta \to 0$ first and then let $c \to \infty$. □

Remark 3.3.2. *The above argument shows that for all functions in $BMO(\mathbb{R}^n)$, if one chooses a sufficiently large localization of M, (3.10) holds, meaning that the mean oscillation of the nonlocal part is small. This also shows itself in the dyadic setting: if one considers the dyadic maximal operator M^d and dyadic BMO, denoted by $BMO_d(\mathbb{R}^n)$, then for a dyadic cube Q_0*

$$M_2^d f(x) = \sup_{\substack{x \in Q \\ l(Q) \geq l(Q_0)}} f_Q = \sup_{Q_0 \subset Q} f_Q, \quad x \in Q_0.$$

Hence, $O(M_2^d f, Q_0) = 0$ and therefore no dilation is needed ($c = 1$).

3.4 Slices of BMO functions and unboundedness of directional and strong maximal operators

In this final section we discuss properties of slices of functions in $BMO(\mathbb{R}^n)$, and for simplicity we restrict ourselves to $BMO(\mathbb{R}^2)$. We begin by asking:

Question. *Suppose φ, ψ are two functions of one variable, when does $f(x,y) = \varphi(x)\psi(y)$ belong to $BMO(\mathbb{R}^2)$?*

To answer this, we need the following lemma which is an application of Fubini's theorem and its proof is found in [18].

Lemma 3.4.1. *Let $A, B \subset \mathbb{R}$ be two measurable sets with finite positive measure, and f be a measurable function on $\mathbb{R}^2$. Then*

$$O(f, A \times B) \approx ⨍_B O(f_y, A)dy + ⨍_A O(f_x, B)dx.$$

Now, take two intervals I, J with $l(I) = l(J)$. Then, an application of the above lemma to $f(x,y) = \varphi(x)\psi(y)$ gives us

$$O(f, I \times J) \approx O(\psi, J) ⨍_I |\varphi| + O(\varphi, I) ⨍_J |\psi|.$$

Taking the supremum over all such I, J we obtain

$$\|f\|_{BMO(\mathbb{R}^2)} \approx \sup_{\delta>0} \left(\sup_{l(I)=\delta} ⨍_I |\varphi|. \sup_{l(J)=\delta} O(\psi, J) + \sup_{l(I)=\delta} O(\varphi, I). \sup_{l(J)=\delta} ⨍_J |\psi| \right). \quad (3.11)$$

When $f \in BMO(\mathbb{R}^2)$ is non-zero on a set of positive measure, the above condition implies that $\varphi, \psi \in BMO(\mathbb{R})$. To see this, note that if φ is non-zero on a set of positive measure, for some Lebesgue point of φ like x, $\varphi(x) \neq 0$. Then from the Lebesgue differentiation theorem, for sufficiently small δ we must have

$\sup_{l(I)=\delta} ⨍_I |\varphi| \gtrsim |\varphi(x)| > 0$. So ψ has bounded mean oscillation on intervals with length less than δ. For intervals J with $l(J) \geq \delta$, $|\psi|$ has bounded averages because otherwise there is a sequence of intervals J_n with $l(J_n) \geq \delta$ and $\lim ⨍_{J_n} |\psi| = \infty$. Then by dividing each of these intervals into sufficiently small pieces of length between $\frac{\delta}{2}$ and δ, we conclude that $|\psi|$ has large averages over such intervals so $\sup_{l(J)=\delta} ⨍_J |\psi| = \infty$. But then $\sup_{l(I)=\delta} O(\varphi, I) = 0$, which means that φ is constant. We summarize the above discussion in the following proposition.

Proposition 3.4.1. *Let $f(x,y) = \varphi(x)\psi(y)$, $f \in BMO(\mathbb{R}^2)$ if and only if (3.11) holds and if $f \neq 0$, then $\varphi, \psi \in BMO(\mathbb{R})$.*

Remark 3.4.1. *When φ and ψ are not constants, the above argument shows that they belong to $bmo(\mathbb{R})$, the nonhomogeneous BMO space, which is a proper subspace of $BMO(\mathbb{R})$. See [28] for more on $bmo(\mathbb{R})$.*

Corollary 3.4.1. *Let $\log^- |x| = \max\{0, -\log |x|\}$ be the negative part of the logarithm and $p, q > 0$ with $p + q \leq 1$. Then the function $f(x,y) = \left(\log^- |x|\right)^p \left(\log^- |y|\right)^q$ is in $BMO(\mathbb{R}^2)$.*

Proof. A direct calculation shows that

$$\sup_{l(I)=\delta} ⨍_I \left(\log^- |x|\right)^p dx \approx \begin{cases} \delta^{-1} & \delta \geq \frac{1}{2} \\ (-\log \delta)^p & \delta < \frac{1}{2} \end{cases},$$

$$\sup_{l(J)=\delta} O\left(\log^- |\cdot|, J\right)^q \approx \begin{cases} \delta^{-1} & \delta \geq \frac{1}{2} \\ (-\log \delta)^{q-1} & \delta < \frac{1}{2} \end{cases}.$$

and the claim follows from Proposition 3.4.1. □

Remark 3.4.2. *The above function f does not have bounded mean oscillation on rectangles, simply because the BMO-norm of the slices becomes larger and larger as we get closer to the origin. See [45] (Example 2.32) for another example.*

Now we answer the third question of this paper.

Theorem 3.4.1. *There exists a sequence of bounded functions* $\{G_N\}$, $N \geq 1$ *such that it is bounded in* $BMO(\mathbb{R}^2)$ *but*

$$\lim_{N\to\infty} \|M_{e_1}(G_N)\|_{BMO(\mathbb{R}^2)} = \infty, \quad \lim_{N\to\infty} \|M_s(G_N)\|_{BMO(\mathbb{R}^2)} = \infty.$$

To prove this we need the following simple lemma.

Lemma 3.4.2. *Let* $Q_0 = [-1,1]^n$, $f \in BMO(\mathbb{R}^n)$ *with support in* Q_0, *and* x_k *be a sequence in* $\mathbb{R}^n$ *with* $|x_k - x_m| \geq 3\sqrt{n}$ *for* $k \neq m$. *Then* $g(x) = \sum f(x - x_k)$ *is in* $BMO(\mathbb{R}^n)$ *and* $\|g\|_{BMO(\mathbb{R}^n)} \lesssim \|f\|_{BMO(\mathbb{R}^n)}$.

Proof. First, by comparing the average of $|f|$ on Q_0 with $Q_0 + 2e_1$ we have

$$\int_{Q_0} |f| = 2^n \left(\fint_{Q_0} |f| - \fint_{Q_0+2e_1} |f| \right) \lesssim O\left(|f|, [-1,3]^n\right) \leq \||f|\|_{BMO(\mathbb{R}^n)} \leq 2\|f\|_{BMO(\mathbb{R}^n)}.$$

Next, take a cube Q and suppose for some k, $Q \cap (x_k + Q_0) \neq \emptyset$. We note that the distance of the support of functions $f(\cdot - x_k)$ from each other is at least $\sqrt{n}$ so if $l(Q) \leq 1$ then $O(g,Q) = O(f(\cdot - x_k), Q) \leq \|f\|_{BMO(\mathbb{R}^n)}$. Otherwise, we have

$$\begin{aligned} O(g,Q) \leq & 2\fint_Q |g| \leq \frac{2}{|Q|} \sum_{Q\cap(x_k+Q_0)\neq\emptyset} \int_{x_k+Q_0} |f(y - x_k)|\, dy \\ & \lesssim \frac{\#\{k | Q \cap (x_k + Q_0) \neq \emptyset\}}{|Q|} \|f\|_{BMO(\mathbb{R}^n)}. \end{aligned}$$

Now to finish the proof, note that $\#\{k|Q \cap (x_k + Q_0) \neq \emptyset\} \lesssim |Q|$ which implies $O(g,Q) \lesssim \|f\|_{BMO(\mathbb{R}^n)}$. □

Proof of Theorem 3.4.1. We may assume $n = 2$, since by a lifting argument similar to (3.6), we can conclude the theorem for higher dimensions. Let f be as in Corollary 3.4.1, N a positive integer, and consider the following function:

$$g_N(x,y) = \sum_{k=0}^{N} \sum_{m=2^k}^{2^{k+1}-1} f\left(x - 3\sqrt{2}m, y - \frac{k}{N}\right).$$

g_N has the following properties:

(i) $\|g_N\|_{BMO(\mathbb{R}^2)} \lesssim 1$ (here our bounds only depend on p, q but not N).

This follows from Corollary 3.4.1 and Lemma 3.4.2 applied to f with $x_{m,k} = \left(3\sqrt{2}m, \frac{k}{N}\right)$.

(ii) $M_s(g_N)(x,y) \geq M_{e_1}(g_N)(x,y) \gtrsim (\log N)^q$ for $0 \leq x, y \leq 1$ and $N \geq 2$.

To see this, let $0 \leq x \leq 1$ and $\frac{l}{N} \leq y < \frac{l+1}{N}$ for some $l < N$. Then consider the average of $(g_N)_y$ on $I = [0, 3 \cdot 2^{l+1}\sqrt{2}]$, which is bounded from below by

$$⨍_I (g_N)_y \geq \frac{1}{3 \cdot 2^{l+1}\sqrt{2}} \sum_{m=2^l}^{2^{l+1}-1} \int_I f\left(t - 3\sqrt{2}m, y - \frac{l}{N}\right) dt.$$

Now note that for $2^l \leq m \leq 2^{l+1} - 1$, I contains the support of $f\left(\cdot - 3\sqrt{2}m, y - \frac{l}{N}\right)$ and since $0 \leq y - \frac{l}{N} \leq \frac{1}{N}$ we have

$$f\left(t - 3\sqrt{2}m, y - \frac{l}{N}\right) \geq (\log N)^q \left(\log^-\left(t - 3\sqrt{2}m\right)\right)^p.$$

From this we get

$$\begin{aligned} M_{e_1}(g_N)(x,y) &\geq ⨍_I (g_N)_y \geq \frac{1}{3 \cdot 2^{l+1}\sqrt{2}} \sum_{m=2^l}^{2^{l+1}-1} \int_I f\left(t - 3\sqrt{2}m, y - \frac{l}{N}\right) dt \\ &\geq \frac{1}{6\sqrt{2}} (\log N)^q \int_{-1}^{1} \left(\log^- |t|\right)^p dt. \end{aligned}$$

At the end we note that for every function g, $M_{e_1}(g) \leq M_s(g)$ holds almost everywhere, and this proves the claim.

(iii) $M_{e_1}(g_N)(x,y) = 0$ for $y < -1$.

This holds simply because g_N is supported in $[3\sqrt{2}-1,\infty)\times[-1,2]$.

(iv) $M_s(g_N)(x,y) \lesssim 1$ for $0 \le x \le 1, y \le -2$.

To prove this final property of g_N, suppose $R = I\times J$ is a rectangle with $(x,y)\in R$. Then if $R \cap supp(g_N) \neq \emptyset$ we have $l(I), l(J) \geq 1$, and we note that

$$\#\left\{(m,k)|R\cap supp\left(f\left(\cdot - 3\sqrt{2}m, \cdot - \frac{k}{N}\right)\right)\neq\emptyset\right\} \lesssim l(I),$$

which implies

$$\fint_R g_N \le l(I)^{-1}\#\left\{(m,k)|R\cap supp\left(f\left(\cdot - 3\sqrt{2}m, \cdot - \frac{k}{N}\right)\right)\neq\emptyset\right\}\int_{\mathbb{R}^2} f \lesssim 1.$$

Now taking the supremum over all rectangles R proves the last property of g_N.

Next, we measure the mean oscillation of $M_{e_1}(g_N)$ on the square $[-3,3]^2$ by

$$O\left(M_{e_1}(g_N),[-3,3]^2\right) \gtrsim \fint_{[0,1]^2} M_{e_1}(g_N) - \fint_{[0,1]\times[-3,-2]} M_{e_1}(g_N).$$

Then from the second and third properties of g_N we obtain

$$O\left(M_{e_1}(g_N),[-3,3]^2\right) \gtrsim (\log N)^q\,, \tag{3.12}$$

and the same is true for M_s by the third and fourth properties of g_N.

At this point we note that the constructed sequence of functions $\{g_N\}$ has all the desired properties claimed in the theorem except that they are not bounded functions. However, this can be fixed by using a truncation argument as follows. For each $N, M \geq 1$, let $g_{N,M}$ be the truncation of g_N at height M, i.e.

$$g_{N,M} := \max\{M, \min\{g_N, -M\}\}.$$

Next, we note that by the first property of $\{g_N\}$ this sequence is bounded in $BMO(\mathbb{R}^2)$ and since $\|g_{N,M}\|_{BMO(\mathbb{R}^2)} \leq 4\|g_N\|_{BMO(\mathbb{R}^2)}$, the double sequence $\{g_{N,M}\}$ is also bounded in $BMO(\mathbb{R}^2)$. Now, for each $N \geq 1$, g_N is a compactly supported function in $L^2(\mathbb{R}^2)$ and the sequence $\{g_{N,M}\}$ converges to g_N in $L^2(\mathbb{R}^2)$ as M goes to infinity. Then, since the operators M_{e_1} and M_s are continuous on this space, we conclude that for each $N \geq 1$, $\{M_{e_1}(g_{N,M})\}$ and $\{M_s(g_{N,M})\}$ converge in $L^2(\mathbb{R}^2)$ to $M_{e_1}(g_N)$ and $M_s(g_N)$, respectively. Therefore, for N' large enough (depending on N), we have

$$O\left(M_{e_1}(g_{N,N'}), [-3,3]^2\right) \geq \frac{1}{2} O\left(M_{e_1}(g_N), [-3,3]^2\right) \gtrsim (\log N)^q,$$

and

$$O\left(M_s(g_{N,N'}), [-3,3]^2\right) \geq \frac{1}{2} O\left(M_s(g_N), [-3,3]^2\right) \gtrsim (\log N)^q.$$

To finish the proof, let $G_N := g_{N,N'}$ and note that $\{G_N\}$ is a sequence of bounded functions such that it is bounded in $BMO(\mathbb{R}^2)$ but

$$\lim_{N\to\infty} \|M_{e_1}(G_N)\|_{BMO(\mathbb{R}^2)} = \infty, \quad \lim_{N\to\infty} \|M_s(G_N)\|_{BMO(\mathbb{R}^2)} = \infty.$$

□

By modifying the above function, one can construct a function in $BMO(\mathbb{R}^2)$ such that none of its horizontal slices are in $BMO(\mathbb{R})$, which provides a negative answer to the forth question of this paper.

Example. Let $\{r_m\}$ be an enumeration of the rational numbers and consider the following function:

$$h(x,y) = \sum_{m\geq 1} f\left(x - 3\sqrt{2}m, y - r_m\right).$$

Then we have

$$O\left(h_y, [3\sqrt{2}m - 1, 3\sqrt{2}m + 1]\right) = O\left(\left(\log^-(\cdot)\right)^p, [-1,1]\right)\left(\log^-(y - r_m)\right)^q.$$

So by density of the rational numbers, for all values of y we get $\sup_{l(I)=2} O(h_y, I) = \infty$, even though $h \in BMO(\mathbb{R}^2)$.

The above example shows that one can not control the maximum mean oscillation of the slices, when we look at intervals with a fixed length. However, in the following theorem, we show that there is a loose control when the length of intervals increases.

Theorem 3.4.2. *Let $f \in BMO(\mathbb{R}^2)$ with $\|f\|_{BMO(\mathbb{R}^2)} = 1$. Then there exist constants $\lambda, c > 0$, independent of f, such that for any sequence of intervals I_k $(k \geq 1)$ with $l(I_k) = 2^k$, and any interval J with $l(J) = 1$, we have*

$$\int_J e^{\lambda \sup_{k\geq 1} \frac{O(f_y, I_k)}{k}} dy \leq c.$$

Proof. Let $E_t = \left\{y \in J | \sup_{k\geq 1} \frac{O(f_y,I_k)}{k} > t\right\}$; then,

$$E_t = \bigcup_{k\geq 1} E_{t,k}, \qquad E_{t,k} = \left\{y \in J | \frac{O\left(f_y, I_k\right)}{k} > t\right\}. \tag{3.13}$$

Now, taking the average over $E_{t,k}$ and applying Lemma 3.4.1 gives us

$$t < \frac{1}{k} \fint_{E_{t,k}} O\left(f_y, I_k\right) dy \lesssim \frac{1}{k} O\left(f, I_k \times E_{t,k}\right).$$

Next, let J_k be the interval with the same center as J and with $l(J_k) = 2^k$, and note that $E_{t,k} \subset J \subset J_k$, so $I_k \times E_{t,k} \subset I_k \times J_k$. Then an application of Lemma 3.3.1 shows that

$$t \lesssim \frac{1}{k} O\left(f, I_k \times E_{t,k}\right) \lesssim \frac{1}{k}\left(1 + \log \frac{|I_k \times J_k|}{|I_k \times E_{t,k}|}\right) \lesssim 1 - \frac{1}{k} \log |E_{t,k}|.$$

So for an appropriate constant $a > 0$, which is independent of f, we have $|E_{t,k}| \lesssim e^{-atk}$ for $t > 0$. From this and (3.13) we get the estimate

$$|E_t| \leq \sum_{k \geq 1} |E_{t,k}| \lesssim e^{-at}, \quad t > 0.$$

Now, an application of Cavalieri's principle gives us

$$\int_J e^{\frac{a}{2} \sup_{k \geq 1} \frac{O(f_y, I_k)}{k}} dy = \frac{a}{2} \int_0^\infty e^{\frac{a}{2} t} |E_t| dt \lesssim 1.$$

Hence, (3.4.2) holds with $\lambda = \frac{a}{2}$, and this finishes the proof. $\square$

Chapter 4

The Operator Norm of Paraproducts on Hardy Spaces

4.1 Introduction

Let Δ_j ($j \in \mathbb{Z}$) be the Littlewood-Paley operators associated with a Schwarz function ψ, and let S_j be its partial sum operators (precise definitions will be given in the next section). For an integer l and tempered distributions f and g, the bilinear form

$$\Pi(f,g) := \sum_{j\in\mathbb{Z}} S_{j-l}(f)\Delta_j(g),$$

is an example of a paraproduct. There are also dyadic counterparts of these bilinear forms, which will be discussed later. Paraproducts are among the most fundamental bilinear forms in harmonic analysis and PDEs. Loosely speaking, they are "half products," and historically, their first appearance was in Bony's para-differential calculus [8], where they were used to extend the work of Coifman and Meyer on pseudo-differential calculus with minimal regularity assumptions on their symbols [17]. Later, they were used in the proof of the $T(1)$ theorem by David and Journé [19], and since T. Figiel's work on the representation of singular integral operators in terms of simpler dyadic ones [24], they have proven to be an essential component of such operators [39]. It is well-known that for real $p > 0$ and sufficiently large l, the operator Π is bounded from $H^p(\mathbb{R}^n) \times BMO(\mathbb{R}^n)$ to $H^p(\mathbb{R}^n)$ [1, 32], where $H^p(\mathbb{R}^n)$ is the real Hardy space and $BMO(\mathbb{R}^n)$ is the space of functions of bounded mean oscillation [21, 41]. Also, for any triple of positive real numbers (p, q, r) with $\frac{1}{p} + \frac{1}{r} = \frac{1}{q}$, the bilinear form Π maps $H^p(\mathbb{R}^n) \times H^r(\mathbb{R}^n)$ to $H^q(\mathbb{R}^n)$ [4, 32]. In addition, the boundedness properties of analogs of these objects in other settings, including multi-parameter and multi-linear settings, have been studied extensively. We refer the reader to [31, 46, 56, 57] and the references therein. The purpose of the present paper is to show that these bounds cannot be improved in the sense that by freezing g, the operator norm of the corresponding operator, Π_g, is comparable to a norm of g predicted by the current bounds. See 4.3.1 and 4.4.1.

Let us begin by reviewing the current known results in this area. The first result on the operator norm of paraproducts seems to be in [7], where it is shown that for the dyadic paraproduct operator π_g, we have

$$\|\pi_g\|_{L^p(\mathbb{R}) \to L^p(\mathbb{R})} \simeq \|g\|_{BMO(\mathbb{R})}, \quad 1 < p < \infty.$$

Also, recently the author in [36] has shown among other things that for π_g it holds that

$$\|\pi_g\|_{L^p(\mathbb{R}^n) \to L^q(\mathbb{R}^n)} \simeq \|g\|_{\dot{L}^r(\mathbb{R}^n)}, \quad \frac{1}{q} = \frac{1}{p} + \frac{1}{r}, \quad 1 < q < p < \infty.$$

The result in [36] is stronger than what we have stated here and was obtained in the Bloom setting, but we are not concerned with that here. To the best of our knowledge, these are all the results known about the operator norm of paraproducts. The ideas used in [36] are similar to those methods employed in [38], where the author completed the characterization of the operator norm of commutators of a non-degenerate Calderón-Zygmund operator T and pointwise multiplication with a locally integrable function b. In fact, it is shown in [38] that for $1 < p, q < \infty$ we have:

$$\| [T, b] \|_{L^p(\mathbb{R}^n) \to L^q(\mathbb{R}^n)} \simeq \begin{cases} \|b\|_{\dot{L}^r(\mathbb{R}^n)} & \frac{1}{q} = \frac{1}{p} + \frac{1}{r}, \quad p > q \\ \|b\|_{BMO(\mathbb{R}^n)} & p = q \\ \|b\|_{\dot{C}^\alpha(\mathbb{R}^n)} & \frac{1}{q} = \frac{1}{p} - \frac{\alpha}{n}, \quad p < q \leq \frac{np}{(n-p)_+}, \end{cases}$$

where in the above $\dot{L}^r(\mathbb{R}^n)$ denotes $L^r(\mathbb{R}^n)$ modulo constants and $\dot{C}^\alpha(\mathbb{R}^n)$ is the homogeneous Hölder space. The strategy of the proof, used both in [38] and [36], is to bound the local mean oscillation of the symbol of the operator by testing it with suitable test functions. Then, using boundedness of the operator will finish the job for all cases except for $p > q$. To resolve this, the local mean oscillation inequality [48,50,51] is employed to obtain an appropriate sparse domination of the symbol. Then, after using a duality argument and a probabilistic linearization, the result follows from boundedness of the operator.

Here, although we are dealing with the operator from H^p to H^q, for all cases where $p \leq q$, our approach is essentially the same as in [7], [38], and [36]. To show that the symbol belongs to $BMO(\mathbb{R}^n)$ or $\dot{\Lambda}^\alpha(\mathbb{R}^n)$, we bound an appropriate mean oscillation of the operator's symbol. However, when $q < p$, and especially when $q < 1$, an argument based on duality will not work. This is because the Hahn-Banach theorem fails for H^q when $0 < q < 1$ [20], so we cannot guarantee that the operator norm of the operator and its adjoint are the same. Instead, by using a suitable sparse domination for the square function of the symbol g, we can handle all cases where $0 < q < p < \infty$. This sparse domination is not new; it is a special case of the general method formulated in [50]. This method shows that almost all known sparse domination results follow a general approach, which we use here. We should also mention that we first encountered this idea in [55], where it is used to construct an atomic decomposition for dyadic H^1, based on the square function. Notably, the construction idea in [55] is derived from [40], where it is shown that functions in Hardy spaces have an atomic decomposition into atoms, whose supporting cubes form a sparse family.

4.2 preliminaries

4.2.1 Notation

Throughout the paper, we use $\lesssim$, $\gtrsim$, and $\simeq$ to suppress constants and parameters in inequalities that are not crucial to our discussion; this will be clear from the context. We use $B_r(x)$ to denote the ball of radius r centered at x. A cube Q in $\mathbb{R}^n$ refers to a cube with sides parallel to the coordinate axes. We denote its Lebesgue measure by $|Q|$, its side length by $l(Q)$, and its dilation by a factor a as aQ. A dyadic cube is a cube Q of the form $Q = 2^k(m + [0,1]^n)$, where $k \in \mathbb{Z}$ and $m \in \mathbb{Z}^n$. For such a cube, all 2^n cubes obtained by bisecting its sides are also dyadic and are called its children. Any dyadic cube Q is a child of a unique cube, called the parent of Q, and denoted by $\hat{Q}$. We use $\mathcal{D}(Q)$ to denote the family of all dyadic cubes within Q, and $\mathcal{D}$ for all dyadic cubes in $\mathbb{R}^n$. For a locally integrable function f, its non-increasing rearrangement is denoted by

$$f^*(t) := \inf \{s \mid |\{|f| > s\}| \le t\}, \quad t > 0,$$

and its average with respect to the Lebesgue measure over a cube Q is given by

$$\langle f \rangle_Q := \frac{1}{|Q|} \int_Q f.$$

Additionally, for $0 < p < \infty$, we use $\operatorname{osc}_p(f, Q)$ to denote its p-mean oscillation and $\operatorname{osc}(f, Q)$ to denote its pointwise oscillation over Q:

$$\operatorname{osc}_p(f, Q) := \left\langle |f - \langle f \rangle_Q|^p \right\rangle_Q^{\frac{1}{p}}, \quad \operatorname{osc}(f, Q) := \sup_{x,y \in Q} |f(x) - f(y)|.$$

4.2.2 Sparse Families

Here, we review the definition of sparse families, one of their useful properties, and the general method of sparse domination introduced in [50].

Definition 4.2.1. *Let $0 < \eta < 1$, and let $\mathcal{C}$ be a family of cubes that are not necessarily dyadic. We say that $\mathcal{C}$ is η-*sparse *if, for each $Q \in \mathcal{C}$, there exists a subset $E_Q \subseteq Q$ such that $|E_Q| \geq \eta |Q|$, and for any two cubes $Q, Q' \in \mathcal{C}$, the sets E_Q and $E_{Q'}$ are disjoint.*

A useful property of sparse families, which is crucial for us here, is that although they may have many overlaps, they behave as if they are disjoint. A simple example of this phenomenon is given by the following lemma, which is well-known and whose proof can be found in [36, 38].

Lemma 4.2.1. *For an η-sparse family of cubes $\mathcal{C}$, nonnegative numbers $\{a_Q\}_{Q \in \mathcal{D}}$, and $0 < p < \infty$, we have*

$$\eta^{\frac{1}{p}} \left(\sum_{Q \in \mathcal{C}} a_Q^p |Q| \right)^{\frac{1}{p}} \lesssim \| \sum_{Q \in \mathcal{C}} a_Q \chi_Q \|_{L^p} \lesssim \eta^{-1} \left(\sum_{Q \in \mathcal{C}} a_Q^p |Q| \right)^{\frac{1}{p}}. \tag{4.1}$$

We now describe the general sparse domination method introduced in [50].

Let $\{f_Q\}_{Q \in \mathcal{D}}$ be a family of measurable functions defined on $\mathbb{R}^n$, which we consider as "localizations" of a function or operator. For every two dyadic cubes P and Q with $P \subseteq Q$, let $f_{P,Q}$ be a measurable function satisfying

$$|f_{P,Q}| \leq |f_P| + |f_Q|,$$

where $f_{P,Q}$ will serve as the "difference" between two localizations. Finally, for such a family of functions, the maximal sharp function is defined as

$$m_Q^{\#} f(x) := \sup_{\substack{x \in P \\ P \subseteq Q}} \operatorname{osc}(f_{P,Q}, P), \quad x \in Q. \tag{4.2}$$

Then, it is proved in [50] that the following holds:

Theorem 4.2.1 (Lerner, Lorist, Ombrosi)**.** *Let $\{f_Q\}_{Q \in \mathcal{D}}$ and $\{f_{P,Q}\}$ be as described above. For a dyadic cube Q_0 and $0 < \eta < 1$, there exists an η-sparse family of cubes $\mathcal{C}$ contained in Q_0 such that for almost every x in Q_0,*

$$f_{Q_0}(x) \lesssim \sum_{Q \in \mathcal{C}} \gamma_Q \chi_Q(x), \tag{4.3}$$

where

$$\gamma_Q = (f_Q \chi_Q)^* \left(\frac{1-\eta}{2^{n+2}} |Q| \right) + (m_Q^{\#} f)^* \left(\frac{1-\eta}{2^{n+2}} |Q| \right). \tag{4.4}$$

As shown in [50], almost all known sparse dominations are special cases of the general theorem above. We refer the reader to [50, 51] and the references therein for a general theory of sparse domination and dyadic calculus.

