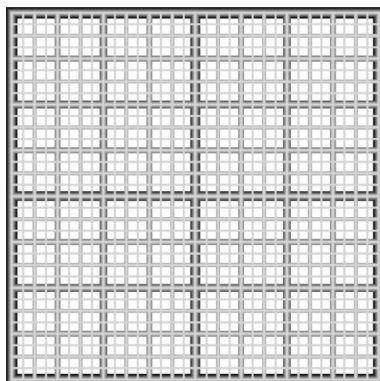




Sparse Domination and the A_2 Conjecture

Maria de Aguiar Sousa e Almeida Madrugo

November 2025

Thesis to obtain the Master of Science Degree in

Applied Mathematics and Computation

Instituto Superior Técnico

Supervisor(s): Prof. Jorge Filipe Drumond Pinto da Silva

Examination Committee

Chairperson: Luís Manuel Gonçalves Barreira

Supervisor: Jorge Filipe Drumond Pinto da Silva

Member of the Committee: Giuseppe Negro

Free at last from the charming,
dyadic prison of my own making¹.

As in the moral of this story:
saved by the stopping-time property.

¹See picture in front cover.

Declaration

I declare that this document is an original work of my own authorship and that it fulfills all the requirements of the Code of Conduct and Good Practices of the Universidade de Lisboa.

Resumo

O estudo de desigualdades com pesos em Análise Harmónica como as conhecemos hoje começou nos anos 70, com a definição dos chamados pesos A_p e a prova de desigualdades com pesos qualitativas para operadores de interesse, nomeadamente os operadores de Calderón-Zygmund. Mais recentemente, houve um interesse crescente nas desigualdades quantitativas correspondentes, isto é, na dependência destas desigualdades na característica A_p (uma noção específica de tamanho) do peso.

De particular interesse era a conjectura A_2 , que afirma que no caso de desigualdades com pesos fortes L^2 para operadores Calderón-Zygmund, a constante da desigualdade depende linearmente da característica A_2 . Em 2012, Hytönen publicou a primeira prova do agora Teorema A_2 , que em anos subsequentes foi reprovada de formas cada vez mais simples, até que Lerner publicou uma prova deste teorema em 2015 utilizando Dominação Esparsa Pontual - uma técnica que requer pouco mais do que as propriedades de grelhas diádicas e a decomposição de Calderón-Zygmund, e que se mostrou extremamente versátil com aplicabilidade incrivelmente alta: nos anos seguintes, muitos artigos foram publicados usando variantes desta técnica para provar desigualdades com pesos (sharp) para uma grande diversidade de operadores.

Nesta tese, apresentamos alguma da teoria clássica de pesos e desenvolvemos a teoria de Dominação Esparsa Pontual, culminando na prova da conjectura A_2 devido a Lerner. De seguida, mostramos alguma da versatilidade desta técnica, resumindo alguns dos resultados recentes nesta área.

Palavras-chave: Desigualdades com Pesos, Análise Harmónica Diádica, Dominação Esparsa, Decomposição de Calderón-Zygmund, Operadores Calderón-Zygmund.

Abstract

The study of weighted inequalities in Harmonic Analysis as we know them today started in the 1970's, with the definition of the so-called A_p weights and the proof of qualitative weighted bounds for operators of interest, namely the class of Calderón-Zygmund operators. In more recent years, there was a growing interest in the corresponding quantitative bounds, that is, on the dependence of these weighted bounds on the A_p characteristic (a specific measure of size) of the weight.

Of particular interest was the A_2 conjecture, which states that in the case of strong weighted L^2 bounds for Calderón-Zygmund operators, the dependence of the boundedness constant on the A_2 characteristic is linear. In 2012, Hytönen published the first proof of the now A_2 theorem, which in following years was reproved in increasingly simpler ways, until Lerner published a proof of this theorem in 2015 via Pointwise Sparse Domination - a technique that requires little more than the properties of dyadic grids and Calderón-Zygmund decomposition, and that turns out to be extremely versatile with incredibly high applicability: in following years a large body of work was published, using variants of this technique to prove (sharp) weighted inequalities for a wide range of operators.