4.2.3 Dyadic Hardy Spaces

Next, we review the definitions and basic properties of the dyadic Hardy spaces that we are concerned with, beginning with the Haar basis.

For a dyadic interval I, let $h_I = |I|^{-\frac{1}{2}}(\chi_{I^-} - \chi_{I^+})$, where I^- and I^+ are the left and right halves of I, respectively. These functions form the well-known Haar basis for $L^2(\mathbb{R})$. In higher dimensions, the Haar basis is defined as follows: First, we let $h_I^1 = h_I$ and $h_I^0 = |I|^{-\frac{1}{2}}\chi_I$. Then, for a dyadic cube $Q = \prod_{j=1}^n I_j$ and a multi-index

$i = (i_1, \ldots, i_n) \neq 0$, where each $i_j \in \{0, 1\}$, we define

$$h_Q^i = \prod_{j=1}^{n} h_{I_j}^{i_j}.$$

For the Haar basis, the associated square function is defined as

$$S_d(f)(x) := \left(\sum_{Q \in \mathcal{D}} \sum_{i \neq 0} \left| \langle f, h_Q^i \rangle \right|^2 \frac{\chi_Q(x)}{|Q|} \right)^{\frac{1}{2}}, \quad x \in \mathbb{R}^n, \tag{4.5}$$

where $\langle \cdot, \cdot \rangle$ denotes the usual inner product in $L^2(\mathbb{R}^n)$. Another important dyadic operator is the dyadic maximal operator, defined as

$$M_d(f)(x) := \sup_{x \in Q} \left| \langle f \rangle_Q \right|, \tag{4.6}$$

where the supremum is taken over all dyadic cubes Q containing x.

Definition 4.2.2. *For $0 < p < \infty$, the dyadic Hardy space $H_d^p(\mathbb{R}^n)$ is the completion of the space of locally integrable functions f such that*

$$\|f\|_{H_d^p(\mathbb{R}^n)} := \|M_d(f)\|_{L^p(\mathbb{R}^n)} < \infty.$$

When $1 \leq p < \infty$, the dyadic Hardy space is a Banach space with the above quantity as its norm. For $0 < p < 1$, however, this quantity is a quasi-norm, making $H_d^p(\mathbb{R}^n)$ a quasi-Banach space. It is also well known that for $1 < p < \infty$, $H_d^p(\mathbb{R}^n)$ is identical to $L^p(\mathbb{R}^n)$. There is also a closely related space that is not identical to $H_d^p(\mathbb{R}^n)$, even though it is denoted by the same notation in the literature. To define it properly, we make the following definition:

Definition 4.2.3. *A dyadic distribution f is a family of complex numbers $\{f_Q^i : Q \in$*

$\mathcal{D}, i \neq 0\}$ *and is formally written as*

$$f = \sum_{Q \in \mathcal{D}} \sum_{i \neq 0} \langle f, h_Q^i \rangle h_Q^i, \quad \langle f, h_Q^i \rangle := f_Q^i.$$

Definition 4.2.4. *For $0 < p < \infty$, the dyadic Hardy space $\dot{H}_d^p(\mathbb{R}^n)$ is the space of all dyadic distributions f such that*

$$\|f\|_{\dot{H}_d^p(\mathbb{R}^n)} := \|S_d(f)\|_{L^p(\mathbb{R}^n)} < \infty.$$

The important fact here is that $\dot{H}_d^p(\mathbb{R}^n)$ is the same as $H_d^p(\mathbb{R}^n)$ modulo constants, meaning that for $f \in L^1_{loc}(\mathbb{R}^n)$ we have

$$\|f\|_{\dot{H}_d^p(\mathbb{R}^n)} \simeq_p \inf_{c \in \mathbb{C}} \|f - c\|_{H_d^p(\mathbb{R}^n)}. \tag{4.7}$$

Now, we turn to the definition of the dual of Hardy spaces.

Definition 4.2.5. *For $0 \leq \alpha < \infty$, the dyadic homogeneous Lipschitz space $\dot{\Lambda}_d^\alpha(\mathbb{R}^n)$ is the space of all locally integrable functions f, modulo constants, such that*

$$\|f\|_{\dot{\Lambda}_d^\alpha(\mathbb{R}^n)} := \sup_{Q \in \mathcal{D}} l(Q)^{-\alpha} osc_1(f, Q) < \infty.$$

In the above, $\dot{\Lambda}_d^0(\mathbb{R}^n)$ is identical with the dyadic BMO, denoted by $BMO_d(\mathbb{R}^n)$, or the space of functions with bounded mean oscillation on dyadic cubes. The first crucial fact about the above definition is that if, for a positive number p, we replace $\mathrm{osc}_1(f, Q)$ with $\mathrm{osc}_p(f, Q)$, we obtain nothing but the same space. More precisely, we have

$$\|f\|_{\dot{\Lambda}_d^\alpha(\mathbb{R}^n)} \simeq \sup_{Q \in \mathcal{D}} l(Q)^{-\alpha} \mathrm{osc}_p(f, Q) < \infty, \quad \alpha \geq 0, \quad p > 0. \tag{4.8}$$

Another important fact that we need later is that for $0 < p \leq 1$, the space $\dot{\Lambda}_d^\alpha(\mathbb{R}^n)$ is the dual of $H^p_d(\mathbb{R}^n)$. Specifically,

$$H^p_d(\mathbb{R}^n)^* \cong \dot{\Lambda}_d^{n(\frac{1}{p}-1)}(\mathbb{R}^n), \quad 0 < p \leq 1. \tag{4.9}$$

See [12, 26, 37] for the proof of these.

4.2.4 Dyadic Paraproducts

Now, we define and review the boundedness properties of dyadic paraproducts, which are the focus here.

Definition 4.2.6. *For a dyadic distribution g, the dyadic paraproduct operator with symbol g is defined as*

$$\pi_g(f) := \sum_{Q\in\mathcal{D}} \sum_{i\neq 0} \langle f\rangle_Q \langle g, h^i_Q\rangle h^i_Q, \quad f \in L^1_{loc}(\mathbb{R}^n). \tag{4.10}$$

In the above, $\pi_g(f)$ is a dyadic distribution, and when f is a linear combination of finitely many Haar functions, it is a well-defined function. The next theorem contains boundedness properties of π_g on dyadic Hardy spaces.

Theorem (A)**.** *For a dyadic distribution g and real numbers $0 < p, q, r < \infty$, the following inequalities hold:*

$$\|\pi_g(f)\|_{\dot{H}^q_d(\mathbb{R}^n)} \lesssim \|g\|_{\dot{H}^r_d(\mathbb{R}^n)} \|f\|_{H^p_d(\mathbb{R}^n)}, \quad \frac{1}{q} = \frac{1}{p} + \frac{1}{r} \tag{4.11}$$

$$\|\pi_g(f)\|_{\dot{H}^{p^*}_d(\mathbb{R}^n)} \lesssim \|g\|_{\dot{\Lambda}^\alpha_d(\mathbb{R}^n)} \|f\|_{H^p_d(\mathbb{R}^n)}, \quad \frac{1}{p^*} = \frac{1}{p} - \frac{\alpha}{n}, \quad 0 < \alpha p < n \tag{4.12}$$

$$\|\pi_g(f)\|_{\dot{H}^p_d(\mathbb{R}^n)} \lesssim \|g\|_{BMO_d(\mathbb{R}^n)} \|f\|_{H^p_d(\mathbb{R}^n)} \tag{4.13}$$

Let us briefly discuss the reasons behind these inequalities. The first inequality

(4.11) follows from the pointwise bound

$$S_d(\pi_g(f))(x) \leq M_d(f)(x) S_d(g)(x),$$

the Hölder inequality, and the maximal characterization of dyadic Hardy spaces (4.7). For inequality (4.12), we have an analog of the classical Hedberg inequality, which we could not find in the literature, so we decided to prove it here.

Proposition 4.2.1. *For $0 < \alpha n < p < \infty$, a dyadic distribution g, and a locally integrable function f, it holds*

$$S_d(\pi_g(f))(x) \lesssim \|g\|_{\dot{\Lambda}_d^{\alpha}(\mathbb{R}^n)} \|f\|_{H_d^p(\mathbb{R}^n)}^{\frac{\alpha p}{n}} M_d(f)(x)^{\frac{p}{p^*}}. \tag{4.14}$$

Proof. After normalization, we may assume $\|g\|_{\dot{\Lambda}_d^{\alpha}(\mathbb{R}^n)} = \|f\|_{H_d^p(\mathbb{R}^n)} = 1$. Let Q be a dyadic cube. The cancellation property of Haar functions ($\int h_Q^i = 0$), and the triangle inequality imply

$$|\langle g, h_Q^i \rangle| \leq \left\langle |g - \langle g \rangle_Q|, |h_Q^i| \right\rangle \leq \operatorname{osc}_1(g, Q)|Q|^{\frac{1}{2}} \leq |Q|^{\frac{\alpha}{n}+\frac{1}{2}}.$$

For $\langle f \rangle_Q$, we have the following two competing bounds:

$$|\langle f \rangle_Q| \leq \inf_{x \in Q} M_d(f)(x), \quad |\langle f \rangle_Q| \lesssim |Q|^{-1} \|f\|_{H_d^p(\mathbb{R}^n)} \|\chi_Q\|_{H_d^p(\mathbb{R}^n)^*}, \tag{4.15}$$

where the right-hand side inequality follows from duality. When $1 < p < \infty$, it is clear that $\|\chi_Q\|_{H_d^p(\mathbb{R}^n)^*} \leq |Q|^{1-\frac{1}{p}}$, and the duality relation (4.9) gives a similar bound holds for $0 < p \leq 1$. To see this, note that if $R \subseteq Q$ is a dyadic cube, then $\operatorname{osc}_1(\chi_Q, R) = 0$. For $Q \subsetneq R$, we have

$$\operatorname{osc}_1(\chi_Q, R) \leq 2 \langle \chi_Q \rangle_R \leq 2|R|^{-1}|Q| \leq 2|R|^{\frac{1}{p}-1}|Q|^{1-\frac{1}{p}},$$

which implies

$$\|\chi_Q\|_{\dot{\Lambda}_d^{n(\frac{1}{p}-1)}(\mathbb{R}^n)} \leq 2|Q|^{1-\frac{1}{p}}. \tag{4.16}$$

Combining (4.15) and (4.16), we get

$$|\langle f\rangle_Q| \lesssim \min\left\{\inf_{x\in Q} M_d(f)(x), |Q|^{-\frac{1}{p}}\right\}, \quad 0<p<\infty.$$

Finally, we estimate the square function as

$$\begin{aligned}
S_d(\pi_g(f))(x) &\leq \sum_{Q\in\mathcal{D}}\sum_{i\neq 0} |\langle f\rangle_Q| |\langle g, h_Q^i\rangle| |Q|^{-\frac{1}{2}}\chi_Q(x) \\
&\lesssim \sum_{Q\in\mathcal{D}} \min\left\{M_d(f)(x), |Q|^{-\frac{1}{p}}\right\} |Q|^{\frac{\alpha}{n}}\chi_Q(x) \\
&\leq M_d(f)(x) \sum_{M_d(f)(x)\leq |Q|^{-\frac{1}{p}}} |Q|^{\frac{\alpha}{n}}\chi_Q(x) + \sum_{M_d(f)(x)>|Q|^{-\frac{1}{p}}} |Q|^{\frac{\alpha}{n}-\frac{1}{p}}\chi_Q(x) \\
&\lesssim M_d(f)(x)^{\frac{p}{p^*}},
\end{aligned}$$

which proves the claim.

□

Now, taking the L^{p^*} norm of (4.14) proves (4.12). The last inequality (4.13) is the most complex one, so we only outline the main steps to prove it. The L^2 boundedness of π_g follows from the Haar characterization of $BMO_d(\mathbb{R}^n)$ in terms of Carleson measures and the Carleson embedding theorem. Boundedness on L^p comes from Calderón-Zygmund theory, and boundedness on H_d^p is derived from the atomic decomposition. We refer the reader to [55] for more on dyadic Hardy spaces and a proof of this. See also [7] for another proof of (4.13). We will now shift from the dyadic setting to the Fourier setting and provide the necessary definitions.

4.2.5 Littlewood-Paley Operators

We begin by fixing some notation. For a Schwartz function f, the Fourier transform is defined by

$$\hat{f}(\xi) := \int_{\mathbb{R}^n} f(x)e^{-2\pi i\langle x,\xi\rangle}dx,$$

and the convolution of two functions is defined as

$$f * g(x) := \int_{\mathbb{R}^n} f(x-y)g(y)dy.$$

Moreover, we use the following notation for translations and dilations of functions:

$$(\tau^{x_0}f)(x) := f(x-x_0), \quad \delta^t(f)(x) := f(t^{-1}x), \quad f_t(x) := t^{-n}\delta^t(f)(x), \quad t>0, \quad x_0, x \in \mathbb{R}^n,$$

and below we summarize some of their useful properties:

$$\tau^{x_0}(f*g) = (\tau^{x_0}f)*g = f*(\tau^{x_0}g), \quad f_t * \delta^s g = \delta^s(f_{s^{-1}t}g).$$

Now, let ψ be a Schwartz function with its Fourier transform supported in an annulus away from the origin and infinity, meaning

$$\operatorname{supp}(\hat{\psi}) \subseteq \{\boldsymbol{a} \leq |\xi| \leq \boldsymbol{b}\}, \quad 0 < \boldsymbol{a} < \boldsymbol{b} < \infty, \tag{4.17}$$

and such that

$$\sum_{j\in\mathbb{Z}} \hat{\psi}(2^{-j}\xi) = 1, \quad \xi \neq 0, \tag{4.18}$$

(see [29] for examples of such functions). Then, the Littlewood-Paley operators associated with ψ are defined as

$$\Delta_j^{\psi}(f) := \psi_{2^{-j}} * f, \quad j \in \mathbb{Z},$$

and the partial sum operators are defined as

$$S_j^{\psi}(f) := \sum_{k\leq j} \Delta_k^{\psi}(f), \quad S_j(f) = \Psi_{2^{-j}} * f,$$

where

$$\hat{\Psi} := \begin{cases} \sum_{k\leq 0} \hat{\psi}(2^k\xi) & \xi \neq 0 \\ 1 & \xi = 0 \end{cases}.$$

In the above, Ψ is a Schwartz function whose Fourier transform is supported in a ball around the origin and equals to 1 in a smaller neighborhood (throughout the paper, we use capital Greek letters for the kernel functions of the partial sum operators). Similarly to the dyadic case, the square function with respect to ψ is defined as

$$S_{\psi}(f)(x) := \left(\sum_{j\in\mathbb{Z}} |\Delta_j^{\psi}(f)(x)|^2 \right)^{\frac{1}{2}}.$$

For simplicity, we drop the dependence on ψ when it is clear from the context. An important feature of the Littlewood-Paley pieces $\Delta_j(f)$ is that their Fourier transform is localized at the scale 2^j, which means they behave like a constant at scale 2^{-j}. This feature can be expressed through Plancherel-Polya-Nikolskij type inequalities.

Theorem 4.2.2. *Let f be a tempered distribution whose Fourier transform is supported in a ball of radius $t > 0$. Then*

- *For $0 < p \leq q \leq \infty$, we have*

$$\|f\|_{L^q(\mathbb{R}^n)} \lesssim t^{n\left(\frac{1}{p}-\frac{1}{q}\right)} \|f\|_{L^p(\mathbb{R}^n)}. \tag{4.19}$$

- *There exists a constant c such that for $0 < h \leq ct^{-1}$, and any sequence of* *numbers* $\{x_k : x_k \in h([0,1]^n + k),\ k \in \mathbb{Z}^n\}$, *we have*

$$\|\{f(x_k)\}_{k\in\mathbb{Z}^n}\|_{l^p(\mathbb{Z}^n)} \simeq h^{-\frac{n}{p}}\|f\|_{L^p(\mathbb{R}^n)}, \quad 0 < p \leq \infty. \tag{4.20}$$

See [68] for the proof.

4.2.6 Real Hardy Spaces

Now, we recall the definition of the real Hardy spaces.

For a Schwartz function φ, the associated maximal operator M_φ and the non-tangential maximal operator M_φ^* are defined as

$$M_\varphi(f)(x) := \sup_{t>0} |\varphi_t * f(x)|, \quad M_\varphi^*(f)(x) := \sup_{t>0} \sup_{|x-y|\leq t} |\varphi_t * f(y)|.$$

Definition 4.2.7. *For a Schwartz function* φ *with* $\int \varphi = 1$ *and* $0 < p < \infty$, *the Hardy space* $H^p(\mathbb{R}^n)$ *is the space of all tempered distributions* f *such that*

$$\|f\|_{H^p(\mathbb{R}^n)} := \|M_\varphi(f)\|_{L^p(\mathbb{R}^n)} < \infty.$$

The above quantity defines a quasi-norm when $0 < p < 1$ and a norm when $1 \leq p < \infty$, which makes $H^p(\mathbb{R}^n)$ into a (quasi) Banach space. The space $H^p(\mathbb{R}^n)$ is independent of φ, and for any other choice of φ, the corresponding (quasi) norms are comparable to each other. Also, for any Schwartz function $\tilde{\varphi}$, we have

$$\|M_{\tilde{\varphi}}^*(f)\|_{L^p(\mathbb{R}^n)} \lesssim \|f\|_{H^p(\mathbb{R}^n)}, \quad 0 < p < \infty,$$

and if $\int \tilde{\varphi} = 1$, the above inequality becomes an equivalence, with bounds depending only on p, n, and finitely many Schwartz semi-norms of φ and $\tilde{\varphi}$. Furthermore, with a minor difference from the dyadic case, the square function characterization holds as well. To be more precise, let $\dot{H}^p(\mathbb{R}^n)$ be the Hardy space $H^p(\mathbb{R}^n)$ modulo polynomials, or equivalently, the space of all tempered distributions with

$$\|f\|_{\dot{H}^p(\mathbb{R}^n)} := \inf_{P\in\mathbb{P}_n} \|f-P\|_{H^p(\mathbb{R}^n)} < \infty,$$

where $\mathbb{P}_n$ is the space of all polynomials on $\mathbb{R}^n$. Then, for any choice of Littlewood-Paley operators $\{\Delta_j\}_{j\in\mathbb{Z}}$, we have

$$\|S(f)\|_{L^p(\mathbb{R}^n)} \simeq \|f\|_{\dot{H}^p(\mathbb{R}^n)}, \quad 0<p<\infty.$$

Finally, just as in the dyadic case, for $1<p<\infty$, $H^p(\mathbb{R}^n)$ coincides with $L^p(\mathbb{R}^n)$ [21, 30, 66]. Next, we recall the duals of Hardy spaces.

Definition 4.2.8. *For $0<\alpha<\infty$, the homogeneous Lipschitz space $\dot{\Lambda}^\alpha(\mathbb{R}^n)$ is the Banach space of all functions f with*

$$\|f\|_{\dot{\Lambda}^\alpha(\mathbb{R}^n)} := \sup_{x\in\mathbb{R}^n}\sup_{h\neq 0} |h|^{-\alpha}|D_h^{[\alpha]+1}(f)(x)| < \infty,$$

where D_h is the forward difference operator defined as $D_h(f)(x) = f(x+h)-f(x)$, and $[\alpha]$ denotes the largest integer not greater than α.

Modulo polynomials, the homogeneous Lipschitz space $\dot{\Lambda}^\alpha(\mathbb{R}^n)$ can be characterized in terms of Littlewood-Paley pieces. Specifically, let $\ddot{\Lambda}^\alpha(\mathbb{R}^n)$ be the space of all tempered distributions f with

$$\|f\|_{\ddot{\Lambda}^\alpha(\mathbb{R}^n)} := \inf_{P\in\mathbb{P}_n} \|f-P\|_{\dot{\Lambda}^\alpha(\mathbb{R}^n)} < \infty.$$

Then, for any choice of Littlewood-Paley operators, we have

$$\|f\|_{\ddot{\Lambda}^\alpha(\mathbb{R}^n)} \simeq \sup_{j\in\mathbb{Z}} 2^{\alpha j}\|\Delta_j(f)\|_{L^\infty(\mathbb{R}^n)}.$$

Similarly to the dyadic case, we have

$$H^p(\mathbb{R}^n)^* \cong \dot{\Lambda}^{n\left(\frac{1}{p}-1\right)}(\mathbb{R}^n), \quad 0 < p < 1, \quad H^1(\mathbb{R}^n)^* \cong BMO(\mathbb{R}^n),$$

where $BMO(\mathbb{R}^n)$ is the space of functions of bounded mean oscillation equipped with the norm

$$\|f\|_{BMO(\mathbb{R}^n)} := \sup_Q \operatorname{osc}_1(f, Q),$$

where the sup is taken over all cubes in $\mathbb{R}^n$. It is well-known that for any positive p, if we replace osc_1 with osc_p, we get an equivalent norm [25, 30, 66]. It is also useful to define $B\dot{M}O(\mathbb{R}^n)$, or $BMO(\mathbb{R}^n)$ modulo polynomials, with the usual definition of the norm

$$\|f\|_{B\dot{M}O(\mathbb{R}^n)} := \inf_{P \in \mathbb{P}_n} \|f - P\|_{BMO(\mathbb{R}^n)}.$$

4.2.7 Fourier Paraproducts

Finally, we discuss the type of paraproducts mentioned in the introduction.

Definition 4.2.9. *Let ψ be a Schwartz function satisfying (4.17). For a tempered distribution g and a Schwartz function φ, the paraproduct operator $\Pi_{g,\varphi}$ with symbol g is formally defined as*

$$\Pi_{g,\varphi}(f) := \sum_{j \in \mathbb{Z}} (\varphi_{2^{-j}} * f) \Delta_j^{\psi}(g).$$

The meaning of convergence in the above sum is not clear unless we impose some restrictions on f and φ. One situation where the above operator is well-defined is when the support of the Fourier transform of φ lies in a compact subset of $\mathbb{R}^n$. In this case, for a Schwartz function f with Fourier transform supported in an annulus, the sum is finite and hence yields a well-defined smooth function. Furthermore, when the support of $\hat{\varphi}$ lies in a ball strictly within the annulus containing the support of

$\hat{\psi}$, the boundedness of this operator on various Hardy spaces is well-known and we present it here as a theorem [30, 32, 66].

Theorem (B). *Suppose φ is a Schwartz function whose Fourier transform is supported in a ball around the origin with radius $\boldsymbol{a}' < \boldsymbol{a}$, where $\boldsymbol{a}$ is as in (4.17). Then, for a distribution g and real numbers $0 < p, q, r < \infty$, the following inequalities hold:*

$$\|\Pi_{g,\varphi}(f)\|_{\dot{H}^q(\mathbb{R}^n)} \lesssim \|g\|_{\dot{H}^r(\mathbb{R}^n)} \|f\|_{H^p(\mathbb{R}^n)}, \quad \frac{1}{q} = \frac{1}{p} + \frac{1}{r}, \tag{4.21}$$

$$\|\Pi_{g,\varphi}(f)\|_{\dot{H}^{p^*}(\mathbb{R}^n)} \lesssim \|g\|_{\dot{\Lambda}^\alpha(\mathbb{R}^n)} \|f\|_{H^p(\mathbb{R}^n)}, \quad \frac{1}{p^*} = \frac{1}{p} - \frac{\alpha}{n}, \quad 0 < \alpha p < n, \tag{4.22}$$

$$\|\Pi_{g,\varphi}(f)\|_{\dot{H}^p(\mathbb{R}^n)} \lesssim \|g\|_{B\dot{M}O(\mathbb{R}^n)} \|f\|_{H^p(\mathbb{R}^n)}. \tag{4.23}$$

The reasons for the validity of these inequalities are similar to those in the dyadic case, although there is a minor difference. The source of this difference is that in the definition of $\Pi_{g,\varphi}$, there is always some overlap between the Fourier supports of consecutive terms. Consequently, we cannot guarantee that a term like $(\varphi_{2^{-j}} * f)\Delta_j(g)$ is a Littlewood-Paley piece of $\Pi_{g,\varphi}$. However, since the Fourier support of the product of two functions is contained within the algebraic sum of their Fourier supports, for $\boldsymbol{a}' < \boldsymbol{a}$, the Fourier support of $(\varphi_{2^{-j}} * f)\Delta_j(g)$ remains away from the origin and around the annulus where the Fourier transform of $\Delta_j(g)$ is supported. Therefore, for a sufficiently large natural number m depending only on $\boldsymbol{a}'$, $\boldsymbol{a}$, and $\boldsymbol{b}$, the Fourier supports of the terms in

$$\Pi_{i,g,\varphi}(f) := \sum_{j \in m\mathbb{Z}+i} (\varphi_{2^{-j}} * f)\Delta_j(g), \quad 0 \le i < m \tag{4.24}$$

are all sufficiently far from each other. Thus, by choosing an appropriate Littlewood-Paley operator $\{\Delta_j^\theta\}_{j\in\mathbb{Z}}$ such that $\hat{\theta}$ equals 1 in a neighborhood of the support of $\hat{\psi}$, we have

$$\Delta_k^\theta(\varphi_{2^{-j}} * f\Delta_j(g)) = \delta_{k,j}\varphi_{2^{-j}} * f\Delta_j(g), \quad k \in \mathbb{Z}, \quad j \in m\mathbb{Z}+i.$$

This implies that

$$S_\theta(\Pi_{i,g,\varphi})(f) = \left(\sum_{j\in m\mathbb{Z}+i} |\varphi_{2^{-j}} * f\Delta_j(g)|^2\right)^{\frac{1}{2}}, \quad 0 \leq i < m.$$

Arguments similar to those used in the dyadic case can be applied to the operators $\Pi_{i,g,\varphi}$. Since

$$\Pi_{g,\varphi} = \sum_{0\leq i<m} \Pi_{i,g,\varphi},$$

we conclude that the same results hold for $\Pi_{g,\varphi}$ [30]. Having established our notation, provided the necessary definitions, and recalled the essential facts, we now proceed to the next section of this article.

4.3 The operator norm of Dyadic paraproducts on dyadic hardy spaces

Our main results in this section are as follows:

Theorem 4.3.1. *Let g be a dyadic distribution. Then we have*

$$\|\pi_g\|_{H_d^p(\mathbb{R}^n)\to\dot{H}_d^q(\mathbb{R}^n)} \simeq \|g\|_{\dot{H}_d^r(\mathbb{R}^n)}, \quad \frac{1}{q} = \frac{1}{p} + \frac{1}{r}, \quad 0 < p,q,r < \infty, \tag{i}$$

$$\|\pi_g\|_{H_d^p(\mathbb{R}^n)\to\dot{H}_d^{p^*}(\mathbb{R}^n)} \simeq \|g\|_{\dot{\Lambda}_d^\alpha(\mathbb{R}^n)}, \quad \frac{1}{p^*} = \frac{1}{p} - \frac{\alpha}{n}, \quad 0 \leq \alpha p < n, \quad 0 < p < \infty. \tag{ii}$$

To prove (i), we need the following rather general theorem.