In this thesis, we present some of the classical theory of weights and develop the theory of Pointwise Sparse Domination, culminating in Lerner's proof of the A_2 conjecture. We then show some of the versatility of this technique, surveying some of the recent results in the area.

Keywords: Weighted Inequalities, Dyadic Harmonic Analysis, Sparse Domination, Calderón-Zygmund Decomposition, Calderón-Zygmund Operators

Contents

Resumo	v
Abstract	vii
List of Tables	xi
List of Figures	xiii
Nomenclature	1
Glossary	1
1 A_p Weights: a Classical Approach	5
1.1 Introduction	5
1.2 The A_p Condition	5
1.2.1 Motivation for the characterization of A_p weights	6
1.2.2 A hint we are on the right track	7
1.2.3 Definition of A_p Weights	9
1.2.4 Some Important Properties	10
1.3 The A_1 case	13
1.4 A_p Weights Exist	16
1.4.1 Constructing A_1 weights	17
1.4.2 Constructing A_p weights	19
1.5 Strong Boundedness for A_p weights	20
1.5.1 A_∞ Weights	24
1.6 Important Weighted Results	25
1.6.1 Extrapolation	25
1.6.2 Calderón-Zygmund Operators	26
2 A Proof of the A_2 Conjecture via Pointwise Sparse Domination	33
2.1 Sparse and Dyadic Families of Sets	33
2.2 Strong Weighted Bounds for Sparse Forms	38
2.3 The Maximal Operator	44
2.3.1 From the Dyadic to the Continuous	44
2.3.2 Sparse Domination of the Dyadic Maximal Function	46
2.4 The A_2 conjecture	50

2.4.1	Reduction to a Cube	51
2.4.2	Reduction to the Iteration Step	52
2.4.3	Proving the Iteration Step	54
2.4.4	Proof of the A_2 Conjecture	60
3	The Sparse Revolution	63
3.1	Generalized Pointwise Domination	63
3.2	Bilinear Sparse Domination	66
3.3	Weak and Endpoint Bounds	70
3.4	Multiparameter Operators	73
	Bibliography	81

List of Tables

List of Figures

1	Approaches to the A_2 theorem by 2012, published in 2014 by Hytönen in [Hyt14]	2
1.1	In blue, the cube Q which contains x and has large average. In black, the 2^n relevant dyadic cubes which intersect Q . In pink, an example of a dilation of a dyadic cube which contains x .	9
1.2	An example of a sequence S_n of cubes containing x and contained in Q with diameter converging to zero.	14
1.3	In black, the cubes Q and $2Q$. In pink, a cube $\tilde{Q}$ which contains the point $y \in Q$ and has positive μ_2 -average, as well as its 5-dilation, which contains Q .	18
1.4	A visual representation of G_m (in drawn black) and G_{m+1} (in filled pink).	22
1.5	Example of a cube Q_i of centre c_i , with a point $x \in (2Q_i)^c$ and a point $y \in Q_i$.	29
2.1	Division of $\mathcal{D}_1$ into 3 disjoint sets of equal size (blue, pink and green), followed by $E(Q)$ for a cube Q of generation 1 in blue, $E(Q)$ for a cube of generation 2 in pink and $E(Q)$ for a cube of generation 3 in green.	34
2.2	Definition of Q_t .	36
2.3	Example of a third step of the iteration. The purple sets correspond exactly to Q_t , whereas the green $E^3(Q)$ were extended to account for the blue $E^3(Q)$, and then the pink $E^3(Q)$ were extended to account for the blue and green $E^3(Q)$.	36
2.4	In black the dyadic cubes containing x . In pink arbitrarily small cubes all such that the smallest cube which contains them is the second biggest black square.	44
2.5	In the first and second picture we have, side by side, the same section of a grid, with focus on two consecutive generations. Each color is associated to a different tripled dyadic grid: tripling every square of a given color gives us one of the 9 dyadic grids, with the one associated to the darker blue squares drawn in dark blue for reference. In the third and fourth picture, we focus on only one color, but add one more generation. In black is the original grid, and in blue the tripled grid associated to the blue squares.	45
2.6	A section of one-quadrant dyadic grid. In dark pink, the support of f , in thick black Q_0 and in thin black its ancestors. For the point y outside, the smallest dyadic square containing x and intersecting the support of f is indicated in light pink.	47