Theorem 4.3.2. *Let $\{g_Q\}_{Q\in\mathcal{D}}$ be a sequence of nonnegative numbers indexed by* *dyadic cubes, where* $0 < q, r, s, p < \infty$ *and* $\frac{1}{q} = \frac{1}{p} + \frac{1}{r}$. *Suppose there exists a constant* A *such that for all step functions* f *with compact support, the following inequality holds:*

$$\left\| \sum_{Q \in \mathcal{D}} |\langle f, \chi_Q \rangle|^s g_Q \chi_Q \right\|_{L^q(\mathbb{R}^n)} \leq A \|f\|^s_{H_d^{sp}(\mathbb{R}^n)}. \tag{4.25}$$

Then we have

$$\left\| \sum_{Q \in \mathcal{D}} g_Q \chi_Q \right\|_{L^r(\mathbb{R}^n)} \lesssim A. \tag{4.26}$$

Let us accept this and prove Theorem 4.3.1.

Proof of Theorem 4.3.1. The upper bounds for the operator norm of π_g are covered by Theorem (A), so we need to prove the lower bounds.

Case (i): For the inequality (i), let $A = \|\pi_g\|_{H_d^p(\mathbb{R}^n) \to \dot{H}_d^q(\mathbb{R}^n)}$. This means

$$\|S_d(\pi_g(f))\|_{L^q(\mathbb{R}^n)} \leq A \|f\|_{H_d^p(\mathbb{R}^n)},$$

which is equivalent to

$$\left\| \sum_{Q \in \mathcal{D}} \sum_{i \neq 0} |\langle f, \chi_Q \rangle|^2 \langle g, h_Q^i \rangle^2 \frac{\chi_Q}{|Q|} \right\|_{L^{\frac{q}{2}}(\mathbb{R}^n)} \leq A^2 \|f\|^2_{H_d^p(\mathbb{R}^n)}.$$

For each fixed $i \neq 0$, we have

$$\left\| \sum_{Q \in \mathcal{D}} |\langle f, \chi_Q \rangle|^2 \langle g, h_Q^i \rangle^2 \frac{\chi_Q(x)}{|Q|} \right\|_{L^{\frac{q}{2}}(\mathbb{R}^n)} \leq A^2 \|f\|^2_{H_d^p(\mathbb{R}^n)},$$

which matches the assumption (4.25) in Theorem 4.3.2 with

$$g_Q = \frac{\langle g, h_Q^i \rangle^2}{|Q|}, \quad s = 2, \quad \frac{1}{\frac{q}{2}} = \frac{1}{\frac{p}{2}} + \frac{1}{\frac{r}{2}}.$$

Thus, by (4.26), we get

$$\left\| \sum_{Q\in\mathcal{D}} \frac{\langle g, h_Q^i\rangle^2 \chi_Q(x)}{|Q|} \right\|_{L^{\frac{r}{2}}(\mathbb{R}^n)} \lesssim A^2.$$

Summing over $i \neq 0$ and using the (quasi) triangle inequality, we obtain

$$\|S_d(g)\|_{L^r(\mathbb{R}^n)} = \|g\|_{\dot{H}_d^r(\mathbb{R}^n)} \lesssim A,$$

proving the claim for Case (i).

Case (ii): For the inequality (ii), let $A = \|\pi_g\|_{H_d^p(\mathbb{R}^n)\to \dot{H}_d^{p^*}(\mathbb{R}^n)}$. We first consider the case where g has a finite Haar expansion. For a dyadic cube R, choose $i \neq 0$ so that $f = |\hat{R}|^{\frac{1}{2}} h_{\hat{R}}^i$ is equal to 1 on R. Then

$$\pi_g(f) = \sum_{Q\subseteq R}\sum_{j\neq 0} \langle g, h_Q^j\rangle h_Q^j + \sum_{Q\nsubseteq R}\sum_{j\neq 0} \langle f, \chi_Q\rangle \langle g, h_Q^j\rangle h_Q^j = g_1 + g_2,$$

which implies

$$\|g_1\|_{\dot{H}_d^{p^*}(\mathbb{R}^n)} \leq \|\pi_g(f)\|_{\dot{H}_d^{p^*}(\mathbb{R}^n)} \leq A\|f\|_{H_d^p(\mathbb{R}^n)} = A|\hat{R}|^{\frac{1}{p}} \lesssim A|R|^{\frac{1}{p}}.$$

Here, $g_1 = (g - \langle g\rangle_R)\chi_R$. Since $|g_1(x)| \leq M_d(g_1)(x)$ for $x \in \mathbb{R}^n$, by the maximal characterization of $\dot{H}_d^{p^*}(\mathbb{R}^n)$, we have

$$\|g_1\|_{L^{p^*}(\mathbb{R}^n)} \leq \|M_d(g_1)\|_{L^{p^*}(\mathbb{R}^n)} \lesssim \|g_1\|_{\dot{H}^{p^*}(\mathbb{R}^n)} \lesssim A|R|^{\frac{1}{p}},$$

which is equivalent to

$$\operatorname{osc}_{p^*}(g, R) \lesssim Al(R)^{\alpha}. \tag{4.27}$$

Combining this with (5.7) yields $\|g\|_{\dot{\Lambda}_d^{\alpha}(\mathbb{R}^n)} \lesssim A$, proving the claim for g with a finite

Haar expansion. To remove the restriction on g, let N be a natural number and define

$$g_N = \sum_{2^{-N} \leq l(Q) \leq 2^N} \sum_{j \neq 0} \langle g, h_Q^j \rangle h_Q^j.$$

Note that $\|\pi_{g_N}\|_{H_d^p(\mathbb{R}^n) \to \dot{H}_d^{p^*}(\mathbb{R}^n)} \leq A$, which implies

$$\|g_N\|_{\dot{\Lambda}_d^\alpha(\mathbb{R}^n)} \lesssim A.$$

For a fixed dyadic cube R, using (4.27) and (5.7), we get

$$\|g'_N\|_{L^2(R, \frac{dx}{|R|})} \lesssim A l(R)^\alpha,$$

where

$$g'_N = (g_N - \langle g_N \rangle_R)\chi_R = \sum_{Q \subseteq R} \sum_{j \neq 0} \langle g_N, h_Q^j \rangle h_Q^j.$$

By the weak compactness of L^2, a subsequence of g'_N converges to a function g'_R with

$$\|g'_R\|_{L^2(R, \frac{dx}{|R|})} \lesssim A l(R)^\alpha.$$

Since $\langle g'_N, h_Q^i \rangle$ converges to $\langle g'_R, h_Q^i \rangle$, we conclude that g'_R coincides with g on R. Hence, g is a locally integrable function, and on each dyadic cube R, it satisfies $\text{osc}_2(g, R) \lesssim A l(R)^\alpha$, and thus,

$$\|g\|_{\dot{\Lambda}_d^\alpha(\mathbb{R}^n)} \lesssim A,$$

which completes the proof. □

Now, we proceed with the proof of Theorem 4.3.2. To do this, we use the following lemma, which is similar to the construction in [55] (Theorem 1.2.4). Here, it is more convenient to introduce a notation. For a family of nonnegative numbers $\{g_Q\}_{Q \in \mathcal{D}}$

and a dyadic cube R, we define

$$(g|R)(x) := \sum_{Q \in \mathcal{D}(R)} g_Q \chi_Q(x).$$

Lemma 4.3.1. *Let $\{g_Q\}_{Q\in\mathcal{D}}$ be a family of nonnegative numbers, Q_0 a fixed dyadic cube, and $0 < \eta < 1$. Then, there exists an η-sparse family of dyadic cubes $\mathcal{C}$ in Q_0 with the property that:*

For each $Q \in \mathcal{C}$, there exists an integer λ_Q such that

$$(g|Q_0) \lesssim \sum_{Q\in\mathcal{C}} 2^{\lambda_Q} \chi_Q \tag{4.28}$$

$$|Q| \lesssim |\left\{(g|Q) > 2^{\lambda_Q - 1}\right\}|. \tag{4.29}$$

Proof. Let

$$f_Q = (g|Q), \quad f_{P,Q} = f_Q - f_P = \sum_{P \subsetneq R \subseteq Q} g_R \chi_R, \quad P \in \mathcal{D}(Q),$$

be as in Theorem 4.2.1. Then, we note that the condition

$$|f_{P,Q}| \leq |f_Q| + |f_P|,$$

holds, and therefore, an application of Theorem 4.2.1 gives a family of cubes in Q_0, which is η-sparse and such that

$$(g|Q_0) \lesssim \sum_{Q\in\mathcal{C}} \gamma_Q \chi_Q, \tag{4.30}$$

where in the above

$$\gamma_Q = (f_Q\chi_Q)^*(\frac{1-\eta}{2^{n+2}}|Q|) + (m_Q^{\#}f)^*(\frac{1-\eta}{2^{n+2}}|Q|).$$

Here, the first thing to note is that $f_{P,Q}$ is constant on P and thus $\mathrm{osc}(f_{P,Q}, P) = 0$, which implies that $m_Q^{\#}f$ vanishes on Q. Now, for $Q \in \mathcal{C}$, if $(f_Q\chi_Q)^*(\frac{1-\eta}{2^{n+2}}|Q|)$ is zero it doesn't appear in (4.30), and we remove Q from $\mathcal{C}$, and if not let λ_Q be an integer such that

$$2^{\lambda_Q - 1} < (f_Q\chi_Q)^*(\frac{1-\eta}{2^{n+2}}|Q|) \leq 2^{\lambda_Q},$$

then (4.28) holds, and from the definition of non-increasing rearrangement we must have

$$\frac{1-\eta}{2^{n+2}}|Q| < |\{f_Q > 2^{\lambda_Q - 1}\}| = |\left\{(g|Q) > 2^{\lambda_Q - 1}\right\}|,$$

which shows that (4.29) holds as well, and this completes the proof. □

We break the proof of Theorem 4.3.2 into two parts, depending on whether $1 < sp < \infty$, in which case our argument heavily relies on sparseness, or $0 < sp \leq 1$, where instead of sparseness, we take advantage of the sub-additivity of the L^{sp} norm.

Proof of Theorem 4.3.2. Case (1). $1 < sp < \infty$:

First, assume that there are only finitely many nonzero coefficients in $\{g_R\}_{R\in\mathcal{D}}$, and let Q_0 be a dyadic cube. Then, an application of Lemma 4.3.1 with $\eta = \frac{1}{2}$ gives us a $\frac{1}{2}$-sparse collection of dyadic cubes $\mathcal{C}$, satisfying (4.28) and (4.29). Now, observe that from Lemma 4.2.1 it follows that

$$\|(g|Q_0)\|^r_{L^r(\mathbb{R}^n)} \lesssim \sum_{Q\in\mathcal{C}} 2^{r\lambda_Q}|Q|. \tag{4.31}$$

Then, let

$$T(f) = \sum_{R\in\mathcal{D}} |\langle f\rangle_R|^s g_R \chi_R, \quad f = \sum_{Q\in\mathcal{C}} 2^{t\lambda_Q}\chi_Q, \quad t = \frac{r}{sp},$$

and note that

$$\langle f\rangle_R \geq 2^{t\lambda_Q}, \quad R \in \mathcal{D}(Q), \quad Q \in \mathcal{C},$$

which implies that

$$\left\{(g|Q) > 2^{\lambda_Q - 1}\right\} \subseteq \left\{T(f) > 2^{(st+1)\lambda_Q - 1}\right\} \cap Q. \tag{4.32}$$

Also, another application of Lemma 4.2.1 gives us

$$\|f\|_{L^{sp}(\mathbb{R}^n)} \simeq \left(\sum_{Q\in\mathcal{C}} 2^{r\lambda_Q}|Q|\right)^{\frac{1}{sp}}. \tag{4.33}$$

Now, we proceed to estimate the right-hand side of (4.31). To this aim, let us partition $\mathcal{C}$ as

$$\mathcal{C}'_k := \{Q \in \mathcal{C} \mid \lambda_Q = k\}, \quad k \in \mathbb{Z},$$

which implies that

$$\sum_{Q\in\mathcal{C}} 2^{r\lambda_Q}|Q| = \sum_{k\in\mathbb{Z}} 2^{kr} \sum_{Q\in\mathcal{C}'_k} |Q|. \tag{4.34}$$

Now, let $\mathcal{C}''_k$ be the collection of maximal cubes in $\mathcal{C}'_k$, and note that since this collection is $\frac{1}{2}$-sparse, we can estimate the last sum as

$$\sum_{Q\in\mathcal{C}'_k} |Q| \leq 2 \sum_{Q\in\mathcal{C}'_k} |E_Q| \leq 2\,|\cup\{Q \in \mathcal{C}'_k\}| = 2 \sum_{Q\in\mathcal{C}''_k} |Q|. \tag{4.35}$$

Also, from the second property (4.29) of the cubes in $\mathcal{C}$, together with (4.32), we obtain

$$\sum_{Q\in\mathcal{C}''_k}|Q| \lesssim \sum_{Q\in\mathcal{C}''_k}\left|\left\{(g|Q) > 2^{k-1}\right\}\right| \le \sum_{Q\in\mathcal{C}''_k}\left|\left\{T(f) > 2^{(st+1)k-1}\right\}\cap Q\right|, \quad k\in\mathbb{Z},$$

and then by noting that cubes in $\mathcal{C}''_k$ are disjoint, we get the following estimate:

$$\sum_{Q\in\mathcal{C}''_k}|Q| \lesssim \left|\left\{T(f) > 2^{(st+1)k-1}\right\}\right|, \quad k\in\mathbb{Z}.$$

Putting the above inequality together with (4.34) and (4.35), we obtain

$$\sum_{Q\in\mathcal{C}} 2^{r\lambda_Q}|Q| \lesssim \sum_{k\in\mathbb{Z}} 2^{kr}\left|\left\{T(f) > 2^{(st+1)k-1}\right\}\right|,$$

which, after noting that $st+1=\frac{r}{q}$, and using the layer cake formula implies that

$$\sum_{Q\in\mathcal{C}} 2^{r\lambda_Q}|Q| \lesssim \int \sum_{2^{\frac{kr}{q}}<2T(f)(x)} 2^{kr}\,dx \lesssim \int_{\mathbb{R}^n} T(f)^q.$$

At the end, we use (4.25) and (4.33) to obtain

$$\sum_{Q\in\mathcal{C}} 2^{r\lambda_Q}|Q| \lesssim \int_{\mathbb{R}^n} T(f)^q \le A^q\|f\|_{L^{sp}}^{sq} \lesssim A^q\left(\sum_{Q\in\mathcal{C}} 2^{r\lambda_Q}|Q|\right)^{\frac{q}{p}},$$

which, together with (4.31), gives us

$$\|(g|Q_0)\|_{L^r(\mathbb{R}^n)} \lesssim \left(\sum_{Q\in\mathcal{C}} 2^{r\lambda_Q}|Q|\right)^{\frac{1}{r}} \lesssim A.$$

Now, since there are only finitely many nonzero terms in $\{g_R\}_{R\in\mathcal{D}}$, for 2^n large dyadic cubes in each octant of $\mathbb{R}^n$, we have

$$\sum_{R\in\mathcal{D}} g_R\chi_R = \sum_{i=1}^{2^n}(g|Q_i), \quad Q_i = [0, \pm 2^N]^n.$$

Then, applying the above inequality to each cube Q_i, and using the (quasi) triangle inequality, we get

$$\|\sum_{R\in\mathcal{D}} g_R\chi_R\|_{L^r(\mathbb{R}^n)} \lesssim A,$$

and this proves the claim when there are only finitely many nonzero terms. To remove this restriction, let

$$(g_N)_R = \begin{cases} g_R & 2^{-N} \leq l(R) \leq 2^N, \\ 0 & \text{Otherwise}, \end{cases}$$

Then from the fact that the assumption (4.25) still holds with A, and the above inequality, we get

$$\|\sum_{2^{-N}\leq l(R)\leq 2^N} g_R\chi_R\|_{L^r(\mathbb{R}^n)} \lesssim A,$$

which, after an application of Fatou's lemma, gives the desired conclusion and completes the proof of the first case.

Case(2). $0 < sp \leq 1$. Here, we cannot use the function f, constructed above, as a test function because functions in $H_d^{sp}(\mathbb{R}^n)$ must have lots of cancellations. However, in this case, the sparseness of the family is not necessary, and we may use sub-additivity which helps us to repeat a similar argument.

So, as in the previous case, suppose only finitely many terms in $\{g_R\}_{R\in\mathcal{D}}$ are nonzero. This time, let

$$G := \sum_{R\in\mathcal{D}} g_R\chi_R,$$

and for $k\in\mathbb{Z}$, let $\mathcal{C}_k$ be the collection of maximal dyadic cubes in $\{G > 2^k\}$. From the layer cake formula we have

$$\|G\|_{L^r(\mathbb{R}^n)}^r \simeq \sum_{k\in\mathbb{Z}} 2^{rk}|\{G > 2^k\}| = \sum_{k\in\mathbb{Z}} 2^{rk} \sum_{Q\in\mathcal{C}_k} |Q|. \tag{4.36}$$

Now, just like in the previous case, we try to estimate the last sum. In order to do this, let $\hat{\mathcal{C}}_k$ be the collection of maximal cubes in $\{\hat{Q} \mid Q \in \mathcal{C}_k\}$, and for each $Q' \in \hat{\mathcal{C}}_k$, let $\tilde{\chi}_{Q'} = |Q'|^{\frac{1}{2}} h^i_{Q'}$ for some $i \neq 0$. We note that this function is either $+1$ or -1 on the children of Q', and belongs to $H^{sp}_d(\mathbb{R}^n)$ with $\|\tilde{\chi}_{Q'}\|^{sp}_{H^{sp}_d(\mathbb{R}^n)} = |Q'|$. Then, let

$$T(f) = \sum_{R\in\mathcal{D}} |\left\langle f \right\rangle_R|^s g_R \chi_R, \quad f = \sum_{k\in\mathbb{Z}} 2^{kt} \sum_{Q\in\hat{\mathcal{C}}_k} \tilde{\chi}_Q \quad t = \frac{r}{sp}.$$

The above function, f, has the following two crucial properties:

$$\|f\|^{sp}_{H^{sp}_d(\mathbb{R}^n)} \lesssim \sum_{j\in\mathbb{Z}} 2^{rk} \sum_{Q\in\mathcal{C}_k} |Q|, \tag{4.37}$$

$$|\left\langle f \right\rangle_R| \gtrsim 2^{kt}, \quad R \in \mathcal{D}(Q), \quad Q \in \mathcal{C}_k, \quad k \in \mathbb{Z}. \tag{4.38}$$

To see the first one, note that from sub-additivity it follows

$$\|f\|^{sp}_{H^{sp}_d(\mathbb{R}^n)} \leq \sum_{k\in\mathbb{Z}} 2^{ktsp} \sum_{Q'\in\hat{\mathcal{C}}_k} \|\tilde{\chi}_{Q'}\|^{sp}_{H^{sp}_d(\mathbb{R}^n)} \leq 2^n \sum_{k\in\mathbb{Z}} 2^{rk} \sum_{Q\in\mathcal{C}_k} |Q|.$$

In order to see the second property, let $Q \in \mathcal{C}_k$ and fix $R \in \mathcal{D}(Q)$. Then, we decompose f as

$$f(x) = \sum_{j\in\mathbb{Z}} 2^{jt} \sum_{\substack{Q'\in\hat{\mathcal{C}}_j \\ Q'\subseteq R}} \tilde{\chi}_{Q'}(x) + \sum_{j\in\mathbb{Z}} 2^{jt} \sum_{\substack{Q'\in\hat{\mathcal{C}}_j \\ R\subsetneq Q'}} \tilde{\chi}_{Q'}(x) = f_1(x) + f_2(x), \quad x \in R.$$

Now, because of the cancellation of the functions $\tilde{\chi}_{Q'}$, we have $\left\langle f_1 \right\rangle_R = 0$. Furthermore, since $\tilde{\chi}_{Q'}$, is either $+1$ or -1 on the children of Q', f_2 is constant on R. Next, we note that for each $j \in \mathbb{Z}$, R is contained in at most one cube in $\hat{\mathcal{C}}_j$, and when this inclusion is strict, the contribution of each term in the right-hand sum on R is either $+2^{jt}$ or -2^{jt}. Therefore,

$$f_2 = \sum_{\substack{Q' \in \hat{\mathcal{C}}_j \\ R \subsetneq Q'}} \pm 2^{jt} \simeq \pm 2^{jl},$$

where in the above l is the largest $j \in \mathbb{Z}$ such that R is strictly contained in a cube $Q' \in \hat{\mathcal{C}}_j$. Then, since $R \subseteq Q \subsetneq \hat{Q}$, and $\hat{Q}$ is contained in a cube in $\hat{\mathcal{C}}_k$, we conclude that $l \geq k$, and this shows that

$$|\langle f \rangle_R| = |\langle f_2 \rangle_R| \simeq 2^{lt} \geq 2^{kt},$$

which proves the second property of f. Next, we proceed to estimate the measure of the level sets of G in terms of $T(f)$. So, let $Q \in \mathcal{C}_k$. We claim that

$$Q \subseteq \{T(f) \gtrsim 2^{(st+1)k}\}, \tag{4.39}$$

and in order to see this, we consider two cases: either $g_Q > 2^{k-1}$ or $g_Q \leq 2^{k-1}$. In the first case, the claim follows from (4.38) as we have

$$T(f) \geq |\langle f \rangle_Q|^s g_Q \chi_Q \gtrsim 2^{(st+1)k} \chi_Q.$$

For the second case, we note that by the maximality of Q in $\{G > 2^k\}$, we have

$$\sum_{Q \subseteq R} g_R \chi_R > 2^k \chi_Q, \tag{4.40}$$

and since we assume $g_Q \leq 2^{k-1}$, we must have

$$\sum_{\hat{Q} \subseteq R} g_R \chi_R > 2^{k-1} \chi_{\hat{Q}},$$

So for Q' a maximal cube in $\mathcal{C}_{k-1}$, we have $\hat{Q} \subseteq Q'$, and then the maximality of Q' implies that

$$\sum_{\hat{Q}' \subseteq R} g_R \chi_R \leq 2^{k-1} \chi_{\hat{Q}'}, \quad Q \subsetneq \hat{Q} \subseteq Q' \subsetneq \hat{Q}'. \tag{4.41}$$

Subtracting (4.41) from (4.40), we obtain

$$\sum_{Q \subseteq R \subseteq Q'} g_R \chi_R > 2^{k-1} \chi_Q.$$

Now, since $Q' \in \mathcal{C}_{k-1}$, it follows from (4.38) that we have

$$|\langle f \rangle_R| \gtrsim 2^{t(k-1)}, \quad Q \subseteq R \subseteq Q',$$

which together with the above inequality implies that

$$T(f) \geq \sum_{Q \subseteq R \subseteq Q'} |\langle f \rangle_R|^s g_R \chi_R \gtrsim 2^{ts(k-1)} \sum_{Q \subseteq R \subseteq Q'} g_R \chi_R \gtrsim 2^{k(st+1)} \chi_Q,$$

which proves our claim in (4.39). Now, it follows from (4.39) and disjointness of cubes in $\mathcal{C}_k$ that

$$\sum_{Q \in \mathcal{C}_k} |Q| \leq |\{T(f) \gtrsim 2^{k(st+1)}\}|, \quad k \in \mathbb{Z},$$

and the rest of the proof follows the same line as in the previous case. Inserting the above estimate in the right-hand side of (4.36) gives us

$$\sum_{k \in \mathbb{Z}} 2^{rk} \sum_{Q \in \mathcal{C}_k} |Q| \leq \sum_{k \in \mathbb{Z}} 2^{rk} |\{T(f) \gtrsim 2^{k(st+1)}\}| \lesssim \int_{\mathbb{R}^n} T(f)^q \lesssim A^q \|f\|^{sq}_{H^{sp}_d(\mathbb{R}^n)},$$

and then using (4.37) and (4.36) we obtain

$$\|G\|_{L^r(\mathbb{R}^n)} \lesssim A,$$

which is the desired result when there are only finitely many nonzero terms in $\{g_R\}_{R\in\mathcal{D}}$, and then the limiting argument presented at the end of the previous case extends the result to the general case, and this completes the proof of this case and Theorem 4.3.2. □

We conclude this section by giving an example which shows that for $0 < q < 1$, the operator norm of the formal adjoint of a dyadic paraproduct π_g can be much smaller than the norm of the operator itself. This happens because of the fact that the Hahn-Banach theorem fails for $H^q_d(\mathbb{R}^n)$ [9, 20].

Example. Let g be a dyadic distribution on $\mathbb{R}$. Then the formal adjoint of π_g is given by

$$\pi^t_g(f) := \sum_{I\in\mathcal{D}(\mathbb{R})} \langle f, h_I\rangle \langle g, h_I\rangle \frac{\chi_I}{|I|}.$$

Now let

$$g = \sum_{\substack{I\in\mathcal{D}[0,1]\\ |I|=2^{-l}}} |I|^{\frac{1}{2}} h_I, \quad \frac{1}{q} = \frac{1}{p} + \frac{1}{r}, \quad 0 < q < 1 < p < \infty,$$

and note that $S(g) = \chi_{[0,1]}$, so we have $\|g\|_{\dot{H}^r_d(\mathbb{R})} = 1$, and Theorem 4.3.1 implies that

$$\|\pi_g\|_{H^p_d(\mathbb{R})\to\dot{H}^q_d(\mathbb{R})} \simeq 1.$$

In fact, here we do not need to use Theorem 4.3.1 to conclude this. This can be simply proven by testing the operator π_g on g itself. However, the norm of the adjoint operator $\pi^t_g : \dot{\Lambda}^{\frac{1}{q}-1}(\mathbb{R}) \to (H^p_d(\mathbb{R}))'$ can be estimated as follows: Let $\|f\|_{\dot{\Lambda}^{\frac{1}{q}-1}} = 1$, and note that

$$|\langle f, h_I\rangle| \le \operatorname{osc}_1(f, I)|I|^{\frac{1}{2}} \le |I|^{\frac{1}{q}-\frac{1}{2}},$$

which implies that

$$|\pi_g^t(f)| \leq \sum_{\substack{I\in\mathcal{D}[0,1]\\ |I|=2^{-l}}} |I|^{\frac{1}{q}-1}\chi_I = 2^{l\left(1-\frac{1}{q}\right)}\chi_{[0,1]},$$

and thus

$$\|\pi_g^t(f)\|_{(H_d^p(\mathbb{R}))'} \simeq \|\pi_g^t(f)\|_{L^{p'}(\mathbb{R})} \leq 2^{l\left(1-\frac{1}{q}\right)},$$

which shows that for large l, the operator norm of π_g^t is much smaller than that of π_g.

4.4 Fourier paraproducts

In this section, we show that similar results hold for the operator $\Pi_{g,\varphi}$. Here, we are faced with many error terms and are forced to assume more than merely the boundedness of $\Pi_{g,\varphi}$. To resolve these difficulties, instead of merely assuming that $\Pi_{g,\varphi}$ is bounded from $H^p(\mathbb{R}^n)$ to $\dot{H}^q(\mathbb{R}^n)$, we assume that the sublinear operator

$$\mathcal{S}_{g,\varphi}(f) := \left(\sum_{j\in\mathbb{Z}} |\varphi_{2^{-j}} * f\Delta_j(g)|^2\right)^{\frac{1}{2}}, \tag{4.42}$$

is bounded from $H^p(\mathbb{R}^n)$ to $L^q(\mathbb{R}^n)$. When $\hat{\varphi}$ has compact support and satisfies the restriction mentioned in Theorem (B), this extra assumption is equivalent to assuming that the operators $\Pi_{i,g,\varphi}$, mentioned in (4.24), are all bounded. The reason for this is that

$$\mathcal{S}_{g,\varphi}(f)(x) \simeq \sum_{0\leq i<m} S_\theta(\Pi_{i,g,\varphi}f)(x), \quad x\in\mathbb{R}^n,$$

and thus

$$\|\mathcal{S}_{g,\varphi}\|_{H^p(\mathbb{R}^n)\to L^q(\mathbb{R}^n)} \simeq \sum_{0\leq i<m} \|\Pi_{i,g,\varphi}\|_{H^p(\mathbb{R}^n)\to\dot{H}^q(\mathbb{R}^n)}.$$

Now, we state the main result of this section.

Theorem 4.4.1. *Let ψ be a Schwartz function as in (4.17) and (4.18), and let φ be a Schwartz function whose Fourier transform is supported in a ball with radius $\boldsymbol{a'} < \boldsymbol{a}$, where $\boldsymbol{a}$ is as in (4.17), and equal to 1 in a smaller neighborhood of the origin. Also, let g be a tempered distribution on $\mathbb{R}^n$, and let the sublinear operator $\mathcal{S}_{g,\varphi}$ be as in (4.42). Then we have*

$$\|\mathcal{S}_{g,\varphi}\|_{H^p(\mathbb{R}^n)\to L^{p^*}(\mathbb{R}^n)} \simeq \|g\|_{\dot{\Lambda}^{\alpha}(\mathbb{R}^n)}, \quad \frac{1}{p^*} = \frac{1}{p} - \frac{\alpha}{n}, \quad 0 < \alpha p < n, \quad 0 < p < \infty, \tag{I}$$

$$\|\mathcal{S}_{g,\varphi}\|_{H^p(\mathbb{R}^n)\to L^{p}(\mathbb{R}^n)} \simeq \|g\|_{B\dot{M}O(\mathbb{R}^n)}, \quad 0 < p < \infty, \tag{II}$$

$$\|\mathcal{S}_{g,\varphi}\|_{H^p(\mathbb{R}^n)\to L^{q}(\mathbb{R}^n)} \simeq \|g\|_{\dot{H}^r(\mathbb{R}^n)}, \quad \frac{1}{q} = \frac{1}{p} + \frac{1}{r}, \quad 0 < q, p, r < \infty. \tag{III}$$

In the above, the upper bounds for the operator norm of $\mathcal{S}_{g,\varphi}$ follow directly from Theorem (B) and the above discussion, and it remains only to prove the lower bounds. As we mentioned before, here we are dealing with some error terms that, at the end of the proofs, have to be absorbed into the left-hand side. To this aim, we bring the following simple lemma.

Lemma 4.4.1. *Let ϕ be a Schwartz function, $g \in BMO(\mathbb{R}^n)$, and Q a cube. Then we have*

$$|\phi_t * (g - \langle g \rangle_Q)|(x) \lesssim (1 + |\log tl(Q)^{-1}|)\|g\|_{BMO(\mathbb{R}^n)}, \quad x \in Q, \quad t > l(Q).$$

Proof. Since ϕ is a Schwartz function, we may assume

$$|\phi(y)| \leq \frac{C(\phi)}{(1+|y|)^{n+\delta}}, \quad y \in \mathbb{R}^n, \quad \delta > 0.$$

Now, we have

$$
\begin{aligned}
|\phi_t * (g - \langle g \rangle_Q)|(x) &\lesssim \int_{\mathbb{R}^n} \frac{t^{-n}}{(1+|\frac{y-x}{t}|)^{n+\delta}} |g(y) - \langle g \rangle_Q| dy \\
&\lesssim \int_Q t^{-n} |g(y) - \langle g \rangle_Q| dy + \sum_{k\geq 0} \int_{2^{k+1}Q \setminus 2^k Q} \frac{t^{-n}}{(1+t^{-1}l(Q)2^k)^{n+\delta}} |g(y) - \langle g \rangle_Q| dy,
\end{aligned}
$$

where in the above we used the fact that for $x \in Q$ and $y \in 2^{k+1}Q \setminus 2^k Q$ we have

$$
|x - y| \simeq 2^k l(Q).
$$

Next, we note that

$$
\int_{2^{k+1}Q} |g(y) - \langle g \rangle_Q| \lesssim \|g\|_{BMO(\mathbb{R}^n)} (k+1) |2^{k+1}Q|, \quad k \geq 0,
$$

which, after plugging in the above, gives us

$$
|\phi_t * (g - \langle g \rangle_Q)|(x) \lesssim \|g\|_{BMO(\mathbb{R}^n)} (t^{-1} l(Q))^n \left(1 + \sum_{k\geq 0} \frac{2^{kn} k}{(1+t^{-1}l(Q)2^k)^{n+\delta}} \right). \quad (4.43)
$$

Then, we estimate the last sum as

$$
\sum_{k\geq 0} \frac{2^{kn} k}{(1+t^{-1}l(Q)2^k)^{n+\delta}} \leq \sum_{t^{-1}l(Q)2^k \leq 1} 2^{kn} k + \sum_{t^{-1}l(Q)2^k > 1} \frac{2^{-k\delta} k}{(t^{-1}l(Q))^{n+\delta}}.
$$

Noting that because of the geometric factor, each sum in the above is dominated by its largest term, we obtain

$$
\sum_{k\geq 0} \frac{2^{kn} k}{(1+t^{-1}l(Q)2^k)^{n+\delta}} \lesssim 1 + (t l(Q)^{-1})^n |\log t l(Q)^{-1}|,
$$

which, together with (4.43) and our assumption that $t > l(Q)$, implies

$$|\phi_t * (g - \langle g \rangle_Q)|(x) \lesssim \|g\|_{BMO(\mathbb{R}^n)}(1 + |\log tl(Q)^{-1}|),$$

and this completes the proof. □

We break the proof of Theorem 4.4.1 into two parts, and first we prove (I) and (II).

Proofs of (I) *and* (II). Let $A = \|\mathcal{S}_{g,\varphi}\|_{H^p(\mathbb{R}^n)\to L^{p^*}(\mathbb{R}^n)}$, and let c be such that

$$\hat{\varphi}(\xi) = 1, \quad |\xi| \le c.$$

We take a Schwartz function f with

$$\text{supp}(\hat{f}) \subseteq B_c \setminus B_{\frac{c}{2}}, \quad f(0) = \int \hat{f} = 1,$$

and for $j \in \mathbb{Z}$ and $x_0 \in \mathbb{R}^n$, consider the function $f_{j,x_0} = \tau^{x_0}\delta^{2^{-j}} f$, whose Fourier transform is supported in $B_{c2^j} \setminus B_{c2^{j-1}}$ and $f_{j,x_0}(x_0) = 1$. Now, since $\hat{\varphi}$ is equal to 1 on B_c, we have $\varphi_{2^{-j}} * f_{j,x_0} = f_{j,x_0}$. Therefore, from the boundedness of $\mathcal{S}_{g,\varphi}$, it follows that

$$\|\varphi_{2^{-j}} * f_{j,x_0}\Delta_j g\|_{L^{p^*}(\mathbb{R}^n)} \le A\|f_{j,x_0}\|_{H^p(\mathbb{R}^n)}.$$

Then, since $\|f_{j,x_0}\|_{H^p(\mathbb{R}^n)} \simeq 2^{-j\frac{n}{p}}$, and $\varphi_{2^{-j}} * f_{j,x_0} = f_{j,x_0}$, we must have

$$\|f_{j,x_0}\Delta_j g\|_{L^{p^*}(\mathbb{R}^n)} \lesssim A2^{-j\frac{n}{p}}.$$

At this point, we note that the Fourier transform of the function $f_{j,x_0}\Delta_j g$ is supported in a ball of radius $\simeq 2^j$. Therefore, using the Plancherel-Polya-Nikolskij inequality

(4.19), we obtain

$$\|f_{j,x_0}\Delta_j g\|_{L^\infty(\mathbb{R}^n)} \lesssim 2^{j\frac{n}{p^*}}\|f_{j,x_0}\Delta_j g\|_{L^{p^*}(\mathbb{R}^n)} \lesssim A2^{jn(\frac{1}{p^*}-\frac{1}{p})} = A2^{-j\alpha},$$

and since $f_{j,x_0}(x_0) = 1$, we must have

$$|\Delta_j g(x_0)| \lesssim A2^{-j\alpha}, \quad x_0 \in \mathbb{R}^n, \quad j \in \mathbb{Z},$$

which shows that

$$\|g\|_{\dot{\Lambda}^\alpha(\mathbb{R}^n)} \simeq \sup_{j\in\mathbb{Z}} 2^{j\alpha}\|\Delta_j g\|_{L^\infty(\mathbb{R}^n)} \lesssim A.$$

This proves (I).

Now we proceed to the proof of (II). To this end, let $A = \|\mathcal{S}_{g,\varphi}\|_{H^p(\mathbb{R}^n)\to L^p(\mathbb{R}^n)}$, and choose a large integer m such that the Fourier transforms of any two consecutive terms in

$$\sum_{j\in m\mathbb{Z}+i} \varphi_{2^{-j}} * f\Delta_j(g), \quad 0 \le i < m,$$

are at sufficiently large distances from each other. Also, for a fixed natural number N, let

$$g_{i,N} = \sum_{\substack{j\in m\mathbb{Z}+i \\ |j|\le N}} \Delta_j g, \quad P_{i,N}(f) = \sum_{\substack{j\in m\mathbb{Z}+i \\ |j|\le N}} \varphi_{2^{-j}} * f\Delta_j g, \quad 0 \le i < m. \tag{4.44}$$

From now on, we fix i, and for simplicity of notation, set $g' = g_{i,N}$ and $P = P_{i,N}$. We note that for a suitable choice of θ, we have

$$\Delta_l^\theta(g') = \delta_{j,l}\Delta_j g, \quad j \in m\mathbb{Z}+i, \quad l \in \mathbb{Z},$$

and thus we may replace $\Delta_j^{\psi} g$ with $\Delta_j^{\theta}(g')$. Therefore, we have

$$g' = \sum_{j\in\mathbb{Z}} \Delta_j^{\theta}(g'), \quad P(f) = \sum_{j\in\mathbb{Z}} \varphi_{2^{-j}} * f\Delta_j^{\theta}(g'), \tag{4.45}$$

and since there are only finitely many nonzero terms in the above expression, we have

$$\|P\|_{H^p(\mathbb{R}^n)\to H^p(\mathbb{R}^n)} \simeq \|P\|_{H^p(\mathbb{R}^n)\to \dot{H}^p(\mathbb{R}^n)} \lesssim A.$$

We now turn to estimating the mean oscillation of g' over a cube. Before doing so, we must ensure that g' belongs to $BMO(\mathbb{R}^n)$. To see this, we note that the previous argument for (I) still holds with $\alpha = 0$. Thus, $\|\Delta_j g\|_{L^\infty(\mathbb{R}^n)} \lesssim A$, and since g' is a finite sum of such terms, it is bounded as well and hence belongs to $BMO(\mathbb{R}^n)$.

So, let Q be a cube with $2^{-k} \leq l(Q) < 2^{-k+1}$, and x_0 be its center. Also, let k_0 be a large number to be determined later, and take the function $h = f_{k-k_0,x_0}$, where f is as in the previous case. Note that $\hat{h}$ is supported on $B_{c2^{k-k_0}} \setminus B_{c2^{k-k_0-1}}$. Then, since the Fourier transform of $\varphi_{2^{-j}}$ is equal to 1 on B_{c2^j} and vanishes outside of $B_{\boldsymbol{a}'2^j}$, we conclude that for a sufficiently large choice of m depending only on c and $\boldsymbol{a}'$, we have

$$\varphi_{2^{-j}} * h = h, \quad j \geq k-k_0, \quad \varphi_{2^{-j}} * h = 0, \quad j < k-k_0, \quad j \in m\mathbb{Z}+i.$$

This observation implies that

$$P(h) = h \sum_{k-k_0\leq j} \Delta_j^{\theta}(g') = h(g' - \sum_{j\leq k-k_0-1} \Delta_j^{\theta}(g')) = h(g' - S_{k-k_0-1}^{\theta}(g')).$$

Now, let Θ be the kernel function of the above partial sum operator. Then, since

$\int \Theta = 1$, we may write

$$g' - S^{\theta}_{k-k_0-1}(g') = g' - \langle g' \rangle_Q - \Theta_{2^{-k+k_0+1}} * (g' - \langle g' \rangle_Q),$$

and thus we obtain

$$P(h) = h\left(g' - \langle g' \rangle_Q - \Theta_{2^{-k+k_0+1}} * (g' - \langle g' \rangle_Q)\right).$$

Also, since $h(x_0) = 1$, we can decompose the above sum as

$$P(h) = g' - \langle g' \rangle_Q - \Theta_{2^{-k+k_0+1}} * (g' - \langle g' \rangle_Q) + (h - h(x_0))\left(g' - \langle g' \rangle_Q - \Theta_{2^{-k+k_0+1}} * (g' - \langle g' \rangle_Q)\right),$$

which is equivalent to

$$\begin{aligned} g' - \langle g' \rangle_Q &= P(h) + \Theta_{2^{-k+k_0+1}} * (g' - \langle g' \rangle_Q) - (h - h(x_0))(g' - \langle g' \rangle_Q) \\ &+ (h - h(x_0))\Theta_{2^{-k+k_0+1}} * (g' - \langle g' \rangle_Q) = P(h) + E_1 - E_2 + E_3. \end{aligned}$$

This implies that

$$\operatorname{osc}_p(g', Q) \lesssim \operatorname{osc}_p(P(h), Q) + \sum_{i=1}^{3} \operatorname{osc}_p(E_i, Q), \tag{4.46}$$

where

$$E_1 = \Theta_{2^{-k+k_0+1}} * (g' - \langle g' \rangle_Q), \quad E_2 = (h - h(x_0))(g' - \langle g' \rangle_Q), \quad E_3 = (h - h(x_0))\Theta_{2^{-k+k_0+1}} * (g' - \langle g' \rangle_Q).$$

Now we estimate the p-mean oscillation of each term. For the first term, we have

$$\operatorname{osc}_p(P(h), Q) \lesssim |Q|^{-\frac{1}{p}} \|P(h)\|_{L^p(\mathbb{R}^n)} \lesssim 2^{k\frac{n}{p}} \|P(h)\|_{H^p(\mathbb{R}^n)} \lesssim A 2^{k\frac{n}{p}} \|h\|_{H^p(\mathbb{R}^n)}.$$

Then, since $\|h\|_{H^p(\mathbb{R}^n)} \simeq 2^{\frac{n(k_0-k)}{p}}$, which follows from the fact that $h = \tau^{x_0}\delta^{2^{k_0-k}}(f)$, the above inequality implies

$$\mathrm{osc}_p(P(h), Q) \lesssim 2^{\frac{nk_0}{p}} A. \tag{4.47}$$

For the second term, we have

$$\mathrm{osc}_p(E_1, Q) \lesssim l(Q) \sup_{x\in Q} \left| \nabla\Theta_{2^{-k+k_0+1}} * (g' - \langle g'\rangle_Q)(x) \right|,$$

where we used the mean value theorem. Since for any $t > 0$, $\nabla\Theta_t = t^{-1}(\nabla\Theta)_t$, we may write

$$\mathrm{osc}_p(E_1, Q) \lesssim 2^{k-k_0-1} l(Q) \sup_{x\in Q} \left| (\nabla\Theta)_{2^{-k+k_0+1}} * (g' - \langle g'\rangle_Q)(x) \right|, \tag{4.48}$$

and by applying Lemma 4.4.1 to $\nabla\Theta$, we get

$$\sup_{x\in Q} \left| (\nabla\Theta)_{2^{-k+k_0+1}} * (g' - \langle g'\rangle_Q)(x) \right| \lesssim \left(1 + \log 2^{-k+k_0+1} l(Q)^{-1}\right) \|g'\|_{BMO(\mathbb{R}^n)},$$

which, noting that $l(Q) \simeq 2^{-k}$, together with (4.48), gives

$$\mathrm{osc}_p(E_1, Q) \lesssim 2^{-k_0} k_0 \|g'\|_{BMO(\mathbb{R}^n)}. \tag{4.49}$$

To estimate the third term, we note that

$$\mathrm{osc}_p(E_2, Q) \lesssim \langle |E_2|^p \rangle_Q^{\frac{1}{p}} \leq \sup_{x\in Q} |h(x) - h(x_0)| \mathrm{osc}_p(g', Q) \lesssim l(Q) \sup_{x\in Q} |\nabla h(x)| \mathrm{osc}_p(g', Q),$$

and since $\|\nabla h\|_{L^\infty} \simeq 2^{k-k_0}$, we obtain

$$\mathrm{osc}_p(E_2, Q) \lesssim 2^{-k_0} \mathrm{osc}_p(g', Q) \lesssim 2^{-k_0} \|g'\|_{BMO(\mathbb{R}^n)}. \tag{4.50}$$

Finally, for the last term, we have

$$\operatorname{osc}_p(E_3, Q) \leq 2 \sup_{x\in Q} |h(x) - h(x_0)| \sup_{y\in Q} \left| \Theta_{2^{-k+k_0+1}} * (g' - \langle g' \rangle_Q)(y) \right|,$$

which, using the mean value theorem for h and applying Lemma 4.4.1, implies

$$\operatorname{osc}_p(E_3, Q) \lesssim l(Q) 2^{k-k_0} \left(1 + \log 2^{-k+k_0+1} l(Q)^{-1}\right) \|g'\|_{BMO(\mathbb{R}^n)} \lesssim 2^{-k_0} k_0 \|g'\|_{BMO(\mathbb{R}^n)}. \tag{4.51}$$

Putting (4.46), (4.47), (4.49), (4.50), and (4.51) together, we obtain

$$\operatorname{osc}_p(g', Q) \lesssim 2^{\frac{nk_0}{p}} A + k_0 2^{-k_0} \|g'\|_{BMO(\mathbb{R}^n)},$$

and taking the supremum over all cubes gives

$$\|g'\|_{BMO(\mathbb{R}^n)} \lesssim 2^{\frac{nk_0}{p}} A + k_0 2^{-k_0} \|g'\|_{BMO(\mathbb{R}^n)}.$$

Now, by choosing k_0 large enough and noting that we already know g' belongs to $BMO(\mathbb{R}^n)$, we conclude that $\|g'\|_{BMO(\mathbb{R}^n)} \lesssim A$. Finally, recall that

$$g' = g_{i,N} = \sum_{\substack{j\in m\mathbb{Z}+i \\ |j|\leq N}} \Delta_j g,$$

and we have

$$g_N = \sum_{|j|\leq N} \Delta_j g = \sum_{0\leq i<m} g_{i,N}.$$

Thus, this sequence must be bounded in $BMO(\mathbb{R}^n)$, and the Banach-Alaoglu theorem implies that there exists a subsequence converging in the weak* topology of $BMO(\mathbb{R}^n)$ to a function G with $\|G\|_{BMO(\mathbb{R})} \lesssim A$. Since the sequence g_N converges in the space of distributions modulo polynomials to g, we conclude that g is equal to G modulo a polynomial, which means that there exists a polynomial U such that

$$\|g - U\|_{BMO(\mathbb{R}^n)} \lesssim A,$$

and this proves (II). □

Remark 4.4.1. *We note that the boundedness of the operator $\mathcal{S}_{g,\varphi}$ on $L^2(\mathbb{R}^n)$ is equivalent to the statement that the measure*

$$d\mu(x,t) = \sum_{j\in\mathbb{Z}} |\Delta_j g|^2 \, dx \, d\delta_{2^{-j}}(t),$$

is a Carleson measure. For $p = 2$, our result can be rephrased as: if the above measure is Carleson, then g, belongs to $B\dot{M}O(\mathbb{R}^n)$. As far as we know, the previous proof of this fact uses the assumption on g, and directly shows that modulo a polynomial g lies in the dual of $H^1(\mathbb{R}^n)$. Fefferman's duality theorem then implies that $g \in B\dot{M}O(\mathbb{R}^n)$ [66] (p. 161). However, here we only used the fact that, to estimate the $BMO(\mathbb{R}^n)$ norm, we may use any p-mean oscillation, which follows from the John-Nirenberg lemma, and the fact that every bounded sequence in $BMO(\mathbb{R}^n)$ has a subsequence converging in the topology of distributions modulo polynomials.

We now proceed to the proof of (III), and in order to do so, we need a series of lemmas.

Lemma 4.4.2. *Let φ be a Schwartz function with $\int \varphi = 1$, and let B_1 be the unit ball in $\mathbb{R}^n$. Then, for $\alpha \geq 2$ and $0 < p < \infty$, there exists a function $\tilde{\chi}$ supported in a ball with radius $c(\alpha, p, \varphi)$ such that*

$$|\varphi_t * \tilde{\chi}| > \frac{1}{3}\chi_{B_1}, \quad t \leq \alpha, \tag{4.52}$$

$$\|\tilde{\chi}\|_{H^p(\mathbb{R}^n)} \leq c'(\varphi, \alpha, p). \tag{4.53}$$

Proof. First, note that if φ is complex valued then the integral of its real part is 1, and therefore without loss of generality we may assume that φ is real valued. Then, we construct an atom with a large amount of cancellation that satisfies the required conditions. To this aim, pick a large number M such that

$$\int_{|x|\geq M} |\varphi| \leq \frac{1}{3}, \quad M > \frac{\alpha}{2},$$

which implies that for $t \leq \alpha$, we have

$$\int_{|x|\geq \alpha M} |\varphi_t|(x)\,dx = \int_{|x|\geq \frac{\alpha}{t}M} |\varphi|(x)\,dx \leq \frac{1}{3}.$$

Also, for $x \in B_1$, we have

$$\varphi_t * \chi_{B_{\alpha M}}(x) = \int_{\mathbb{R}^n} \varphi_t(x-y)\chi_{B_{\alpha M}}(y)\,dy = \int_{B_{\alpha M}(x)} \varphi_t(z)\,dz = 1 - \int_{B^c_{\alpha M}(x)} \varphi_t(z)\,dz,$$

and since $B^c_{\alpha M}(x) \subseteq B^c_M(0)$ (because $\alpha \geq 2$), we have

$$\varphi_t * \chi_{B_{\alpha M}}(x) \geq \frac{2}{3}, \quad x \in B_1. \tag{4.54}$$

Now, take a natural number $N > n(\frac{1}{p}-1)$, and let B' be a ball of radius 1 at distance D from the origin, say $B' = B_1(2De)$ for some unit vector e. Then, we choose P, a polynomial of degree at most N, and set

$$\tilde{\chi}(x) = \chi_{B_{\alpha M}}(x) + P(x)\chi_{B'}(x).$$

To make $\tilde{\chi}$ into an atom, we need to find P such that

$$\int \tilde{\chi}(x)Q(x)\,dx = 0, \quad Q \in \mathbb{P}_N, \tag{4.55}$$

where $\mathbb{P}_N$ is the space of real valued polynomials in n variables with degree no more than N. Additionally, P has to be chosen such that $P\chi_{B'}$ has a small contribution to $\varphi_t * \tilde{\chi}$ on the unit ball, meaning that

$$|\varphi_t * (P\chi_{B'})(x)| \leq \frac{1}{3}, \quad x \in B_1, \quad t \leq \alpha. \tag{4.56}$$

To achieve the first task, consider the inner product on $\mathbb{P}_N$, defined as

$$\langle Q_1, Q_2 \rangle := \int_{B_1} Q_1(x) Q_2(x)\, dx,$$

and pick an orthonormal basis $\{Q_\beta\}_{|\beta| \leq N}$ with respect to this inner product. Then, translate these polynomials to the center of B', and set

$$Q'_\beta = \tau^{2De} Q_\beta, \quad |\beta| \leq N, \tag{4.57}$$

which gives an orthonormal basis for $\mathbb{P}_N$, equipped with the new inner product

$$\langle Q_1, Q_2 \rangle' := \int_{B'} Q_1(x) Q_2(x)\, dx.$$

Now, the cancellation condition (4.55) becomes

$$\left\langle P, Q'_\beta \right\rangle' = -\int_{B_{\alpha M}} Q'_\beta, \quad |\beta| \leq N,$$

which has a unique solution P defined as

$$P(x) = \sum_{|\beta| \leq N} \left\langle P, Q'_\beta \right\rangle' Q'_\beta(x) = \sum_{|\beta| \leq N} c_\beta x^\beta.$$

Then, to see why (4.56) holds note that (4.57) implies that for each β, the coefficients of Q'_β, grow no more than a constant (which depends only on n, and N) times D^N, and therefore

$$|\langle P, Q'_\beta \rangle'| \leq \int_{B_{\alpha M}} |Q'_\beta| \leq C(n, N, \alpha, M) D^N,$$

which implies that

$$|c_\beta| \leq C'(n, N, , \alpha, M) D^{2N}, \quad D > 1, \quad |\beta| \leq N.$$

On the other hand, the Schwartz function φ decays faster than any polynomial, so for a constant $C(\varphi, N)$, we have

$$|\varphi(x)| \leq C(\varphi, N) \frac{1}{(1 + |x|)^{3N+n}}, \quad x \in \mathbb{R}^n.$$

Now, for $x \in B_1$ we write

$$|\varphi_t * (P\chi_{B'})(x)| \leq \int_{B'} |\varphi_t(x - y)||P(y)| \, dy,$$

and we note that since $x \in B_1$ and $y \in B'$, we have $|x - y| \simeq D$. Also, since P is a polynomial of degree at most N, we have

$$|P(y)| \leq C(n, N, \alpha, M) D^{3N}, \quad y \in B',$$

which implies that

$$|\varphi_t * (P\chi_{B'})(x)| \leq \int_{B'} \varphi_t(x - y)|P(y)| \, dy \leq C(n, \varphi, N, \alpha, M) D^{3N} t^{-n} \int_{B'} \frac{1}{|t^{-1}D|^{3N+n}}.$$

Then, it is enough to recall that $|B'| = 1$, and $t \leq \alpha$ to obtain

$$|\varphi_t * (P\chi_{B'})(x)| \leq C(n, \varphi, N, \alpha, M) D^{-n}, \quad x \in B_1.$$

Now, we choose D to be sufficiently large so that the right-hand side of the above inequality is less than $\frac{1}{3}$. This shows that we can find P such that (4.56) holds, which, together with (4.54), implies that $\tilde{\chi}$ satisfies (4.52). Finally, since $\tilde{\chi}$ satisfies (4.55) and is a bounded function supported in a ball with a radius depending only on φ, p, α, and n, we conclude

$$\|\tilde{\chi}\|_{H^p(\mathbb{R}^n)} \leq C(\varphi, p, \alpha, n),$$

which completes the proof. □

We continue by proving the following lemma, which is designed to verify some a priori bounds.

Lemma 4.4.3. *Let $0 < p, q, r < \infty$ with $\frac{1}{q} = \frac{1}{p} + \frac{1}{r}$, φ a Schwartz function with $\int \varphi = 1$, and u, a smooth function with compact Fourier support. Also, assume*

$$\|(\varphi * f)u\|_{L^q(\mathbb{R}^n)} \leq A\|f\|_{H^p(\mathbb{R}^n)},$$

holds for compactly supported functions f. Then $\|u\|_{L^r(\mathbb{R}^n)} < \infty$.

Proof. After a rescaling, we may assume that $\operatorname{supp}(\hat{u}) \subseteq B_1$. Our strategy is to use the second Plancherel-Polya-Nikolskij inequality (4.20) and show that for a sufficiently small choice of h, there exists a sequence $\{x_k\}_{k\in\mathbb{Z}^n}$ with

$$\|\{u(x_k)\}\|_{l^r(\mathbb{Z}^n)} < \infty, \quad x_k \in h(k + [0,1]^n), \quad k \in \mathbb{Z}^n, \tag{4.58}$$

which implies the claim once we apply the Plancherel-Polya-Nikolskij inequality (4.20). In order to do this, fix h and partition $\mathbb{R}^n$ into cubes of the form

$$Q_k = h(k + [0,1]^n), \quad k \in \mathbb{Z}^n,$$

then let $\tilde{\chi}$ be the function provided by Lemma 4.4.2 with φ, p, and $\alpha = 2$. Now, for each cube Q_k, consider the function $f_k = \tau^{hk}\delta^{2\sqrt{n}}\tilde{\chi}$, which, for $h \leq 1$, has the property that

$$|\varphi * f_k| \gtrsim \chi_{B_{2\sqrt{n}}(hk)} \geq \chi_{Q_k}, \quad \|f_k\|_{H^p(\mathbb{R}^n)} \lesssim 1, \quad |f_k| \lesssim \chi_{B_c(hk)},$$

where c is a large number provided by the above lemma. Next, choose an arbitrary collection of numbers $\{a_k\}_{k\in\mathbb{Z}^n}$ such that only finitely many of them are nonzero, and let $\{\epsilon_k = \pm 1\}_{k\in\mathbb{Z}^n}$ be a family of independent random variables with Bernoulli distribution. Then, for the function f defined as

$$f = \sum_{k\in\mathbb{Z}^n} \epsilon_k a_k f_k,$$

we have

$$\|(\varphi * f)u\|^q_{L^q(\mathbb{R}^n)} = \int \left| \sum_{k\in\mathbb{Z}^n} \epsilon_k a_k \varphi * f_k \right|^q |u|^q \leq A^q \|f\|^q_{H^p(\mathbb{R}^n)}. \tag{4.59}$$

Now, for $0 < p \leq 1$, let us use sub-additivity and estimate the right-hand side as

$$\|f\|_{H^p(\mathbb{R}^n)} \leq \left(\sum_{k\in\mathbb{Z}^n} |a_k|^p \|f_k\|^p_{H^p(\mathbb{R}^n)} \right)^{\frac{1}{p}} \lesssim \|\{a_k\}_{k\in\mathbb{Z}^n}\|_{l^p(\mathbb{Z}^n)},$$

and for $1 < p < \infty$ we estimate as

$$\|f\|_{H^p(\mathbb{R}^n)} \lesssim \| \sum_{k\in\mathbb{Z}^n} |a_k| \chi_{B_c(hk)} \|_{L^p(\mathbb{R}^n)} \lesssim_h \|\{a_k\}_{k\in\mathbb{Z}^n}\|_{l^p(\mathbb{Z}^n)}.$$

Therefore, from (4.59) we must have

$$\int \left| \sum_{k\in\mathbb{Z}^n} \epsilon_k a_k \varphi * f_k \right|^q |u|^q \lesssim_h A^q \|\{a_k\}_{k\in\mathbb{Z}^n}\|^q_{l^p(\mathbb{Z}^n)}.$$

Now, taking the expectation we obtain

$$\int \mathbb{E} \left| \sum_{k\in\mathbb{Z}^n} \epsilon_k a_k \varphi * f_k \right|^q |u|^q = \mathbb{E} \int \left| \sum_{k\in\mathbb{Z}^n} \epsilon_k a_k \varphi * f_k \right|^q |u|^q \lesssim_h A^q \|\{a_k\}_{k\in\mathbb{Z}^n}\|^q_{l^p(\mathbb{Z}^n)},$$

and then applying Khintchine inequality gives us

$$\int \left(\sum_{k \in \mathbb{Z}^n} |a_k \varphi * f_k|^2 \right)^{\frac{q}{2}} |u|^q \lesssim_h A^q \|\{a_k\}_{k\in\mathbb{Z}^n}\|^q_{l^p(\mathbb{Z}^n)}.$$

Now, we note that $|\varphi * f_k| \gtrsim \chi_{Q_k}$, and the cubes Q_k are disjoint, which implies that

$$\left(\sum_{k \in \mathbb{Z}^n} |a_k|^q \int_{Q_k} |u|^q \right)^{\frac{1}{q}} \lesssim_h A\|\{a_k\}_{k\in\mathbb{Z}^n}\|_{l^p(\mathbb{Z}^n)}.$$

Then, we choose a natural number N, and set $\{a_k\}_{k\in\mathbb{Z}^n}$ to be

$$a_k^p = a_k^q \left\langle |u|^q \right\rangle_{Q_k}, \quad |k| \leq N, \quad \text{otherwise} a_k = 0,$$

which after plugging into the above inequality, and letting N goes to infinity implies that

$$\left(\sum_{k \in \mathbb{Z}^n} \left(\left\langle |u|^q \right\rangle_{Q_k} \right)^{\frac{r}{q}} \right)^{\frac{1}{r}} \lesssim_h A.$$

Finally, since u is continuous for any cube Q_k, there exists a choice of x_k such that

$$|u(x_k)| = \left\langle |u|^q \right\rangle_{Q_k}^{\frac{1}{q}}, \quad x_k \in \mathbb{Z}^n,$$

which shows that (4.58) holds, and this completes the proof. □

In the next lemma, we find a sparse domination of the square function of the symbol of the operator, and as in the dyadic case, it is more convenient to introduce a notation. For a dyadic cube R and α, we set

$$S_\alpha(g|R) := \left(\sum_{2^{-j} \leq \alpha l(R)} |\Delta_j^\psi(g)|^2 \right)^{\frac{1}{2}} \chi_R.$$

Also, we use the convention that $2^{-\infty} = 0$.

Lemma 4.4.4. *Let g be a tempered distribution, $s > 0$, and $\alpha > \sqrt{n}$. Then, for a dyadic cube Q_0 and $0 < \eta < 1$, there exists $\mathcal{C}$, an η-sparse family of cubes in Q_0, with the following properties:*

For any $Q \in \mathcal{C}$, there exists $\lambda_Q \in \mathbb{Z} \cup \{-\infty\}$ such that

$$S_\alpha(g|Q_0)(x) \lesssim \sum_{Q \in \mathcal{C}} 2^{\lambda_Q} \chi_Q(x) + \alpha^{-1} \sum_{Q \in \mathcal{C}} \left\langle M^*_{\nabla\psi}(g)^s \right\rangle_Q^{\frac{1}{s}} \chi_Q(x), \tag{4.60}$$

$$|Q| \lesssim |\{S_\alpha(g|Q) \geq 2^{\lambda_Q - 1}\}|. \tag{4.61}$$

Proof. Using the notation of Theorem 4.2.1, for any dyadic cube Q and $P \subseteq Q$, let

$$f_Q = S_\alpha(g|Q), \quad f_{P,Q} = \left(\sum_{\alpha l(P) < 2^{-j} \leq \alpha l(Q)} |\Delta_j^\psi(g)|^2 \chi_Q \right)^{\frac{1}{2}}.$$

First, we note that, by Minkowski's inequality for the l^2 norm, the condition $|f_{P,Q}| \leq |f_P| + |f_Q|$ holds. Then, it follows from Theorem 4.2.1 that there exists an η-sparse family of dyadic cubes $\mathcal{C}$ such that

$$f_Q(x) \lesssim \sum_{Q \in \mathcal{C}} \gamma_Q \chi_Q(x),$$

where

$$\gamma_Q = (f_Q \chi_Q)^*(\eta'|Q|) + (m_Q^{\#} f \chi_Q)^*(\eta'|Q|), \quad \eta' = \frac{1-\eta}{2^{n+2}}. \tag{4.62}$$

So, it is enough to choose numbers λ_Q, such that for all $Q \in \mathcal{C}$ (4.61) holds, and in addition

$$(f_Q \chi_Q)^*(\eta'|Q|) \lesssim 2^{\lambda_Q}, \quad (m_Q^{\#} f \chi_Q)^*(\eta'|Q|) \lesssim \alpha^{-1} \left\langle M^*_{\nabla\psi}(g)(x)^s \right\rangle_Q^{\frac{1}{s}},$$

and this proves the claim. First of all, if $(f_Q \chi_Q)^*(\eta'|Q|) = 0$, we set $\lambda_Q = -\infty$, and then (4.61) trivially holds. Now suppose $(f_Q\chi_Q)^*(\eta'|Q|) \neq 0$, and let λ_Q be the integer such that

$$2^{\lambda_Q-1} < (f_Q\chi_Q)^*(\eta'|Q|) \leq 2^{\lambda_Q},$$

then from the definition of non-increasing rearrangement we have

$$\eta'|Q| \leq |\{f_Q\chi_Q > 2^{\lambda_Q-1}\}|,$$

which shows that (4.61), holds for Q as well. Also, in both case we have

$$(f_Q\chi_Q)^*(\eta'|Q|) \leq 2^{\lambda_Q}.$$

So, it remains to estimate $(m_Q^{\#}f\chi_Q)^*(\eta'|Q|)$. To this aim, recall that

$$m_Q^{\#}f(x) = \sup_{\substack{x\in P\\ P\subseteq Q}} \operatorname{osc}(f_{P,Q}, P), \quad x \in Q,$$

and fix $x \in Q$ and $P \subseteq Q$ with $x \in P$. Then for any $y, z \in P$, by the triangle inequality for the l^2 norm, we have

$$|f_{P,Q}(y) - f_{P,Q}(z)| \leq \left(\sum_{\alpha l(P)<2^{-j}\leq \alpha l(Q)} |\Delta_j^{\psi}(g)(y) - \Delta_j^{\psi}(g)(z)|^2 \right)^{\frac{1}{2}}. \tag{4.63}$$

Now, for a fixed j, applying the mean value theorem gives us

$$|\Delta_j^{\psi}(g)(y) - \Delta_j^{\psi}(g)(z)| \lesssim l(P) \sup_{w\in Q} |\nabla\Delta_j(g)(w)|,$$

which, by recalling that

$$\Delta_j g(w) = \psi_{2^{-j}} * g(w), \quad \nabla\psi_{2^{-j}} = 2^j(\nabla\psi)_{2^{-j}},$$

is equivalent to

$$|\Delta_j^{\psi}(g)(y) - \Delta_j^{\psi}(g)(z)| \lesssim 2^j l(P) \sup_{w \in Q} |(\nabla\psi)_{2^{-j}}(g)(w)|.$$

Also, since $|w - x| \leq \sqrt{n} l(P) \leq \sqrt{n}\alpha^{-1}2^{-j}$, it follows from the definition of the non-tangential maximal function that for $\alpha > \sqrt{n}$, we have

$$|\Delta_j^{\psi}(g)(y) - \Delta_j^{\psi}(g)(z)| \lesssim 2^j l(P) M^*_{\nabla\psi}(g)(x).$$

Now, we estimate (4.63) as

$$|f_{P,Q}(y) - f_{P,Q}(z)| \lesssim M^*_{\nabla\psi}(g)(x) l(P) \sum_{\alpha l(P) < 2^{-j}} 2^j \lesssim \alpha^{-1} M^*_{\nabla\psi}(g)(x),$$

which proves that

$$\operatorname{osc}(f_{P,Q}, P) \lesssim \alpha^{-1} M^*_{\nabla\psi}(g)(x),$$

and after taking the supremum over $P \subseteq Q$ gives us

$$m_Q^{\#} f(x) \lesssim \alpha^{-1} M^*_{\nabla\psi}(g)(x).$$

Next, we note that for any function h and $0 < \lambda < 1$, Chebyshev's inequality implies that

$$h^*(\lambda|Q|) \leq \lambda^{-\frac{1}{s}} \langle |h|^s \rangle_Q^{\frac{1}{s}},$$

which, together with the above estimate on $m_Q^{\#} f$, gives us

$$(m_Q^{\#} f \chi_Q)^*(\eta'|Q|) \lesssim \alpha^{-1} \left\langle |M^*_{\nabla\psi}(g)|^s \right\rangle_Q^{\frac{1}{s}},$$

and this completes the proof.

□

The last lemma that we need is the following well-known result, which we prove it here.

Lemma 4.4.5. *For $0 < s < r < \infty$, a function h, and an η-sparse collection of cubes $\mathcal{C}$, we have*

$$\left\| \sum_{Q \in \mathcal{C}} \left(\langle |h|^s \rangle_Q \right)^{\frac{1}{s}} \chi_Q \right\|_{L^r(\mathbb{R}^n)} \lesssim \|h\|_{L^r(\mathbb{R}^n)}.$$

Proof. Let M be the cubic Hardy-Littlewood maximal operator. For $Q \in \mathcal{C}$, let E_Q be the disjoint parts as in the definition of sparse families. Then we have

$$\langle |h|^s \rangle_Q^{\frac{1}{s}} \leq M(|h|^s)^{\frac{1}{s}}(x), \quad x \in Q,$$

which, after taking the r-average over E_Q, gives us

$$\langle |h|^s \rangle_Q^{\frac{1}{s}} \leq \left\langle M(|h|^s)^{\frac{r}{s}} \right\rangle_{E_Q}^{\frac{1}{r}}.$$

Now we have

$$\left\| \sum_{Q \in \mathcal{C}} \langle |h|^s \rangle_Q^{\frac{1}{s}} \chi_Q \right\|_{L^r(\mathbb{R}^n)} \leq \left\| \sum_{Q \in \mathcal{C}} \left\langle M(|h|^s)^{\frac{r}{s}} \right\rangle_{E_Q}^{\frac{1}{r}} \chi_Q \right\|_{L^r(\mathbb{R}^n)},$$

which, after applying Lemma 4.2.1 and using the disjointness of the sets E_Q, implies

$$\left\| \sum_{Q \in \mathcal{C}} \langle |h|^s \rangle_Q^{\frac{1}{s}} \chi_Q \right\|_{L^r(\mathbb{R}^n)} \lesssim \left(\int M(|h|^s)^{\frac{r}{s}} \right)^{\frac{1}{r}}.$$

Finally, using the boundedness of M on $L^{\frac{r}{s}}(\mathbb{R}^n)$, we obtain

$$\left\| \sum_{Q \in \mathcal{C}} \left(\langle |h|^s \rangle_Q \right)^{\frac{1}{s}} \chi_Q \right\|_{L^r(\mathbb{R}^n)} \lesssim \|h\|_{L^r(\mathbb{R}^n)},$$

which completes the proof.

□

Now, we proof the third part of Theorem 4.4.1.

Proof of (III). Let $A = \|\mathcal{S}_{g,\varphi}\|_{H^p(\mathbb{R}^n)\to L^q(\mathbb{R}^n)}$. As in the proof of (II), consider the function $g_{i,N}$ and the operator $P_{i,N}$ as defined in (4.44). For simplicity of notation, we use g' and P instead of $g_{i,N}$ and $P_{i,N}$, respectively, as in (4.45). Furthermore, we note that, as in the previous case, $\|P\|_{H^p(\mathbb{R}^n)\to H^q(\mathbb{R}^n)} \lesssim A$.

We begin by showing that g' belongs to $H^r(\mathbb{R}^n)$. Since it is a finite sum of terms $\Delta_j^\theta(g')$, it is enough to show that for each $j \in \mathbb{Z}$, the function $\Delta_j^\theta(g')$ belongs to $L^r(\mathbb{R}^n)$. To this aim, we note that the boundedness of the operator P implies that

$$\|(\varphi_{2^{-j}} * f)\Delta_j^\theta(g')\|_{L^q(\mathbb{R}^n)} \leq A\|f\|_{H^p(\mathbb{R}^n)},$$

holds for all compactly supported functions. Since $\Delta_j^\theta(g')$ has compact Fourier support and $\int \varphi = 1$, an application of Lemma 4.4.3 implies that $\Delta_j^\theta(g')$ belongs to $L^r(\mathbb{R}^n)$, and thus $\|g'\|_{H^r(\mathbb{R}^n)} < \infty$. Next, we fix a dyadic cube Q_0 and apply Lemma 4.4.4 to g' with $s < r$, $\eta = \frac{1}{2}$, and a large α which will be determined later. We then have a sparse collection of cubes $\mathcal{C}$ and numbers $\lambda_Q \in \mathbb{Z} \cup \{-\infty\}$ such that

$$S_\alpha(g'|Q_0) \lesssim \sum_{Q\in\mathcal{C}} 2^{\lambda_Q}\chi_Q + \alpha^{-1}\sum_{Q\in\mathcal{C}} \left\langle M^*_{\nabla\psi}(g')^s\right\rangle_Q^{\frac{1}{s}}\chi_Q, \tag{4.64}$$

$$|Q| \lesssim |\{S_\alpha(g'|Q) \geq 2^{\lambda_Q - 1}\}|. \tag{4.65}$$

Now, (4.64) implies that

$$\|S_\alpha(g'|Q_0)\|_{L^r(\mathbb{R}^n)} \lesssim \|\sum_{Q\in\mathcal{C}} 2^{\lambda_Q}\chi_Q(x)\|_{L^r(\mathbb{R}^n)} + \alpha^{-1}\|\sum_{Q\in\mathcal{C}} \left\langle M^*_{\nabla\psi}(g')(x)^s\right\rangle_Q^{\frac{1}{s}}\chi_Q\|_{L^r(\mathbb{R}^n)}. \tag{4.66}$$

For the first term, we use Lemma 4.2.1 and get

$$\| \sum_{Q\in\mathcal{C}} 2^{\lambda_Q}\chi_Q(x)\|_{L^r(\mathbb{R}^n)} \lesssim \left(\sum_{Q\in\mathcal{C}} 2^{r\lambda_Q}|Q| \right)^{\frac{1}{r}}, \tag{4.67}$$

and an application of Lemma 4.4.5 provides the following estimate for the second term:

$$\| \sum_{Q\in\mathcal{C}} \langle M^*_{\nabla\psi}(g')(x)^s\rangle_Q^{\frac{1}{s}} \chi_Q\|_{L^r(\mathbb{R}^n)} \lesssim \|M^*_{\nabla\psi}(g')\|_{L^r(\mathbb{R}^n)} \lesssim \|g'\|_{H^r(\mathbb{R}^n)},$$

where, in the last line, we used the boundedness of non-tangential maximal functions on $H^r(\mathbb{R}^n)$. Putting the above two bounds together with (4.66), we obtain

$$\|S_\alpha(g'|Q_0)\|_{L^r(\mathbb{R}^n)} \lesssim \left(\sum_{Q\in\mathcal{C}} 2^{r\lambda_Q}|Q| \right)^{\frac{1}{r}} + \alpha^{-1}\|g'\|_{H^r(\mathbb{R}^n)}. \tag{4.68}$$

Now, we proceed to estimate the main term of the above inequality. To do this, let $\tilde{\chi}$ be the function provided by Lemma 4.4.2. Then, for each cube $Q \in \mathcal{C}$, with center c_Q let

$$\tilde{\chi}_Q(x) = \tilde{\chi}(\frac{x-c_Q}{2\sqrt{n}l(Q)}) = \tau^{c_Q}\delta^{2\sqrt{n}l(Q)}\tilde{\chi}(x), \quad x \in \mathbb{R}^n.$$

Here, we summarize properties of the above function:

$$|\varphi_t * \tilde{\chi}_Q| \geq \frac{1}{3}\chi_Q, \quad 0 < t \leq \alpha l(Q), \tag{4.69}$$

$$|\tilde{\chi}_Q| \lesssim \chi_{cQ}, \tag{4.70}$$

$$\|\tilde{\chi}_Q\|_{H^p(\mathbb{R}^n)} \lesssim |Q|^{\frac{1}{p}}. \tag{4.71}$$

To see the first property note that from Lemma 4.4.2 we have

$$|\varphi_t * \tilde{\chi}_Q| = |\tau^{c_Q}\delta^{2\sqrt{n}l(Q)}(\varphi_{\frac{t}{2\sqrt{n}l(Q)}} * \tilde{\chi})| \geq \frac{1}{3}\chi_{B_{2\sqrt{n}l(Q)}(c_Q)} \geq \frac{1}{3}\chi_Q, \quad \frac{t}{2\sqrt{n}l(Q)} \leq \alpha.$$

The second and third properties also follow from the properties of $\tilde{\chi}$ and from using dilation and translation. From now on, we fix a finite sub-collection of $\mathcal{C}$ and denote it by $\mathcal{C}'$. Then, take a sequence of independent Bernoulli random variables $\{\epsilon_Q = \pm 1\}_{Q\in\mathcal{C}'}$ and consider the following random function

$$f_\epsilon = \sum_{Q\in\mathcal{C}'} \epsilon_Q 2^{t\lambda_Q}\tilde{\chi}_Q, \quad t = \frac{r}{p}.$$

The first thing to note is that

$$\|f_\epsilon\|_{H^p(\mathbb{R}^n)} \lesssim \left(\sum_{Q\in\mathcal{C}'} 2^{r\lambda_Q}|Q|\right)^{\frac{1}{p}}, \quad 0 < p < \infty. \tag{4.72}$$

To see this, we note that for $0 < p \leq 1$ sub-additivity implies

$$\|f_\epsilon\|_{H^p(\mathbb{R}^n)} \leq \left(\sum_{Q\in\mathcal{C}'} 2^{tp\lambda_Q}\|\tilde{\chi}_Q\|^p_{H^p(\mathbb{R}^n)}\right)^{\frac{1}{p}} \lesssim \left(\sum_{Q\in\mathcal{C}'} 2^{r\lambda_Q}|Q|\right)^{\frac{1}{p}},$$

where the last estimate follows from (4.71). Also, for $1 < p < \infty$, from (4.70) we have

$$\|f_\epsilon\|_{H^p(\mathbb{R}^n)} \lesssim \|f_\epsilon\|_{L^p(\mathbb{R}^n)} \lesssim \|\sum_{Q\in\mathcal{C}'} 2^{t\lambda_Q}\chi_{cQ}\|_{L^p(\mathbb{R}^n)},$$

which after noting that the collection of concentric dilations $\{cQ : Q \in \mathcal{C}'\}$ is $c^{-n}\eta$-sparse, and using Lemma 4.2.1 implies that

$$\|f_\epsilon\|_{H^p(\mathbb{R}^n)} \lesssim \|\sum_{Q\in\mathcal{C}'} 2^{t\lambda_Q}\chi_{cQ}\|_{L^p(\mathbb{R}^n)} \lesssim \left(\sum_{Q\in\mathcal{C}'} 2^{r\lambda_Q}|Q|\right)^{\frac{1}{p}},$$

which proves the claim. Next, recall that in the defining expression of the operator P

$$P(f) = \sum_{j\in\mathbb{Z}} \varphi_{2^{-j}} * f\Delta^\theta_j(g') = \sum_{\substack{j\in m\mathbb{Z}+i \\ |j|\leq N}} \varphi_{2^{-j}} * f\Delta^\psi_j(g),$$

every two consecutive terms have Fourier support at sufficiently large distance from each other provided by the large magnitude of m. Therefore, if we take a sequence of independent Bernoulli random variables $\{\omega_j = \pm 1\}_{j\in\mathbb{Z}}$ and modify the operator P as

$$P_\omega(h) := \sum_{\substack{j\in m\mathbb{Z}+i \\ |j|\le N}} \omega_j \varphi_{2^{-j}} * h\Delta_j^{\psi}(g),$$

we still have

$$S_\theta(P_\omega(h))(x) = \left(\sum_{\substack{j\in m\mathbb{Z}+i \\ |j|\le N}} |\varphi_{2^{-j}} * h\Delta_j^{\psi}(g)|(x)^2 \right)^{\frac{1}{2}} \le \mathcal{S}_{g,\varphi}(h)(x), \quad x\in\mathbb{R}^n,$$

which implies that

$$\|P_\omega\|_{H^p(\mathbb{R}^n)\to H^q(\mathbb{R}^n)} \simeq \|P_\omega\|_{H^p(\mathbb{R}^n)\to \dot{H}^q(\mathbb{R}^n)} \lesssim A. \tag{4.73}$$

Now, for a fixed ω and ϵ we have

$$P_\omega(f_\epsilon) = \sum_{j\in\mathbb{Z}} \Delta_j(g')\omega_j\varphi_{2^{-j}} * \sum_{Q\in\mathcal{C}'} \epsilon_Q 2^{t\lambda_Q}\tilde{\chi}_Q = \sum_{Q\in\mathcal{C}'} \epsilon_Q 2^{t\lambda_Q} \sum_{j\in\mathbb{Z}} \omega_j\varphi_{2^{-j}} * \tilde{\chi}_Q\Delta_j(g').$$

Then, we get

$$\|P_\omega(f_\epsilon)\|_{L^q(\mathbb{R}^n)} \le \|P_\omega(f_\epsilon)\|_{H^q(\mathbb{R}^n)} \le A\|f_\epsilon\|_{H^p(\mathbb{R}^n)},$$

and from (4.72) we obtain

$$\|P_\omega(f_\epsilon)\|_{L^q(\mathbb{R}^n)} \lesssim A\left(\sum_{Q\in\mathcal{C}'} 2^{r\lambda_Q}|Q|\right)^{\frac{1}{p}},$$

or equivalently

$$\int_{\mathbb{R}^n} \left| \sum_{Q\in\mathcal{C}'} \epsilon_Q 2^{t\lambda_Q} \sum_{j\in\mathbb{Z}} \omega_j \varphi_{2^{-j}} * \tilde{\chi}_Q(x)\Delta_j(g')(x) \right|^q dx \lesssim A^q \left(\sum_{Q\in\mathcal{C}'} 2^{r\lambda_Q}|Q| \right)^{\frac{q}{p}}.$$

Taking expectation with respect to ϵ first, and using Khintchine inequality gives us

$$\int_{\mathbb{R}^n} \left(\sum_{Q\in\mathcal{C}'} 2^{2t\lambda_Q} \left| \sum_{j\in\mathbb{Z}} \omega_j \varphi_{2^{-j}} * \tilde{\chi}_Q(x)\Delta_j(g')(x) \right|^2 \right)^{\frac{q}{2}} dx \lesssim A^q \left(\sum_{Q\in\mathcal{C}'} 2^{r\lambda_Q}|Q| \right)^{\frac{q}{p}},$$

and then taking expectation with respect to ω implies that

$$\int_{\mathbb{R}^n} \mathbb{E} \left(\sum_{Q\in\mathcal{C}'} 2^{2t\lambda_Q} \left| \sum_{j\in\mathbb{Z}} \omega_j \varphi_{2^{-j}} * \tilde{\chi}_Q(x)\Delta_j(g')(x) \right|^2 \right)^{\frac{q}{2}} dx \lesssim A^q \left(\sum_{Q\in\mathcal{C}'} 2^{r\lambda_Q}|Q| \right)^{\frac{q}{p}}.$$

Now, let us call

$$F(x) = \mathbb{E} \left(\sum_{Q\in\mathcal{C}'} 2^{2t\lambda_Q} \left| \sum_{j\in\mathbb{Z}} \omega_j \varphi_{2^{-j}} * \tilde{\chi}_Q(x)\Delta_j(g')(x) \right|^2 \right)^{\frac{q}{2}}, \quad x \in \mathbb{R}^n, \tag{4.74}$$

then the above inequality is nothing but

$$\|F\|_{L^1(\mathbb{R}^n)} \lesssim A^q \left(\sum_{Q\in\mathcal{C}'} 2^{r\lambda_Q}|Q| \right)^{\frac{q}{p}}. \tag{4.75}$$

Our next task is to show that

$$|Q| \lesssim |\{F \gtrsim 2^{r\lambda_Q}\} \cap Q|, \quad Q \in \mathcal{C}'. \tag{4.76}$$

In order to do this, we note that

$$F(x) \geq 2^{tq\lambda_Q} \mathbb{E} \left| \sum_{j\in\mathbb{Z}} \omega_j \varphi_{2^{-j}} * \tilde{\chi}_Q(x)\Delta_j(g')(x) \right|^q,$$

which after using Khintchine inequality implies that

$$F(x) \gtrsim 2^{tq\lambda_Q} \left(\sum_{j\in\mathbb{Z}} |\varphi_{2^{-j}} * \tilde{\chi}_Q(x) \Delta_j(g')(x)|^2 \right)^{\frac{q}{2}}, \quad x \in \mathbb{R}^n.$$

Also, from (4.69)

$$|\varphi_{2^{-j}} * \tilde{\chi}_Q| \geq \frac{1}{3}\chi_Q, \quad 2^{-j} \leq \alpha l(Q),$$

we obtain

$$F(x) \gtrsim 2^{tq\lambda_Q} \left(\sum_{2^{-j}\leq \alpha l(Q)} |\Delta_j(g')(x)|^2 \right)^{\frac{q}{2}} = 2^{tq\lambda_Q} S_\alpha(g'|Q)(x)^q, \quad x \in Q.$$

Now, recall (4.65) stating that

$$|Q| \lesssim |\{S_\alpha(g'|Q)(x) \geq 2^{\lambda_Q - 1}\}|,$$

which together with the above inequality implies that

$$|Q| \lesssim |\{F(x) \gtrsim 2^{q(t+1)\lambda_Q}\} \cap Q|.$$

Now, it is enough to note that

$$q(t+1) = r, \quad t = \frac{r}{p}, \quad \frac{1}{q} = \frac{1}{p} + \frac{1}{r},$$

which proves (4.76). Having this inequality in hand, we can follow the same line of reasoning as in the dyadic case, which we will do now. First, let us partition cubes in $\mathcal{C}'$ as

$$\mathcal{C}'_k = \{Q \in \mathcal{C}' : \lambda_Q = k\}, \quad k \in \mathbb{Z} \cup \{-\infty\}.$$

Then we have

$$\sum_{Q\in\mathcal{C}'} 2^{r\lambda_Q}|Q| = \sum_{k\in\mathbb{Z}} 2^{kr} \sum_{Q\in\mathcal{C}'_k} |Q|.$$

So, if $\mathcal{C}''_k$ is the collection of maximal cubes in $\mathcal{C}'_k$, it follows from sparseness of $\mathcal{C}'_k$ that we have

$$\sum_{k\in\mathbb{Z}} 2^{kr} \sum_{Q\in\mathcal{C}'_k} |Q| \lesssim \sum_{k\in\mathbb{Z}} 2^{kr} \sum_{Q\in\mathcal{C}''_k} |Q|,$$

which together with (4.76), and noting that maximal cubes in $\mathcal{C}''_k$ are disjoint implies that

$$\sum_{Q\in\mathcal{C}'} 2^{r\lambda_Q}|Q| \lesssim \sum_{k\in\mathbb{Z}} 2^{kr} \sum_{Q\in\mathcal{C}''_k} |\{F \gtrsim 2^{kr}\}\cap Q| \lesssim \sum_{k\in\mathbb{Z}} 2^{kr}|\{F \gtrsim 2^{kr}\}| \lesssim \|F\|_{L^1(\mathbb{R}^n)}.$$

Next, we use (4.75) and obtain

$$\sum_{Q\in\mathcal{C}'} 2^{r\lambda_Q}|Q| \lesssim A^q \left(\sum_{Q\in\mathcal{C}'} 2^{r\lambda_Q}|Q|\right)^{\frac{q}{p}},$$

and noting that the right hand side is finite we get

$$\left(\sum_{Q\in\mathcal{C}'} 2^{r\lambda_Q}|Q|\right)^{\frac{1}{r}} \lesssim A,$$

and after taking the supremum over all finite sub-collections of $\mathcal{C}$ we get

$$\left(\sum_{Q\in\mathcal{C}} 2^{r\lambda_Q}|Q|\right)^{\frac{1}{r}} \lesssim A.$$

Then, we recall (4.68) and obtain

$$\|S_\alpha(g'|Q_0)\|_{L^r(\mathbb{R}^n)} \lesssim A + \alpha^{-1}\|g'\|_{H^r(\mathbb{R}^n)}.$$

Now, we consider 2^n large dyadic cubes of the form $[0, \pm 2^L]^n$, apply the above inequality to each of them and then after letting L tends to infinity we finally get

$$\|S(g')\|_{L^r(\mathbb{R}^n)} \lesssim A + \alpha^{-1}\|g'\|_{H^r(\mathbb{R}^n)}.$$

So, there exists a polynomial U such that

$$\|g' - U\|_{H^r(\mathbb{R}^n)} \lesssim \|S(g')\|_{L^r(\mathbb{R}^n)} \lesssim A + \alpha^{-1}\|g'\|_{H^r(\mathbb{R}^n)},$$

however since we already showed that $g' \in L^r(\mathbb{R}^n)$, the polynomial U must be zero and we conclude

$$\|g'\|_{H^r(\mathbb{R}^n)} \lesssim A + \alpha^{-1}\|g'\|_{H^r(\mathbb{R}^n)}.$$

Now by choosing α large enough and noting that $\|g'\|_{H^r(\mathbb{R}^n)}$ is finite, we get $\|g'\|_{H^r(\mathbb{R}^n)} \lesssim A$. Then, recall that $g' = g_{i,N}$ and

$$g_N = \sum_{|j|\leq N} \Delta_j g = \sum_{0\leq i<m} g_{i,N},$$

which implies that the sequence of functions

$$S_N(g) = (\sum_{|j|\leq N} |\Delta_j g|^2)^{\frac{1}{2}},$$

is bounded in $L^r(\mathbb{R}^n)$, and thus Fatou lemma implies

$$\|S(g)\|_{L^r(\mathbb{R}^n)} \lesssim A,$$

which means that there exists a polynomial U' such that

$$\|g - U'\|_{H^r(\mathbb{R}^n)} \lesssim A,$$

and this prove (III), and completes the proof of Theorem 4.4.1. □

Chapter 5

The Operator Norm of Paraproducts on Bi-parameter Hardy Spaces

5.1 Introduction

The one-parameter dyadic paraproduct operator with symbol g is defined as

$$\pi_g(f) := \sum_{I \in \mathcal{D}} g_I \langle f \rangle_I h_I,$$

where $\mathcal{D}$ denotes the collection of all dyadic intervals on the real line, h_I is the L^2-normalized Haar wavelet associated with the interval I, g_I represents the Haar coefficient of g, and $\langle f \rangle_I$ is the average of f over the interval I. Bilinear forms (in terms of f and g) of this type are among the most important ones in harmonic analysis, with many applications to PDEs. This is mainly due to the fact that many bilinear forms can be decomposed in terms of paraproducts and their adjoints. For instance, the product of two functions f and g can be written as

$$fg = \pi_g(f) + \pi_f(g) + \pi_g^*(f),$$

where π_g^* denotes the adjoint of π_g. For this reason, the boundedness properties of these operators play a crucial role in analyzing various problems in harmonic analysis and PDEs. We refer the reader to [5] for a brief introduction and to [58] for an excellent exposition of this subject. See also [1, 4, 7, 8, 17, 32] for various boundedness properties of paraproducts. Because of their importance, it is natural to wonder about the norm of these operators acting between various function spaces. The first result in this direction appeared in [7], where it was shown that

$$\|\pi_g\|_{L^p(\mathbb{R}) \to \dot{L}^p(\mathbb{R})} \simeq \|g\|_{\mathrm{BMO}_d(\mathbb{R})}, \quad 1 < p < \infty.$$

Here, $\dot{L}^p(\mathbb{R})$ is the Lebesgue space $L^p(\mathbb{R})$ modulo constants, with the quotient norm defined as

$$\|f\|_{\dot{L}^p(\mathbb{R})} := \inf_{c \in \mathbb{R}} \|f - c\|_{L^p(\mathbb{R})},$$

and $\mathrm{BMO}_d(\mathbb{R})$ stands for the dyadic BMO, the space of functions with bounded mean oscillation on the dyadic intervals of the real line. In addition, recently in [36], the authors extended the above result to the off-diagonal range of exponents. More precisely, they showed, among other things, that

$$\|\pi_g\|_{L^p(\mathbb{R}) \to \dot{L}^q(\mathbb{R})} \simeq \|g\|_{\dot{L}^r(\mathbb{R})} \quad \text{where} \quad \frac{1}{q} = \frac{1}{p} + \frac{1}{r}, \quad 1 < p, q, r < \infty.$$

The next progress in the subject comes from our recent work in [65], where our main contribution was to replace Lebesgue spaces with Hardy spaces and to lift the restrictions on the exponents on the right-hand side of the above two results.

Specifically, it was shown that

$$\|\pi_g\|_{H^p_d(\mathbb{R})\to \dot{H}^p_d(\mathbb{R})} \simeq \|g\|_{\mathrm{BMO}_d(\mathbb{R})}, \quad 0<p<\infty, \tag{5.1}$$

$$\|\pi_g\|_{H^p_d(\mathbb{R})\to \dot{H}^q_d(\mathbb{R})} \simeq \|g\|_{\dot{H}^r_d(\mathbb{R})}, \quad \frac{1}{q}=\frac{1}{p}+\frac{1}{r}, \quad 0<p,q,r<\infty, \tag{5.2}$$

where the $H^p_d(\mathbb{R})$-norm is defined as the $L^p(\mathbb{R})$-norm of the dyadic maximal function, and the $\dot{H}^p_d(\mathbb{R})$-norm refers to the $L^p(\mathbb{R})$-norm of the dyadic square function. We also obtained similar results for Fourier paraproducts in the continuous setting, and we refer the reader to [65] for precise statements in this context.

The proof idea in [36], which is similar to that in [38], heavily relies on the duality of Lebesgue spaces. However, as demonstrated in [65], this approach fails when $0 < q < 1$. In [65], we therefore adopted a direct method. By using a suitable pointwise sparse domination of the square function of the symbol g, we were able to construct a test function f such that, when testing the operator π_g on f, we could recover the $L^r(\mathbb{R})$-norm of the square function of g, achieving the desired result.

In the present paper, we focus on operators acting on bi-parameter Hardy spaces. There are various types of bi-parameter paraproducts, and the one we study is the most similar to the one-parameter operator. It is defined as

$$\pi_g(f) := \sum_{R\in\mathcal{D}\otimes\mathcal{D}} g_R \langle f\rangle_R h_R,$$

where the sum is taken over the collection of all dyadic rectangles in the plane (see the next section for precise definitions and notation). We refer the reader to [58] for an exposition of multi-parameter paraproducts. See also [46,56,57]. To obtain an analog of (5.1) and (5.2) for this operator, we employ a similar strategy and demonstrate how the one-parameter arguments in [65] can be modified to work in the multi-parameter setting. As in our previous work, we first present our arguments in the dyadic setting and then extend them to the continuous setting. Before doing so, let us fix some definitions and notation.

5.2 preliminaries

As mentioned before, by $\mathcal{D}$ we mean the collection of all dyadic intervals in $\mathbb{R}$, and $\mathcal{D}\otimes\mathcal{D}$ stands for the collection of all dyadic rectangles in the plane. For $f \in L^1_{loc}(\mathbb{R}^2)$ and E a measurable set of finite positive measure, we denote the average of f over E by

$$\langle f \rangle_E := |E|^{-1} \int_E f.$$

For such a function, the bi-parameter dyadic maximal operator is defined as

$$M_d(f)(x) := \sup_{\substack{x\in R \\ R\in\mathcal{D}\otimes\mathcal{D}}} |\langle f \rangle_R|.$$

When the supremum is taken over all rectangles (not necessarily dyadic) with sides parallel to the axes, the resulting operator is denoted by M and is referred to as the strong maximal operator. We also need the following version of this operator:

$$M_s(f)(x) := M(|f|^s)^{\frac{1}{s}}(x), \quad 0 < s < \infty,$$

as well as the bi-parameter Fefferman-Stein inequality, which states that

$$\|\big(\sum_j M_s(f_j)^2\big)^{\frac{1}{2}}\|_{L^p(\mathbb{R}^2)} \lesssim \|\big(\sum_j |f_j|^2\big)^{\frac{1}{2}}\|_{L^p(\mathbb{R}^2)}, \quad 0 < s < p < \infty.$$

The next notation deals with an enlargement of open sets Ω, which occurs quite often in multi-parameter theory. We use the following somewhat standard notation:

$$\tilde{\Omega} := \{M(\chi_\Omega) > \frac{1}{2}\},$$

and recall that

$$\Omega \subset \tilde{\Omega}, \quad |\tilde{\Omega}| \lesssim |\Omega|,$$

which follows from the boundedness of M on, say, $L^2(\mathbb{R}^2)$. See [16, 22, 23] for the proof of the above assertions and other properties of maximal operators in the product setting.

Now, to modify our one-parameter arguments in [65], we need to generalize some properties of sparse families of cubes to simple families of measurable sets.

Definition 5.2.1. *A sequence of measurable sets $\{\Omega_i\}_{i\geq 0}$ of finite measure is called contracting if*

$$\Omega_{i+1} \subset \Omega_i, \quad |\Omega_{i+1}| \leq \frac{1}{2}|\Omega_i|, \quad i = 0, 1, 2, \ldots$$

In the next lemma, we show that when dealing with L^p norms, a contracting family can be treated as a disjoint family.

Lemma 5.2.1. *Let $\{\Omega_i\}_{i\geq 0}$ be a contracting family. Then, for any sequence of non-negative numbers $\{a_i\}_{i\geq 0}$ and any $0 < p < \infty$, we have*

$$\|\sum_{i\geq 0} a_i \chi_{\Omega_i}\|_{L^p} \simeq_p \big(\sum_{i\geq 0} a_i^p |\Omega_i|\big)^{\frac{1}{p}}.$$

Proof. First, note that

$$\big(\sum_{i\geq 0} a_i^p |\Omega_i|\big)^{\frac{1}{p}} \lesssim \|\sum_{i\geq 0} a_i \chi_{\Omega_i \setminus \Omega_{i+1}}\|_{L^p} \leq \|\sum_{i\geq 0} a_i \chi_{\Omega_i}\|_{L^p}.$$

So it remains to prove the other direction. For $0 < p \leq 1$, from sub-additivity we get

$$\|\sum_{i\geq 0} a_i \chi_{\Omega_i}\|_{L^p}^p \leq \sum_{i\geq 0} \|a_i \chi_{\Omega_i}\|_{L^p}^p = \sum_{i\geq 0} a_i^p |\Omega_i|.$$

Therefore, we are left to prove

$$\|\sum_{i\geq 0} a_i \chi_{\Omega_i}\|_{L^p} \lesssim \big(\sum_{i\geq 0} a_i^p |\Omega_i|\big)^{\frac{1}{p}}, \quad 1 < p < \infty. \tag{5.3}$$

To this aim, take a function $g \in L^{p'}$, where p' is the Hölder's conjugate of p and note that

$$\int g \sum_{i\geq 0} a_i \chi_{\Omega_i} \leq 2 \sum_{i\geq 0} a_i \left\langle |g| \right\rangle_{\Omega_i} |\Omega_i \backslash \Omega_{i+1}| \leq 2 \int m(g) \sum_{i\geq 0} a_i \chi_{\Omega_i \backslash \Omega_{i+1}},$$

where in the above

$$m(g)(x) := \sup_{\substack{x\in\Omega_i \\ i\geq 0}} \left\langle |g| \right\rangle_{\Omega_i}.$$

For now, let us assume that this operator is bounded on $L^{p'}$ with norm that depends only on p. Then, the above last inequality gives us

$$\int g \sum_{i\geq 0} a_i \chi_{\Omega_i} \leq 2 \|\sum_{i\geq 0} a_i \chi_{\Omega_i \backslash \Omega_{i+1}}\|_{L^p} \|m(g)\|_{L^{p'}} \lesssim \big(\sum_{i\geq 0} a_i^p |\Omega_i|\big)^{\frac{1}{p}} \|g\|_{L^{p'}},$$

which implies the desired inequality in (5.3). To show why m is bounded on Lebesgue spaces, note that m is L^∞-bounded with norm 1 and is of weak-$(1,1)$ type. These two facts, along with interpolation, imply that m is bounded on $L^{p'}$ for $1 < p < \infty$. The weak-$(1,1)$ bound for m follows from the fact that for $\lambda > 0$, we have

$$|\{m(g) > \lambda\}| = |\bigcup_{\left\langle |g| \right\rangle_{\Omega_i} > \lambda} \Omega_i| = |\Omega_{i_0}| \leq \lambda^{-1} \|g\|_{L^1},$$

where $i_0 = \min\{i \geq 0 \mid \left\langle |g| \right\rangle_{\Omega_i} > \lambda\}$. This completes the proof.

□

Another useful property of contracting families is that their large portions form a Carleson family of sets.

Lemma 5.2.2. *Let* $\{\Omega_i\}_{i\geq 0}$ *be a contracting family and suppose* $\{E_i\}_{i\geq 0}$ *is a family of measurable sets such that for* $0<\eta\leq 1$*, we have*

$$E_i \subset \Omega_i, \quad |E_i| \geq \eta|\Omega_i|, \quad i\geq 0.$$

Then, for any $A\subset\{0,1,2,\ldots\}$ *we have*

$$\sum_{i\in A}|E_i| \leq \frac{2}{\eta}|\bigcup_{i\in A}E_i|.$$

Proof. Let $k=\min\{i\in A\}$ and note that

$$\sum_{i\in A}|E_i| \leq \sum_{i\in A}|\Omega_i| \leq 2|\Omega_k| \leq \frac{2}{\eta}|E_k| \leq \frac{2}{\eta}\bigcup_{i\in A}E_i|.$$

□

For more on Carleson families of measurable sets, we refer the reader to [35] and the references therein.

5.2.1 Bi-parameter Dyadic Hardy Spaces

Next, we turn to the definition of bi-parameter Hardy spaces in the dyadic setting. For a dyadic rectangle $R\in\mathcal{D}\otimes\mathcal{D}$, we define

$$h_R := h_I\otimes h_J, \quad \text{where } R=I\times J,$$

with h_I and h_J being the L^2-normalized Haar wavelets associated with intervals I and J, respectively. As it is well-known $\{h_R\}_{R\in\mathcal{D}\otimes\mathcal{D}}$ forms an orthonormal basis for

$L^2(\mathbb{R}^2)$ and for $f \in L^2(\mathbb{R}^2)$ we have

$$f = \sum_{R \in \mathcal{D}\otimes\mathcal{D}} f_R h_R, \quad f_R := \langle f, h_R \rangle .$$

To define dyadic Hardy spaces rigorously, we first make the following definition.

Definition 5.2.2. *A dyadic distribution f is a family of real numbers $\{f_R\}_{R \in \mathcal{D}\otimes\mathcal{D}}$, and formally is written as*

$$f := \sum_{R \in \mathcal{D}\otimes\mathcal{D}} f_R h_R.$$

For such an object the dyadic square function is defined as

$$S_d(f)(x) := \Big(\sum_{R \in \mathcal{D}\otimes\mathcal{D}} f_R^2 \frac{\chi_R}{|R|}(x) \Big)^{\frac{1}{2}}.$$

Definition 5.2.3. *For $0 < p < \infty$, the space $H^p_d(\mathbb{R} \otimes \mathbb{R})$ is the completion of the space of all real valued locally integrable functions f with*

$$\|f\|_{H^p_d(\mathbb{R}\otimes\mathbb{R})} := \|M_d(f)\|_{L^p(\mathbb{R}^2)} < \infty.$$

Definition 5.2.4. *For $0 < p < \infty$, the space $\dot{H}^p_d(\mathbb{R} \otimes \mathbb{R})$ is the space of all dyadic distributions f with*

$$\|f\|_{\dot{H}^p_d(\mathbb{R}\otimes\mathbb{R})} := \|S_d(f)\|_{L^p(\mathbb{R}^2)} < \infty.$$

Now, let us explain the relation between these two (quasi)-norms. When a priori f is a bounded function with compact support, or more generally a function in $L^1_{loc}(\mathbb{R}^2) \cap L^q(\mathbb{R}^2)$ for some $0 < q < \infty$, the two mentioned quantities are equivalent, with constants independent of this a priori information. The inequality

$$\|M_d(f)\|_{L^p(\mathbb{R}^2)} \lesssim_p \|S_d(f)\|_{L^p(\mathbb{R}^2)}, \tag{5.4}$$

follows from sub-linearity of the operator M_d, and an atomic decomposition of f, which can be easily derived from the square function [15, 70]. The other direction

$$\|S_d(f)\|_{L^p(\mathbb{R}^2)} \lesssim_p \|M_d(f)\|_{L^p(\mathbb{R}^2)}, \tag{5.5}$$

is harder to prove and follows from the following distributional inequality due to Brossard [10]. See also [70].

Theorem (Brossard)**.** *There there exists a constant C such that for any compactly supported function $f \in L^2(\mathbb{R}^2)$, and any $\delta > 0$, we have*

$$\int_{\{\{M_d(f)>\delta\}^{\sim}\}^c} S_d(f)^2 \leq C\left(\delta^2 |\,\{M_d(f) > \delta\}\,| + \int_{\{M_d(f)\leq\delta\}} M_d(f)^2\right). \tag{5.6}$$

In [10], the inequality (5.6) was proved in the general setting of bi-parameter regular martingales. The analog of this inequality for bi-harmonic functions is due to Merryfield [53], and in the one-parameter setting, this inequality was previously established by Fefferman and Stein [21]. To the best of our knowledge, the only application of this type of inequalities so far has been to prove (5.5). However, in the next sections, we will show that the inequality (5.6) can be useful in certain constructions (see lemmas 5.3.2 and 5.4.1). It is also worth mentioning that all the arguments presented in the following sections extend verbatim to any number of parameters. We chose to work in the bi-parameter setting only because we could not find the analog of (5.6) for a higher number of parameters, though we believe such a theorem should hold [22, 53].

Here, we would like to mention that in the literature, the above two spaces, which are different, are both referred to as $H^p_d(\mathbb{R}\otimes\mathbb{R})$. This is also true for the one-parameter version $H^p_d(\mathbb{R})$ and the continuous versions $H^p(\mathbb{R})$, $H^p(\mathbb{R}\otimes\mathbb{R})$, etc. However, these spaces are not identical, and their equivalence must be understood through some a priori information or by using quotient norms. The reason lies in the cancellation within the square function, which is absent in the maximal function. For instance, since all bi-parameter Haar coefficients of a function of the form

$$E(x_1, x_2) = f_1(x_1) + f_2(x_2)$$

are zero, adding or subtracting such functions does not affect the bi-parameter dyadic square function but does change the dyadic maximal function.

The last function space to recall is the bi-parameter dyadic BMO.

Definition 5.2.5. *The space $BMO_d(\mathbb{R} \otimes \mathbb{R})$ is the space of all dyadic distributions f with*

$$\|f\|_{BMO_d(\mathbb{R}\otimes\mathbb{R})} := \sup_{\Omega} \big(|\Omega|^{-1} \sum_{R\subseteq\Omega} f_R^2\big)^{\frac{1}{2}} < \infty.$$

In the above, the supremum is taken over all open subsets of the plane, and as Carleson famously showed [14], it is not sufficient to consider only rectangles, which is in sharp contrast with the one-parameter theory, where intervals or cubes can replace open sets. The next fact to recall is the bi-parameter John-Nirenberg inequality, which states that for any $0 < p < \infty$, we have

$$\|f\|_{BMO_d(\mathbb{R}\otimes\mathbb{R})} \simeq_p \sup_{\Omega} \langle S_d(f|\Omega)^p\rangle_{\Omega}^{\frac{1}{p}}, \quad S_d(f|\Omega) := \big(\sum_{R\subseteq\Omega} f_R^2 \frac{\chi_R}{|R|}\big)^{\frac{1}{2}}, \tag{5.7}$$

(see [58] for a proof). Throughout the paper, we will use the notation on the right-hand side for localizations of the square function. Last but not least, we recall the well-known C. Fefferman's duality

$$\dot{H}^1_d(\mathbb{R}\otimes\mathbb{R})^* \cong BMO_d(\mathbb{R}\otimes\mathbb{R}), \tag{5.8}$$

the proof of which in this setting is due to Bernard [6]. We refer the reader to [70] for the proofs and an exposition of Hardy spaces in the general setting of martingales.

5.2.2 Bi-parameter Hardy spaces in the Continuous Setting

Let ψ be a Schwartz function on $\mathbb{R}^2$ with

$$\text{supp}(\hat{\psi}) \subseteq \{\xi = (\xi_1, \xi_2) \mid 0 < \mathbf{a} \leq |\xi_1|, |\xi_2| \leq \mathbf{b} < \infty\}, \tag{5.9}$$

$$\sum_{(j_1,j_2)\in\mathbb{Z}^2} \hat{\psi}(2^{-j_1}\xi_1, 2^{-j_2}\xi_2) = 1, \quad \xi_1, \xi_2 \neq 0. \tag{5.10}$$

Then, the bi-parameter Littlewood-Paley projections and the associated square function of a tempered distribution f are defined as

$$\Delta_j(f) := \psi_{2^{-j}} * f, \quad \hat{\psi}_{2^{-j}}(\xi) := \hat{\psi}(2^{-j_1}\xi_1, 2^{-j_2}\xi_2), \quad j \in \mathbb{Z}^2, \tag{5.11}$$

$$S_\psi(f)(x) := \big(\sum_{j\in\mathbb{Z}^2} |\Delta_j(f)(x)|^2\big)^{\frac{1}{2}}, \quad x \in \mathbb{R}^2. \tag{5.12}$$

In addition, for a Schwartz function φ with $\int \varphi = 1$ let

$$M_\varphi(f)(x) := \sup_{t_1,t_2>0} |\varphi_t * f(x)|, \quad \varphi_t(x) := \frac{1}{t_1 t_2}\varphi(\frac{x_1}{t_1}, \frac{x_2}{t_2}), \quad x = (x_1, x_2) \in \mathbb{R}^2, \quad t_1, t_2 > 0.$$

be the bi-parameter smooth vertical maximal operator and

$$M_\varphi^*(f)(x) := \sup_{\substack{|x_1-y_1|\leq t_1, |x_2-y_2|\leq t_2 \\ t_1,t_2>0}} |\varphi_t * f(y_1, y_2)|,$$

be its non-tangential analog.

Definition 5.2.6. *For $0 < p < \infty$, the space $H^p(\mathbb{R}\otimes\mathbb{R})$ consists of all tempered distributions f with*

$$\|f\|_{H^p(\mathbb{R}\otimes\mathbb{R})} := \|M_\varphi(f)\|_{L^p(\mathbb{R}^2)} < \infty,$$

and $\dot{H}^p(\mathbb{R} \times \mathbb{R})$ *is the space of all tempered distributions with*

$$\|f\|_{\dot{H}^p(\mathbb{R}\otimes\mathbb{R})} := \|S_\psi(f)\|_{L^p(\mathbb{R}^2)} < \infty.$$

In the above, both spaces are independent of the choice of φ, ψ, and $H^p(\mathbb{R} \otimes \mathbb{R})$ is identical to $L^p(\mathbb{R}^2)$ for $1 < p < \infty$. Similar to the dyadic setting, with an a priori information such as $f \in L^q(\mathbb{R}^2)$ for some $0 < q < \infty$, we have

$$\|f\|_{H^p(\mathbb{R}\otimes\mathbb{R})} \simeq_{p,\varphi,\psi} \|f\|_{\dot{H}^p(\mathbb{R}\otimes\mathbb{R})}.$$

See [53] and [22] for the proof of these.

Next, we recall quasi-orthogonal expansions and their properties. Let, θ be a Schwartz function whose Fourier transform is compactly supported and is equal to 1 on the support of $\hat{\psi}$. Then we may write

$$\Delta_j(g) = \theta_{2^{-j}} * \Delta_j(g) = \sum_{R\in\mathcal{D}\otimes\mathcal{D}_j} \int_R \theta_{2^{-j}}(x-y)\Delta_j(g)(y)dy, \quad j \in \mathbb{Z}^2,$$

where in th above $\mathcal{D} \otimes \mathcal{D}_j$ denotes the collection of dyadic rectangles with sides $2^{-j_1} \times 2^{-j_2}$ $(j = (j_1, j_2))$. Now, let

$$\lambda_R(g) = \sup_{y\in R} |\Delta_j(g)(y)||R|^{\frac{1}{2}}, \quad a_R(g) = \lambda_R(g)^{-1} \int_R \theta_{2^{-j}}(x-y)\Delta_j(g)(y)dy, \quad R \in \mathcal{D}\otimes\mathcal{D}_j. \tag{5.13}$$

Then, one can show that

$$\|g\|_{\dot{H}^p(\mathbb{R}\otimes\mathbb{R})} \simeq \|\sum_{R\in\mathcal{D}\otimes\mathcal{D}} \lambda_R(g)h_R\|_{\dot{H}^p_d(\mathbb{R}\otimes\mathbb{R})}, \quad g = \sum_{R\in\mathcal{D}\otimes\mathcal{D}} \lambda_R(g)a_R(g), \tag{5.14}$$

roughly giving an isomorphism between $\dot{H}^p(\mathbb{R} \otimes \mathbb{R})$ and $\dot{H}^p_d(\mathbb{R} \otimes \mathbb{R})$. In addition, for any sub-collection $\mathcal{C}$ of dyadic rectangles we have

$$\|\sum_{R\in\mathcal{C}} \lambda_R(g) a_R(g)\|_{\dot{H}^p(\mathbb{R}\otimes\mathbb{R})} \lesssim \|\sum_{R\in\mathcal{C}} \lambda_R(g) h_R\|_{\dot{H}^p_d(\mathbb{R}\otimes\mathbb{R})}. \tag{5.15}$$

The above inequalities follow from almost orthogonality of functions $a_R(g)$, Fefferman-Stein vector valued inequality, and a well-known inequality which captures the local constancy of band-limited functions. Bellow we bring its bi-parameter version.

Theorem. *Let f be a function with supp$(\hat{f}) \subseteq \{\xi \mid |\xi_1| \leq t_1, |\xi_2| \leq t_2\}$, then we have*

$$|f(y)| \lesssim_s (1 + t_1 t_2 |x_1 - y_1||x_2 - y_2|) M_s(f)(x), \quad 0 < s < \infty. \tag{5.16}$$

See [30], p. 94, for the proof in the one-parameter setting. Here, we would like to mention that we could not find a proof of (5.14) and (5.15) as stated above in the literature. However, the one-parameter arguments presented in [30] work for any number of parameters with minor changes. See also [15].

At the end, let us recall the continuous BMO in the product setting. Similar to the one-parameter theory, BMO($\mathbb{R}\otimes\mathbb{R}$) can be defined in terms of Carleson measures, but we do not use this fact here and instead mention its quasi-orthogonal characterization, which is quite similar to (5.14). More precisely, BMO($\mathbb{R} \otimes \mathbb{R}$) is the space of all tempered distributions g such that the dyadic distribution

$$\sum_{R\in\mathcal{D}\otimes\mathcal{D}} \lambda_R(g) h_R$$

belongs to $BMO_d(\mathbb{R} \otimes \mathbb{R})$, and the equivalence of norms holds:

$$\|g\|_{BMO(\mathbb{R}\otimes\mathbb{R})} \simeq \big(\sup_{\Omega} |\Omega|^{-1} \sum_{R\subseteq\Omega} \lambda_R(g)^2\big)^{\frac{1}{2}}.$$

See [15] for the proof.

5.3 Bi-parameter Dyadic Paraproducts

In the bi-parameter theory, there are different types of paraproducts arising in the product of two functions [58]. The one considered here is of the form

$$\pi_g(f) := \sum_{R\in\mathcal{D}\otimes\mathcal{D}} \langle f\rangle_R g_R h_R,$$

where g is a dyadic distribution, $f \in L^1_{\text{loc}}(\mathbb{R}^2)$, and $\pi_g(f)$ is understood as a dyadic distribution. It is easy to see that

$$S_d(\pi_g(f)) \leq S_d(g)M_d(f),$$

and thus, by Hölder's inequality, we obtain

$$\|\pi_g\|_{H^p_d(\mathbb{R}\otimes\mathbb{R})\to\dot{H}^q_d(\mathbb{R}\otimes\mathbb{R})} \leq \|g\|_{\dot{H}^r_d(\mathbb{R}\otimes\mathbb{R})}, \quad \frac{1}{q} = \frac{1}{p} + \frac{1}{r}, \quad 0 < p, q, r < \infty.$$

Additionally, atomic decomposition together with the John-Nirenberg inequality imply that

$$\|\pi_g\|_{H^p_d(\mathbb{R}\otimes\mathbb{R})\to\dot{H}^p_d(\mathbb{R}\otimes\mathbb{R})} \lesssim \|g\|_{BMO_d(\mathbb{R}\otimes\mathbb{R})}, \quad 0 < p < \infty.$$

In this section we will show that in both cases the reverse direction holds, and this is the content of our main theorem.

Theorem 5.3.1. *Let g be a dyadic distribution. Then we have*

$$\|\pi_g\|_{H^p_d(\mathbb{R}\otimes\mathbb{R})\to\dot{H}^q_d(\mathbb{R}\otimes\mathbb{R})} \simeq \|g\|_{\dot{H}^r_d(\mathbb{R}\otimes\mathbb{R})}, \quad \frac{1}{q} = \frac{1}{p} + \frac{1}{r}, \quad 0 < p, q, r < \infty, \tag{I}$$

$$\|\pi_g\|_{H^p_d(\mathbb{R}\otimes\mathbb{R})\to\dot{H}^p_d(\mathbb{R}\otimes\mathbb{R})} \simeq \|g\|_{BMO_d(\mathbb{R}\otimes\mathbb{R})}, \quad 0 < p < \infty. \tag{II}$$

Using the duality relation (5.8), we get another characterization of $BMO_d(\mathbb{R}\otimes\mathbb{R})$, which is dual to (II).

Corollary 5.3.1. *Let g be a function with only finitely many non-zero Haar coefficients. Then for the operator*

$$\pi_g^*(f) := \sum_{R\in\mathcal{D}\otimes\mathcal{D}} f_R g_R \frac{\chi_R}{|R|},$$

we have

$$\|g\|_{BMO_d(\mathbb{R}\otimes\mathbb{R})} \simeq \|\pi_g^*\|_{BMO_d(\mathbb{R}\otimes\mathbb{R})\to BMO_d(\mathbb{R}\otimes\mathbb{R})}.$$

The proof of theorem 5.3.1 is based on the following two lemmas.

Lemma 5.3.1. *Let g be a dyadic distribution with Haar coefficients that are zero except for finitely many, and such that for an open subset Ω_0 with finite measure, we have*

$$g = \sum_{R\subseteq\Omega_0} g_R h_R.$$

Then, there exists an absolute constant $0 < \eta_0 < 1$ (independent of g and Ω_0) such that for any $0 < \eta < \eta_0$, the following holds;

There exist a contracting family of open sets $\{\Omega_i\}_{i\geq 0}$ starting from Ω_0, and a sequence of numbers $\{\lambda_i \in \mathbb{Z}\cup\{-\infty\}\}_{i\geq 0}$ such that

$$\eta|\Omega_i| \leq |\{S_d(g|\Omega_i) > 2^{\lambda_i - 1}\}|, \quad \text{if} \quad \lambda_i\in\mathbb{Z} \quad i\geq 0, \tag{5.17}$$

$$\|g\|_{\dot{H}^r_d(\mathbb{R}\otimes\mathbb{R})} \simeq_{\eta,r} \big(\sum_{i\geq 0} 2^{r\lambda_i}|\Omega_i|\big)^{\frac{1}{r}}, \quad 0 < r < \infty. \tag{5.18}$$

In the above we use the convention that $2^{-\infty} = 0$, and $r\cdot(-\infty) = -\infty$.

Lemma 5.3.2. *For $0 < p, \varepsilon \leq 1$ and any open set of finite positive measure $\Omega\subset\mathbb{R}^2$,* *there exists a function* $\tilde{\chi}_\Omega \in H^p_d(\mathbb{R}\otimes\mathbb{R}) \cap L^2(\mathbb{R}^2)$ *such that*

$$\|\tilde{\chi}_\Omega\|_{H^p_d(\mathbb{R}\otimes\mathbb{R})} \lesssim_{p,\varepsilon} |\Omega|^{\frac{1}{p}}, \tag{5.19}$$

$$|\Omega'| \geq (1-\varepsilon)|\Omega|, \quad \Omega' = \Big\{x \in \Omega \mid \inf_{\substack{x\in R\\ R\subseteq\Omega}} \langle\tilde{\chi}_\Omega\rangle_R \geq \frac{1}{2}\Big\}. \tag{5.20}$$

Let us accept these two lemmas and prove Theorem 5.3.1.

Proof of Theorem 5.3.1. First, we consider the case (I), and to this aim let

$$A = \|\pi_g\|_{H^p_d(\mathbb{R}\otimes\mathbb{R})\to\dot{H}^q_d(\mathbb{R}\otimes\mathbb{R})}.$$

We must show that

$$\|g\|_{\dot{H}^r_d(\mathbb{R}\otimes\mathbb{R})} \lesssim A. \tag{5.21}$$

Now, observe that if g' is a dyadic distribution with finitely many non-zero Haar coefficients which are the same as those of g, the operator norm of $\pi_{g'}$ is not larger than A. This follows because

$$S_d(\pi_{g'}(f)) \leq S_d(\pi_g(f)).$$

If we can show that (5.21) holds with g replaced by g', then by the monotone convergence theorem, we can easily deduce (5.21) for g. Therefore, without loss of generality, we assume that g has only finitely many non-zero coefficients, and that the associated dyadic rectangles are all contained in Ω_0. We then apply Lemma 5.3.1 to g with sufficiently small η, which yields a finite contracting sequence of open sets $\{\Omega_i\}_{i\geq 0}$ and numbers $\{\lambda_i\}_{i\geq 0}$ with the described properties. Next, when $0 < p \leq 1$, we apply Lemma 5.3.2 with $\varepsilon = \frac{1}{2}\eta$ to each Ω_i to obtain the function $\tilde{\chi}_{\Omega_i}$. For $1 < p < \infty$, we set $\tilde{\chi}_{\Omega_i} = \chi_{\Omega_i}$ for $i \geq 0$.

As the next step, we take a sequence of independent Bernoulli random variables $\{\omega_i\}_{i\geq 0}$ with $\mathbb{P}(\omega_i = \pm 1) = \frac{1}{2}$ and construct the following random function:

$$f_\omega = \sum_{i\geq 0} \omega_i 2^{t\lambda_i} \tilde{\chi}_{\Omega_i}, \quad \text{where } t = \frac{r}{q} - 1.$$

Now, from (5.4), we have

$$\int |\sum_{R\in\mathcal{D}\otimes\mathcal{D}} g_R \langle f_\omega \rangle_R h_R|^q \lesssim A^q \|f_\omega\|^q_{H^p_d(\mathbb{R}\otimes\mathbb{R})}.$$

Additionally, if we choose another sequence of independent Bernoulli random variables $\{\epsilon_R\}_{R\in\mathcal{D}\otimes\mathcal{D}}$ and multiply these signs to the coefficients of g, the operator norm of the resulting dyadic paraproduct remains unchanged. This yields

$$\int |\sum_{R,i\geq 0} \epsilon_R \omega_i g_R 2^{t\lambda_i} \langle \tilde{\chi}_{\Omega_i} \rangle_R h_R|^q = \int |\sum_{R\in\mathcal{D}\otimes\mathcal{D}} \epsilon_R g_R \langle f_\omega \rangle_R h_R|^q \lesssim A^q \|f_\omega\|^q_{H^p_d(\mathbb{R}\otimes\mathbb{R})}.$$

Now, taking the expectation with respect ot both variables and using the bi-parameter version of Khintchine inequality gives us

$$\int \Big(\sum_{R,i\geq 0} g_R^2 2^{2t\lambda_i} \langle \tilde{\chi}_{\Omega_i} \rangle_R^2 \frac{\chi_R}{|R|}\Big)^{\frac{q}{2}} \lesssim A^q \mathbb{E} \|f_\omega\|^q_{H^p_d(\mathbb{R}\otimes\mathbb{R})}.$$

Next, call the function under the sign of integral F and rewrite the above inequality as

$$\|F\|_{L^1(\mathbb{R}^2)} \lesssim A^q \mathbb{E} \|f_\omega\|^q_{H^p_d(\mathbb{R}\otimes\mathbb{R})}, \quad F = \Big(\sum_{R,i\geq 0} g_R^2 2^{2t\lambda_i} \langle \tilde{\chi}_{\Omega_i} \rangle_R^2 \frac{\chi_R}{|R|}\Big)^{\frac{q}{2}}. \tag{5.22}$$

At this point, let us first estimate the right hand side of the above inequality, and assume initially that $0 < p \leq 1$. From sub-additivity, (5.19) in Lemma 5.3.2, and the fact that $tp = r$ we obtain

$$\|f_\omega\|^p_{H^p_d(\mathbb{R}\otimes\mathbb{R})} \leq \sum_{i\geq 0} 2^{pt\lambda_i}\|\tilde{\chi}_{\Omega_i}\|^p_{H^p_d(\mathbb{R}\otimes\mathbb{R})} \lesssim \sum_{i\geq 0} 2^{pt\lambda_i}|\Omega_i| = \sum_{i\geq 0} 2^{r\lambda_i}|\Omega_i|.$$

In the case that $1 < p < \infty$ we also have

$$\|f_\omega\|_{H^p_d(\mathbb{R}\otimes\mathbb{R})} \simeq \|f_\omega\|_{L^p(\mathbb{R}^2)} \leq \|\sum_{i\geq 0} 2^{t\lambda_i}\chi_{\Omega_i}\|_{L^p(\mathbb{R}^2)} \simeq \big(\sum_{i\geq 0} 2^{r\lambda_i}|\Omega_i|\big)^{\frac{1}{p}},$$

where in the above we used Lemma 5.2.1. Therefore, from the above and (5.22) we must have

$$\|F\|_{L^1(\mathbb{R}^2)} \lesssim A^q\big(\sum_{i\geq 0} 2^{r\lambda_i}|\Omega_i|\big)^{\frac{q}{p}}. \tag{5.23}$$

The next step is to observe that (5.22) implies

$$F \geq 2^{tq\lambda_i}\big(\sum_{R\subseteq\Omega_i} g_R^2\,\langle\tilde{\chi}_{\Omega_i}\rangle_R^2\,\frac{\chi_R}{|R|}\big)^{\frac{q}{2}}.$$

Then, since from (5.20) we have

$$\langle\tilde{\chi}_{\Omega_i}\rangle_R \geq \frac{1}{2}, \quad R\subseteq\Omega_i, \quad x\in R, \quad x\in\Omega_i',$$

we conclude that

$$F(x) \geq 2^{(1+t)q\lambda_i-2q} = 2^{r\lambda_i-2q}, \quad x\in E_i = \Omega_i'\cap\left\{S_d(g|\Omega_i) > 2^{\lambda_i-1}\right\}, \quad i\geq 0. \tag{5.24}$$

Now, since

$$|\Omega_i'| \geq (1-\frac{1}{2}\eta)|\Omega_i|, \quad |\left\{S_d(g|\Omega_i) > 2^{\lambda_i-1}\right\}| \geq \eta|\Omega_i|, \quad i\geq 0,$$

we have that

$$|E_i| \geq \frac{1}{2}\eta|\Omega_i|, \quad E_i \subseteq \Omega_i, \quad i \geq 0,$$

which together with (5.24) and Lemma 5.2.2 implies that

$$\sum_{i\geq 0} 2^{r\lambda_i}|\Omega_i| = \sum_{k\in\mathbb{Z}} 2^{rk} \sum_{\lambda_i = k} |\Omega_i| \lesssim_\eta \sum_{k\in\mathbb{Z}} 2^{rk} \sum_{\lambda_i=k} |E_i| \lesssim \tag{5.25}$$

$$\sum_{k\in\mathbb{Z}} 2^{rk}|\bigcup_{\lambda_i=k} E_i| \leq \sum_{k\in\mathbb{Z}} 2^{rk}|\{F > 2^{rk-2q}\}| \simeq \|F\|_{L^1(\mathbb{R}^2)}. \tag{5.26}$$

From this and (5.23) we obtain

$$\sum_{i\geq 0} 2^{r\lambda_i}|\Omega_i| \lesssim A^q\big(\sum_{i\geq 0} 2^{r\lambda_i}|\Omega_i|\big)^{\frac{q}{p}},$$

which after noting that the sum appearing on both sides is finite gives us

$$\big(\sum_{i\geq 0} 2^{r\lambda_i}|\Omega_i|\big)^{\frac{1}{r}} \lesssim A.$$

Finally, we recall (5.18) and obtain

$$\|g\|_{\dot{H}^r_d(\mathbb{R}\otimes\mathbb{R})} \lesssim A,$$

which is the desired inequality in (5.21), and this finishes the proof of case (I).

Now we turn to the case (II), which can be proven by almost the same argument, and we outline only the required changes. This time, we need to show that for any open subset of finite measure Ω_0 we have

$$\int S_d(g|\Omega_0)^p \lesssim A^p|\Omega_0|, \quad A = \|\pi_g\|_{H^p_d(\mathbb{R}\otimes\mathbb{R})\to \dot{H}^p_d(\mathbb{R}\otimes\mathbb{R})}.$$

We note that if $1 < p < \infty$, the result follows trivially by inserting the function χ_{Ω_0}

into the operator and observing that

$$S_d(g|\Omega_0) \leq S_d(\pi_g(\chi_{\Omega_0})).$$

Thus, we only need to consider the case $0 < p \leq 1$. In this scenario, the only change required in the argument is to set $t = 0$ and note that

$$\|f_\omega\|^p_{H^p_d(\mathbb{R}\otimes\mathbb{R})} \lesssim |\Omega_0|.$$

Then, the inequality (5.23) is replaced with

$$\|F\|_{L^1(\mathbb{R}^2)} \lesssim A^p|\Omega_0|,$$

and in (5.25) and (5.26) we replace r with p, yielding

$$\int S_d(g|\Omega_0)^p \simeq \sum_{i\geq 0} 2^{p\lambda_i}|\Omega_i| \lesssim \|F\|_{L^1(\mathbb{R}^2)} \lesssim A^p|\Omega_0|.$$

This establishes the result we sought, thereby completing the proof of case (II) and the proof of Theorem 5.3.1.

□

Now, we turn to the proof of Lemma 5.3.1, which is based on an iteration of a well-known argument in multi-parameter theory.

Proof of Lemma 5.3.1. Let λ_0 be the smallest number in $\mathbb{Z} \cup \{-\infty\}$ such that

$$|\{S_d(g|\Omega_0) > 2^{\lambda_0}\}| \leq \eta|\Omega_0|. \tag{5.27}$$

The above inequality is satisfied for sufficiently large values of $\lambda_0 \in \mathbb{Z}$, ensuring the existence of such a number. Also, if $\lambda_0 \in \mathbb{Z}$, we must have

$$\eta|\Omega_0| \leq |\{S_d(g|\Omega_0) > 2^{\lambda_0 - 1}\}|.$$

Next, let

$$\Omega_1 := \{S_d(g|\Omega_0) > 2^{\lambda_0}\}^{\sim} \cap \Omega_0,$$

and note that if η is small enough, we have

$$|\Omega_1| \lesssim \eta|\Omega_0| \leq \frac{1}{2}|\Omega_0|.$$

Now, we repeat the process by replacing Ω_0 with Ω_1, and get λ_1 and Ω_2. Continuing this, we end up with a sequence of contracting open sets $\{\Omega_i\}_{i \geq 0}$ and numbers $\lambda_i \in \mathbb{Z} \cup \{-\infty\}$ such that

$$\Omega_{i+1} = \{S_d(g|\Omega_i) > 2^{\lambda_i}\}^{\sim} \cap \Omega_i, \quad i \geq 0, \tag{5.28}$$

$$\eta|\Omega_i| \leq |\{S_d(g|\Omega_i) > 2^{\lambda_i - 1}\}|, \quad \text{if} \quad \lambda_i \in \mathbb{Z}. \tag{5.29}$$

Clearly these sets and numbers satisfy (5.17), and it remains to show that (5.18) holds as well. First, we show that

$$\sum_{i \geq 0} 2^{r\lambda_i}|\Omega_i| \lesssim \|S_d(g)\|^r_{L^r(\mathbb{R}^2)},$$

which follows from

$$\sum_{i \geq 0} 2^{r\lambda_i}|\Omega_i| = \sum_{k \in \mathbb{Z}} \sum_{\lambda_i = k} 2^{rk}|\Omega_i| \leq \eta^{-1} \sum_{k \in \mathbb{Z}} 2^{rk} \sum_{\lambda_i = k} |\{S_d(g|\Omega_i) > 2^{k-1}\}|,$$

and noting that from (5.29), we are allowed to apply Lemma 5.2.2 to get

$$\sum_{i\geq 0} 2^{r\lambda_i}|\Omega_i| \lesssim \sum_{k\in\mathbb{Z}} 2^{rk}|\{S_d(g) > 2^{k-1}\}| \simeq \|S_d(g)\|^r_{L^r(\mathbb{R}^2)}.$$

In order to get the other direction, we decompose g as

$$g = \sum_{i\geq 0} g_i, \quad g_i := \sum_{\substack{R\subseteq\Omega_i \\ R\nsubseteq\Omega_{i+1}}} g_R h_R, \quad i \geq 0, \tag{5.30}$$

with the convention that if $\Omega_i = \emptyset$, then $g_i = 0$. Then, we consider two separate cases either $0 < r \leq 2$ or $2 < r < \infty$. For the first case, we apply sub-linearity and get

$$\|S_d(g)\|^r_{L^r(\mathbb{R}^2)} = \int \Big|\sum_{i\geq 0} S_d(g_i)^2\Big|^{\frac{r}{2}} \leq \sum_{i\geq 0}\int_{\Omega_i} \left|S_d(g_i)^2\right|^{\frac{r}{2}} \leq \sum_{i\geq 0} |\Omega_i|^{1-\frac{r}{2}} \left(\int S_d(g_i)^2\right)^{\frac{r}{2}},$$

where in the last estimate we used Hölder's inequality. Therefore, it is enough to show that

$$\int S_d(g_i)^2 \lesssim 2^{2\lambda_i}|\Omega_i|, \quad i \geq 0, \tag{5.31}$$

which together with the previous estimate gives us

$$\|S_d(g)\|^r_{L^r(\mathbb{R}^2)} \lesssim \sum_{i\geq 0} 2^{r\lambda_i}|\Omega_i|,$$

which is what we are looking for. Now, to see why (5.31) holds, note that

$$\int S_d(g_i)^2 = \sum_{\substack{R\subseteq\Omega_i \\ R\nsubseteq\Omega_{i+1}}} g_R^2 \leq 2 \sum_{\substack{R\subseteq\Omega_i \\ R\nsubseteq\Omega_{i+1}}} g_R^2 \left|\{S_d(g|\Omega_i) > 2^{\lambda_i}\}^c \cap R\right| |R|^{-1},$$

where the last inequality follows from the fact that $R \nsubseteq \Omega_{i+1}$, and this gives us

$$\int S_d(g_i)^2 \leq 2\int_{\{S_d(g|\Omega_i)>2^{\lambda_i}\}^c} S_d(g|\Omega_i)^2 \leq 2\cdot 2^{2\lambda_i}|\Omega_i|, \quad i \geq 0,$$

showing that (5.31) holds, and the proof of this case is finished. Next, we consider the case $2 < r < \infty$, which follows from duality together with a similar argument to the one presented above. Let φ be a function with $\|\varphi\|_{L^{(\frac{r}{2})'}(\mathbb{R}^2)} = 1$. Then we have

$$\int S_d(g)^2\varphi = \sum_{i\geq 0}\int S_d(g_i)^2\varphi = \sum_{i\geq 0}\sum_{\substack{R\subseteq\Omega_i\\ R\nsubseteq\Omega_{i+1}}} g_R^2\langle\varphi\rangle_R \leq$$

$$2\sum_{i\geq 0}\sum_{\substack{R\subseteq\Omega_i\\ R\nsubseteq\Omega_{i+1}}} g_R^2|\{S_d(g|\Omega_i) > 2^{\lambda_i}\}^c\cap R||R|^{-1}|\langle\varphi\rangle_R| \leq$$

$$2\sum_{i\geq 0}\int_{\{S_d(g|\Omega_i)>2^{\lambda_i}\}^c} S_d(g|\Omega_i)^2 M_d(\varphi) \leq 2\int\sum_{i\geq 0} 2^{2\lambda_i}\chi_{\Omega_i}M_d(\varphi) \leq$$

$$2\|\sum_{i\geq 0} 2^{2\lambda_i}\chi_{\Omega_i}\|_{L^{\frac{r}{2}}(\mathbb{R}^2)}\|M_d(\varphi)\|_{L^{(\frac{r}{2})'}(\mathbb{R}^2)} \lesssim \left(\sum_{i\geq 0} 2^{r\lambda_i}|\Omega_i|\right)^{\frac{2}{r}},$$

where for the last estimate Lemma 5.2.1 is used, and this shows that

$$\|S_d(g)\|_{L^r(\mathbb{R}^2)} \lesssim \big(\sum_{i\geq 0} 2^{r\lambda_i}|\Omega_i|\big)^{\frac{1}{r}},$$

which completes the proof of this case and Lemma 5.3.1. □

Remark 5.3.1. *The decomposition* (5.30)

$$g = \sum_{i\geq 0} g_i, \quad g_i := \sum_{\substack{R\subseteq\Omega_i\\ R\nsubseteq\Omega_{i+1}}} g_R h_R, \quad i\geq 0,$$

is an atomic decomposition of g with the property that the supporting open sets of its atoms form a contracting family. Indeed, the above argument shows that g_i are L^p-atoms (although not normalized) for $H^r_d(\mathbb{R}\otimes\mathbb{R})$ for any positive values of p and r. See [40] for the counterpart of this result in the one-parameter theory. See also [55] p. 42.

As our last job in this section we give the proof of Lemma 5.3.2.

Proof of Lemma 5.3.2. Let $0 < \delta < 1$ be a small number to be determined later, and consider the following three enlargements of Ω

$$\Omega_1 := \{M_d(\chi_\Omega) > \delta\}, \quad \Omega_2 := \tilde{\Omega}_1, \quad \Omega_3 := \{M_d(\chi_{\Omega_2}) > \delta\},$$

and let

$$\tilde{\chi}_\Omega = \chi_\Omega - f, \quad f = \sum_{R \not\subseteq \Omega_3} \langle \chi_\Omega, h_R \rangle \, h_R.$$

Then since

$$\tilde{\chi}_\Omega = \sum_{R \subseteq \Omega_3} \langle \chi_\Omega, h_R \rangle \, h_R,$$

from Hölder's inequality and the fact that $|\Omega_3| \lesssim_\delta |\Omega|$ we have

$$\int S_d(\tilde{\chi}_\Omega)^p \leq |\Omega_3|^{1-\frac{p}{2}} \left(\int S_d(\chi_\Omega)^2 \right)^{\frac{p}{2}} \lesssim_{\delta,p} |\Omega|.$$

So, $\tilde{\chi}_\Omega$ satisfies the first property in (5.19), and it remains to choose δ so small that the second property holds as well. To this aim, we note that

$$\langle \tilde{\chi}_\Omega \rangle_R + \langle f \rangle_R = 1, \quad R \subseteq \Omega,$$

and thus

$$\Omega' = \big\{ x \in \Omega | \quad \inf_{\substack{x \in R \\ R \subseteq \Omega}} \langle \tilde{\chi}_\Omega \rangle_R \geq \frac{1}{2} \big\} \supseteq \{M_d(f) \geq \frac{1}{2}\}^c \cap \Omega,$$

which implies that if f is such that we have

$$|\{M_d(f) \geq \frac{1}{2}\}| \leq \varepsilon |\Omega|, \tag{5.32}$$

then $\tilde{\chi}_\Omega$ satisfies (5.20) and we are done. Now since we have

$$|\{M_d(f) \geq \frac{1}{2}\}| \lesssim \int_{\mathbb{R}^2} |f|^2, \tag{5.33}$$

it is enough to show that the right hand is small. Here, we observe that

$$\int_{\mathbb{R}^2} |f|^2 = \int_{\mathbb{R}^2} S_d(f)^2 = \int_{\Omega_2} S_d(f)^2 + \int_{\Omega_2^c} S_d(f)^2, \tag{5.34}$$

and may estimate the first term by

$$\int_{\Omega_2} S_d(f)^2 = \sum_{R\not\subseteq\Omega_3} \langle \chi_\Omega, h_R\rangle^2 \frac{|\Omega_2 \cap R|}{|R|} \leq \delta \sum_{R\not\subseteq\Omega_3} \langle \chi_\Omega, h_R\rangle^2 \leq \delta|\Omega|, \tag{5.35}$$

where in the above we used the fact that $R \not\subseteq \Omega_3$. For the second term we have

$$\int_{\Omega_2^c} S_d(f)^2 \leq \int_{\Omega_2^c} S_d(\chi_\Omega)^2 \lesssim \delta^2 |\{M_d(\chi_\Omega) > \delta\}| + \int_{\{M_d(\chi_\Omega)\leq\delta\}} M_d(\chi_\Omega)^2,$$

which follows from (5.6). Now, using boundedness of M_d of $L^{\frac{3}{2}}(\mathbb{R}^2)$ yields

$$\int_{\Omega_2^c} S_d(f)^2 \lesssim \delta^2 |\{M_d(\chi_\Omega) > \delta\}| + \delta^{\frac{1}{2}} \int_{\{M_d(\chi_\Omega)\leq\delta\}} M_d(\chi_\Omega)^{\frac{3}{2}} \lesssim \delta^{\frac{1}{2}}|\Omega|, \tag{5.36}$$

and putting (5.33),(5.34), (5.35), and (5.36) together gives us

$$|\{M_d(f) \geq \frac{1}{2}\}| \lesssim \int_{\mathbb{R}^2} |f|^2 \lesssim \delta^{\frac{1}{2}}|\Omega|,$$

showing that by choosing δ small enough (5.32) holds, and this completes the proof.

□

Here, it is worth mentioning that in the one-parameter theory the same construction yields a function with the stronger property that

$$\tilde{\chi}_\Omega(x) \geq \frac{1}{2}, \quad x \in \Omega.$$

To see this, let $f = \sum_{I\not\subseteq\tilde{\Omega}} f_I h_I$ and $\tilde{\chi}_\Omega = \chi_\Omega - f$. Then, the function f, is constant on each maximal dyadic interval of $\tilde{\Omega}$, and thus is not larger than $\frac{1}{2}$ on Ω, which proves the above inequality. Regarding this, we ask the following question:

Question. Let $0 < p \leq 1$, and Ω be an open subset with $|\Omega| < \infty$. Does there exist a function $\tilde{\chi}_\Omega$ with the following two properties?

$$\|\tilde{\chi}_\Omega\|_{H^p_d(\mathbb{R}\otimes\mathbb{R})} \lesssim_p |\Omega|^{\frac{1}{p}}, \quad \tilde{\chi}_\Omega(x) \geq \frac{1}{2}, \quad x \in \Omega.$$

5.4 Bi-Parameter Fourier Paraproducts

In this section we explain how similar results can be obtained for Fourier paraproducts of the form

$$\Pi_g(f)(x) := \sum_{j\in\mathbb{Z}^2} \varphi_{2^{-j}} * f(x)\Delta_j(g)(x), \quad x \in \mathbb{R}^2,$$

where in the above $\hat{\varphi} \subseteq \{\xi \mid |\xi_1|, |\xi_2| \leq \mathbf{a}'\}$ with $\mathbf{a}' < \mathbf{a}$, and $\mathbf{a}$ is the same as in (5.9). Then, one can show that

$$\|\Pi_g\|_{H^p(\mathbb{R}\otimes\mathbb{R})\to\dot{H}^q(\mathbb{R}\otimes\mathbb{R})} \lesssim \|g\|_{\dot{H}^r(\mathbb{R}\otimes\mathbb{R})}, \quad \frac{1}{q} = \frac{1}{p} + \frac{1}{r}, \quad 0 < p, q, r < \infty,$$
$$\|\Pi_g\|_{H^p(\mathbb{R}\otimes\mathbb{R})\to\dot{H}^p(\mathbb{R}\otimes\mathbb{R})} \lesssim \|g\|_{BMO(\mathbb{R}\otimes\mathbb{R})}, \quad 0 < p < \infty.$$

Indeed, the above inequalities follow from the support properties of $\hat{\varphi}$ and $\hat{\psi}$ and

$$\|\mathcal{S}_g(f)\|_{L^q(\mathbb{R}^2)} \lesssim \|g\|_{\dot{H}^r(\mathbb{R}^2)}\|f\|_{H^p(\mathbb{R}\otimes\mathbb{R})}, \quad \frac{1}{q} = \frac{1}{p} + \frac{1}{r}, \quad 0 < p, q, r < \infty, \tag{5.37}$$

$$\|\mathcal{S}_g(f)\|_{L^p(\mathbb{R}^2)} \lesssim \|g\|_{BMO(\mathbb{R}\otimes\mathbb{R})}\|f\|_{H^p(\mathbb{R}\otimes\mathbb{R})}, \quad 0 < p < \infty, \tag{5.38}$$

where in the above

$$\mathcal{S}_g(f)(x) := \Big(\sum_{j\in\mathbb{Z}^2} |\varphi_{2^{-j}} * f(x)\Delta_j(g)(x)|^2\Big)^{\frac{1}{2}}, \quad x \in \mathbb{R}^2,$$

[58]. Here, we show that the converse of (5.37) and (5.38) holds.

Theorem 5.4.1. *Let g be a tempered distribution, and $\mathcal{S}_g$ be as above. Then,*

$$\|\mathcal{S}_g\|_{H^p(\mathbb{R}\otimes\mathbb{R})\to L^q(\mathbb{R}^2)} \simeq \|g\|_{\dot{H}^r(\mathbb{R}\otimes\mathbb{R})}, \quad \frac{1}{q}=\frac{1}{p}+\frac{1}{r}, \quad 0<p,q,r<\infty, \tag{i}$$

$$\|\mathcal{S}_g\|_{H^p(\mathbb{R}\otimes\mathbb{R})\to L^p(\mathbb{R}^2)} \simeq \|g\|_{BMO(\mathbb{R}\otimes\mathbb{R})}, \qquad 0<p<\infty. \tag{ii}$$

Proof of Theorem. We prove only case (i), since the other case follows with exactly the same argument. Let, $A=\|\mathcal{S}_g\|_{H^p(\mathbb{R}\otimes\mathbb{R})\to L^q(\mathbb{R}^2)}$, we need to show that

$$\|g\|_{\dot{H}^r(\mathbb{R}\otimes\mathbb{R})} \lesssim A.$$

To this aim, recall the notation introduced in (5.13), let $x_R\in R$ be such that

$$|\Delta_j g(x_R)| = \sup_{y\in R}|\Delta_j(g)(y)|, \quad R\in\mathcal{D}\otimes\mathcal{D}_j, \quad j\in\mathbb{Z}^2,$$

and define

$$\pi'(f) := \sum_{j\in\mathbb{Z}^2}\sum_{R\in\mathcal{D}\otimes\mathcal{D}_j}\varphi_{2^{-j}} * f(x_R)\lambda_R(g)h_R.$$

Then we note that the function $\varphi_{2^{-j}} * f\Delta_j(g)$, has Fourier support in $\{\xi \mid |\xi_1|\lesssim 2^{j_1}, |\xi_2|\lesssim 2^{j_2}\}$, and thus we may apply (5.16) and get

$$|\varphi_{2^{-j}}*f(x_R)|\sup_{y\in R}|\Delta_j(g)(y)| \le \sup_{y\in R}|\varphi_{2^{-j}}*f(y)\Delta_j(g)(y)| \lesssim M_s(\varphi_{2^{-j}}*f\Delta_j(g))(x), \quad x\in R,$$

which implies that

$$\sum_{j\in\mathbb{Z}^2}\sum_{R\in\mathcal{D}\otimes\mathcal{D}}|\varphi_{2^{-j}} * f(x_R)\lambda_R(g)|^2\frac{\chi_R}{|R|} \lesssim \sum_{j\in\mathbb{Z}^2} M_s(\varphi_{2^{-j}} * f\Delta_j(g))^2(x).$$

Then applying Fefferman-Stein inequality with $0 < s < q$ yields

$$\|\pi'(f)\|_{\dot{H}^q_d(\mathbb{R}\otimes\mathbb{R})} \lesssim \|\big(\sum_{j\in\mathbb{Z}^2} M_s(\varphi_{2^{-j}} * f\Delta_j(g))^2\big)^{\frac{1}{2}}\|_{L^q(\mathbb{R}^2)} \lesssim \|\mathcal{S}_g(f)\|_{L^q(\mathbb{R}^2)} \leq A\|f\|_{H^p(\mathbb{R}\otimes\mathbb{R})}.$$

So we have

$$\|\pi'\|_{H^p(\mathbb{R}\otimes\mathbb{R})\to\dot{H}^q_d(\mathbb{R}\otimes\mathbb{R})} \lesssim A.$$

Now, it follows from (5.14) that our task is the show that

$$\|\sum_{R\in\mathcal{D}\otimes\mathcal{D}} \lambda_R(g)h_R\|_{\dot{H}^q_d(\mathbb{R}\otimes\mathbb{R})} \lesssim A,$$

which follows from an argument identical to the one presented in the proof of Theorem 5.3.1 if we replace $\tilde{\chi}_\Omega$ with its counterpart $\tilde{\tilde{\chi}}_\Omega$ in the following lemma, and this completes the proof.

□

Lemma 5.4.1. *Let $0 < p, \varepsilon \leq 1$, and φ be a Schwartz function with $\int \varphi = 1$. Then for any open set Ω with $|\Omega| < \infty$, there exists a function $\tilde{\tilde{\chi}}_\Omega \in H^p(\mathbb{R}\otimes\mathbb{R}) \cap L^2(\mathbb{R}^2)$ such that*

$$\|\tilde{\tilde{\chi}}_\Omega\|_{H^p(\mathbb{R}\otimes\mathbb{R})} \lesssim_{p,\varepsilon} |\Omega|^{\frac{1}{p}}, \tag{5.39}$$

$$|\Omega'| \geq (1-\varepsilon)|\Omega|, \quad \Omega' = \big\{x \in \Omega \mid \inf_{\substack{x,y\in R \\ R\subseteq\Omega, R\in\mathcal{D}\otimes\mathcal{D}_j, j\in\mathbb{Z}^2}} \varphi_{2^{-j}} * \tilde{\tilde{\chi}}_\Omega(y) \geq \frac{1}{2}\big\}. \tag{5.40}$$

Proof. First, we note that since φ decays rapidly and $\int \varphi = 1$, we may choose a large constant α, depending only on φ such that

$$|\varphi|_{2^{-j}} * \chi_{(\alpha R)^c}(y) < \frac{1}{4}, \quad y \in R, \quad R \in \mathcal{D}\otimes\mathcal{D}_j, \quad j \in \mathbb{Z}^2,$$

where αR is the concentric dilation of R with α. Thus, if O is an open subset with the property that $\alpha R \subseteq O$ whenever $R \subseteq \Omega$ we must have

$$\varphi_{2^{-j}} * \chi_O(y) > \frac{3}{4}, \quad y \in R \subseteq \Omega, \quad R \in \mathcal{D} \otimes \mathcal{D}_j, \quad j \in \mathbb{Z}^2. \tag{5.41}$$

So, let

$$O = \{M(\chi_\Omega) \geq \alpha^{-1}\},$$

be the first enlargement of Ω, where M is the strong maximal operator. Then, let

$$\chi_O = \sum_{R \in \mathcal{D} \otimes \mathcal{D}} \lambda_R(\chi_O) a_R(\chi_O),$$

be the quasi-orthogonal expansion of χ_O as in (5.14), and define the function

$$E := \sum_{R \in \mathcal{D} \otimes \mathcal{D}} \lambda_R(\chi_O) h_R.$$

Now, we follow the same strategy as in Lemma 5.3.2 as it follows. Take a small number $0 < \delta < 1$, to be determined later and set

$$O_1 := \{M_d(E) > \delta\}, \quad O_2 := \tilde{O}_1, \quad O_3 := \{M_d(\chi_{O_2}) > \delta\},$$

and decompose E and χ_O as

$$E = \sum_{R \subseteq O_3} E_R h_R + \sum_{R \not\subseteq O_3} E_R h_R = \tilde{E} + \tilde{f} \tag{5.42}$$

$$\chi_O = \sum_{R \subseteq O_3} \lambda_R(\chi_O) a_R(\chi_O) + \sum_{R \not\subseteq O_3} \lambda_R(\chi_O) a_R(\chi_O) = \tilde{\tilde{\chi}}_\Omega + f. \tag{5.43}$$

Next, we note that (5.15), Hölder's inequality, and (5.14) imply that

$$\|\tilde{\tilde{\chi}}_\Omega\|^p_{H^p(\mathbb{R}\otimes\mathbb{R})} \lesssim \|\tilde{E}\|^p_{\dot{H}^p_d(\mathbb{R}\otimes\mathbb{R})} \leq \int_{O_3} S_d(E|O_3)^p \leq |O_3|^{1-\frac{p}{2}} \left(\int S_d(E)^2 \right)^{\frac{p}{2}} \lesssim_{\delta,p} |O| \lesssim_{\delta,p} |\Omega|. \tag{5.44}$$

Therefore, $\tilde{\tilde{\chi}}_\Omega$ satisfies (5.39), and it remains to show that (5.40) holds as well. Now, similar to Lemma 5.3.2 we may estimate the L^2-norm of f as

$$\int_{\mathbb{R}^2} |f|^2 \lesssim \int_{\mathbb{R}^2} |\tilde{f}|^2 = \int_{O_2} |S_d(\tilde{f})|^2 + \int_{O_2^c} |S_d(\tilde{f})|^2 \leq \sum_{R \not\subseteq O_3} E_R^2 \frac{|R \cap O_2|}{|R|} + \int_{O_2^c} |S_d(E)|^2 \leq$$
$$\delta \sum_{R \not\subseteq O_3} E_R^2 + \int_{\{\{M_d(E)>\delta\}^\sim\}^c} |S_d(E)|^2 \lesssim \delta \int_{\mathbb{R}^2} |E|^2 + \delta^2 |\{M_d(E) > \delta\}| +$$
$$\int_{M_d(E) \leq \delta} M_d(E)^2 \lesssim \delta \int_{\mathbb{R}^2} |E|^2 + \delta^{\frac{1}{2}} \int_{\mathbb{R}^2} |E|^{\frac{3}{2}} \lesssim \delta^{\frac{1}{2}} |O| \lesssim \delta^{\frac{1}{2}} |\Omega|.$$

In the above we used (5.15) and (5.6). Next, recall (5.41) and note that

$$\varphi_{2^{-j}} * \tilde{\tilde{\chi}}_\Omega(y) + \varphi_{2^{-j}} * f(y) > \frac{3}{4}, \quad y \in R \subseteq \Omega, \quad R \in \mathcal{D} \otimes \mathcal{D}_j, \quad j \in \mathbb{Z}^2,$$

which implies that

$$\Omega' \supseteq \Big\{x \mid \sup_{\substack{x,y \in R \\ R \subseteq \Omega, R \in \mathcal{D} \otimes \mathcal{D}_j, j \in \mathbb{Z}^2}} \varphi_{2^{-j}} * f(y) > \frac{1}{4}\Big\}^c \cap \Omega \supseteq \Big\{M_\varphi^*(f)(x) > \frac{1}{4}\Big\}^c \cap \Omega,$$

where $M_\varphi^*(f)$ is the bi-parameter non-tangential maximal function of f. Finally, from boundedness of M_φ^* on $L^2(\mathbb{R}^2)$ and the smallness of L^2-norm of f we get

$$|\Omega \backslash \Omega'| \leq |\{M_\varphi^*(f)(x) > \frac{1}{4}\}| \lesssim \int_{\mathbb{R}^2} |f|^2 \lesssim \delta^{\frac{1}{2}} |\Omega|,$$

and thus if we choose δ small enough (5.40) holds and this completes the proof. □

Bibliography

[1] Auscher P., Hofmann S., Muscalu C., Tao T., and Thiele C., *Carleson measures, trees, extrapolation, and $T(b)$ theorems*, Publ. Mat. 46 (2002), no. 2, 257–325.

[2] Bennett C., *Another characterization of BLO*, Proc. Amer. Math. Soc. 85 (1982), no. 4, 552–556.

[3] Bennett C., DeVore R. A., and Sharpley R., *Weak-L^∞ and BMO*, Ann. of Math. (2) 113 (1981), no. 3, 601–611.

[4] Bényi Á., Maldonado D., Nahmod A. R., and Torres R. H., *Bilinear paraproducts revisited*, Math. Nachr. 283 (2010), no. 9, 1257–1276.

[5] Bényi Á., Maldonado D., and Naibo V., *What is ... a paraproduct?*, Notices Amer. Math. Soc. 57 (2010), no. 7, 858–860.

[6] Bernard A., *Espaces H^1 de martingales à deux indices. Dualité avec les martingales de type "BMO"*, Bull. Sci. Math. (2), 103 (1979), no. 3, 297–303.

[7] Blasco O., *Dyadic BMO, paraproducts and Haar multipliers*, Interpolation theory and applications, Contemp. Math., vol. 445, 11–18. Amer. Math. Soc., Providence, RI, 2007.

[8] Bony J. M., *Calcul symbolique et propagation des singularités pour les équations aux dérivées partielles non linéaires*, Ann. Sci. École Norm. Sup. (4), 14 (1981), no. 2, 209–246.

[9] Bownik M., *Boundedness of operators on Hardy spaces via atomic decompositions*, Proc. Amer. Math. Soc. 133 (2005), no. 12, 3535–3542.

[10] Brossard J., *Comparison des "normes" L_p du processus croissant et de la variable maximale pour les martingales régulières à deux indices. Théorème local correspondant*, Ann. Probab. 8 (1980), no. 6, 1183–1188.

[11] Brudnyi A., and Brudnyi Y., *Methods of Geometric Analysis in Extension and Trace Problems. Volumes 1 and 2*, Monographs in Mathematics, vol. 102, Birkhäuser, 2002.

[12] Burkholder, D. L. and Gundy, R. F., *Extrapolation and interpolation of quasi-linear operators on martingales*, Acta Math. 124 (1970), 249–304.

[13] Butaev, A. and Dafni, G., *Approximation and Extension of Functions of Vanishing Mean Oscillation*, J. Geom. Anal. 31 (2021), no. 7, 6892–6921.

[14] Carleson L., *A Counter Example for Measures Bounded on H^p for the Bi-disc*, Institut Mittag-Leffler, 1974.

[15] Chang S. Y. A., and Fefferman R., *A continuous version of duality of H^1 with BMO on the bidisc*, Ann. of Math. (2), 112 (1980), no. 1, 179–201.

[16] Chang S. Y. A., and Fefferman R., *Some recent developments in Fourier analysis and H^p-theory on product domains*, Bull. Amer. Math. Soc. (N.S.), 12 (1985), no. 1, 1–43.

[17] Coifman R. R., and Meyer Y., *Au delà des opérateurs pseudo-différentiels*, Astérisque, vol. 57, Société Mathématique de France, Paris, 1978. With an English summary.

[18] Dafni G., and Gibara R., *BMO on shapes and sharp constants*, Advances In Harmonic Analysis And Partial Differential Equations. 748 (2020), 1-33.

[19] David G., and Journé J. L., *A boundedness criterion for generalized Calderón-Zygmund operators*, Ann. of Math. (2), 120 (1984), no. 2, 371–397.

[20] Duren P. L., Romberg B. W., and Shields A. L., *Linear functionals on H^p spaces with* $0 < p < 1$, J. Reine Angew. Math., 238 (1969), 32–60.

[21] Fefferman C., and Stein E. M., *H^p spaces of several variables*, Acta Math. 129 (1972), no. 3-4, 137–193.

[22] Fefferman R., *Multiparameter Fourier analysis*, Beijing lectures in harmonic analysis (Beijing, 1984), Ann. of Math. Stud., vol. 112, 47–130, Princeton Univ. Press, Princeton, NJ, 1986.

[23] Fefferman R., and Stein E. M., *Singular integrals on product spaces*, Adv. in Math. 45 (1982), no. 2, 117–143.

[24] Figiel T., *Singular integral operators: a martingale approach*, London Math. Soc. Lecture Note Series, vol. 158, Cambridge Univ. Press, Cambridge, 1990, 95–110.

[25] García-Cuerva J., and Rubio de Francia J. L., *Weighted norm inequalities and related topics*, North-Holland Mathematics Studies, vol. 116, North-Holland Publishing Co., Amsterdam, 1985, Notas de Matemática, 104.

[26] Garsia A. M., *Martingale inequalities: Seminar notes on recent progress*, Mathematics Lecture Note Series, W. A. Benjamin, Inc., Reading, Mass.-London-Amsterdam, 1973.

[27] Garnett J., and Jones P. W., *BMO from Dyadic BMO*, Pacific J. Math. 99 (1982), no. 2, 351–371.

[28] Goldberg D. A., *local version of real Hardy spaces*, Duke Math. J. 46, (1979), 27-42.

[29] Grafakos L., *Classical Fourier analysis*, 2nd ed., Graduate Texts in Mathematics, vol. 249, Springer, New York, 2008.

[30] Grafakos L., *Modern Fourier analysis*, Second, Graduate Texts in Mathematics, vol. 250, Springer, New York, 2009.

[31] Grafakos L., He D., Kalton N. J., and Mastyło M., *Multilinear paraproducts revisited*, Rev. Mat. Iberoam 31 (2015), no. 2, 609–616.

[32] Grafakos L., and Kalton N. J., *The Marcinkiewicz multiplier condition for bilinear operators*, Studia Math. 146 (2001), no. 2, 115–156.

[33] Gundy R. F., and Stein E. M., H^p theory for the poly-disc, Proc. Nat. Acad. Sci. U.S.A. 76 (1979), no. 3, 1026–1029.

[34] Hardy G. H., *The mean value of the modulus of an analytic function*, Proc. London Math. Soc. (2) 14 (1915), 269–277.

[35] Hänninen T. S., *Equivalence of sparse and Carleson coefficients for general sets*, Ark. Mat. 56 (2018), no. 2, 333–339.

[36] Hänninen T. S., Lorist E., and Sinko J., *Weighted $l^p \to l^q$-boundedness of commutators and paraproducts in the bloom setting*, arXiv:2303.14855, (2023).

[37] Herz, C., *H_p-spaces of martingales,* $0 < p \leq 1$, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 28 (1973-74), 189–205.

[38] Hytönen T. P., *The L^p-to-L^q boundedness of commutators with applications to the Jacobian operator*, J. Math. Pures Appl. (9), 156 (2021), 351–391.

[39] Hytönen T. P., *The sharp weighted bound for general Calderón-Zygmund operators*, Ann. of Math. (2) 175 (2012), no. 3, 1473–1506.

[40] Janson S., and Jones P. W., *Interpolation between H^p spaces: the complex method*, J. Functional Analysis 48 (1982), no. 1, 58–80.

[41] John F., and Nirenberg L., *On functions of bounded mean oscillation*, Comm. Pure Appl. Math. 14 (1961), no. 3, 415-426.

[42] Jones P. W., *Extension Theorems for BMO*, Indiana Univ. Math. J. 29 (1980), no. 1, 41–66.

[43] Jones P. W., *Quasiconformal Mappings and Extendability of Functions in Sobolev Spaces*, Acta Math. 147 (1981), no. 1-2, 71–88.

[44] Jonsson A., and Wallin H., *Function Spaces on Subsets of* $\mathbb{R}^n$, Math. Rep. 2, 1984.

[45] Korenovskii A., *Mean oscillations and equimeasurable rearrangements of functions*, Springer, Berlin, 2007.

[46] Lacey M., and Metcalfe J., *Paraproducts in one and several parameters*, Forum Math. 19 (2007), no. 2, 325–351.

[47] Leoni G., *A First Course in Sobolev Spaces*, 2nd ed., Graduate Studies in Mathematics, vol. 181, American Mathematical Society, 2017.

[48] Lerner A. K. , *A pointwise estimate for the local sharp maximal function with applications to singular integrals*, Bull. Lond. Math. Soc. 42 (2010), no. 5, 843–856.

[49] Lerner A. K., *A simple proof of the A_2 conjecture*, Int. Math. Res. Not. IMRN (2013), no. 14, 3159–3170.

[50] Lerner A. K. , Lorist E. , and Ombrosi S. , Operator-free sparse domination, *Forum Math. Sigma* 10 (2022), Paper No. e15, 28.

[51] Lerner A. K., and Nazarov F., *Intuitive dyadic calculus: the basics*, Expo. Math. **37** (2019), no. 3, 225–265.

[52] Luiro H., *Continuity of the maximal operator in Sobolev spaces*, Proc. Amer. Math. Soc. 135 (2007), no. 1, 243–251.

[53] Merryfield K. G., *H^p Spaces in Poly-half Spaces*, ProQuest LLC, Ann Arbor, MI, 1980. Thesis (Ph.D.)–The University of Chicago.

[54] Miyachi, A., *H^p Spaces over Open Subsets of* $\mathbb{R}^n$, Studia Math. 95 (1990), no. 3, 205–228.

[55] Müller P. F. X., *Isomorphisms between H^1 spaces*, Instytut Matematyczny Polskiej Akademii Nauk. Monografie Matematyczne (New Series), vol. 66, Birkhäuser Verlag, Basel, 2005.

[56] Muscalu C., Pipher J., Tao T., and Thiele C., *Bi-parameter paraproducts*, Acta Math. 193 (2004), no. 2, 269–296.

[57] Muscalu C., Pipher J., Tao T., and Thiele C., *Multi-parameter paraproducts*, Rev. Mat. Iberoam. 22 (2006), no. 3, 963–976.

[58] Muscalu C., and Schlag W., *Classical and multilinear harmonic analysis. Vol. II*, Cambridge Studies in Advanced Mathematics, vol. 138, Cambridge University Press, Cambridge, 2013.

[59] Riesz F., *Über die Randwerte einer analytischen Funktion*, Math. Z. 18 (1923), no. 1, 87–95.

[60] Saari O., *Poincaré inequalities for the maximal function*, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5). 19 (2019), no. 3, 1065–1083

[61] Sarason D., *Functions of vanishing mean oscillation*, Trans. Amer. Math. Soc. 207 (1975), 391–405.

[62] Shaabani S., Extension domains for Hardy spaces, Studia Math. 272 (2023), no. 2, 139–158.

[63] Shaabani S., Maximal operators on BMO and slices, Canad. Math. Bull. 67 (2024), no. 1, 94–107.

[64] Shaabani S., *The Operator Norm of Paraproducts on Bi-parameter Hardy Spaces*, to appear in Studia Math., arXiv:2408.08366v2, 2025.

[65] Shaabani S., *The Operator Norm of Paraproducts on Hardy Spaces*, to appear in Math. Ann., arXiv:2402.13084v3, 2024.

[66] Stein, E. M. and Murphy, T. S., *Harmonic Analysis: Real-Variable Methods, Orthogonality, and Oscillatory Integrals*, Princeton Math. Series, vol. 43, Princeton University Press, 1993.

[67] Stein E. M., and Weiss G. L., On the theory of harmonic functions of several variables. I. The theory of H^p-spaces, Acta Math. 103 (1960), 25–62.

[68] Triebel, H., *Theory of function spaces*, Monographs in Mathematics, vol. 78, Birkhäuser Verlag, Basel, 1983.

[69] Uchiyama A., *The Construction of Certain BMO Functions and the Corona Problem*, Pacific J. Math. 99 (1982), no. 1, 183–204.

[70] Weisz F., *Martingale Hardy spaces and their applications in Fourier analysis*, Lecture Notes in Mathematics, vol. 1568, Springer-Verlag, Berlin, 1994.