2.7	Illustration of the first two steps. In light green G_0 as defined above, the set of “good points”. The bad set B_1 in orange is then written as a disjoint union of dyadic squares. For each of which, we define $E(Q_{1,i})$ consisting of the “good points” in that square. We then define G_1 as the union of these new good points, and $B_1 = B_0 \setminus G_1$. In these steps, we added Q_0 and $Q_{1,1}, \dots, Q_{1,5}$ to S , with the corresponding E -sets represented in green. Note that these need not be rectangles and will usually in fact be more complicated sets.	48
2.8	In black, the cube Q which contains the support of f . In green, $Q_{1,1}, \dots, Q_{1,8}$ and in pink $Q_{2,1}, \dots, Q_{2,8}$, the first two iterations of the construction of the $Q_{i,j}$.	52
2.9	Visual representation of the proof structure of Proposition 2.4.3. Note that technically the Q_j should be elements of $\mathcal{D}(Q)$ and their area should be smaller, this picture is for illustrative effects only.	56
2.10	In the center, the cube Q which contains x and ξ , contained in the cube $3Q$. In the outside, the cube B_x with centre x and radius $2\text{diam}(Q)$, which contains $3Q$ and therefore Q , x and ξ as well.	59
3.1	Examples of possible $\mathcal{P}_0, \mathcal{P}_1, \mathcal{P}_2$: to obtain these sets, choose any point inside each of the blue, red and green squares, respectively.	76

Introduction

Historical Background

The study of weighted inequalities in Harmonic Analysis as we know them today started in the 1970's, with the definition of the so-called A_p weights² (also sometimes referred to in the literature as Muckenhoupt weights³). For $1 < p < \infty$, these weights ω are exactly the L^1_{loc} non-negative functions for which the maximal function is a bounded operator $L^p(\omega) \rightarrow L^p(\omega)$ and a bit of work shows that an equivalent characterization is that

$$[\omega]_p := \sup_{\text{cubes } Q} \left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{1}{|Q|} \int_Q \omega^{1-p'} \right)^{p-1}$$

is finite. This value $[\omega]_p$ is called the A_p characteristic of ω and it is in a way a measure of its size.

A lot has been known about weighted inequalities for A_p weights since their original study⁴:

- Calderón-Zygmund operators satisfy strong weighted L^p bounds when the weight is in A_p ($1 < p < \infty$), due to R. Coifman and C. Fefferman in [CF74];
- In the particular case of the Hilbert transform, ω being in A_p is necessary in order for the Hilbert transform to satisfy strong weighted bounds, as in the case for the maximal function (see [HMW73]);
- In 1984, Rubio de Francia proved an extrapolation theorem (see [Rub84]), which states that proving weighted bounds for any $\omega \in A_p$ immediately proves weighted bounds for weights in any other A_p class $A_{p'}$. Moreover, this extrapolation can be made quantitative, that is, in a way where we keep track of the dependence on the A_p characteristic of the weight.

In more recent years, there was a growing interest in quantitative versions (that is, with explicit dependence on the A_p characteristic $[\omega]_p$) of the strong weighted bounds above. Originally studied in [Buc93], these quantitative versions have proved to be helpful in the context of PDE, for example in [AIS01], where a linear dependence of the Beurling-Ahlfors operator weighted bound on the A_p characteristic is crucial to show sharp regularity results⁵. Of particular interest was the A_2 conjecture, stating that weighted strong L^2 bounds for Calderón-Zygmund operators are linear in the A_2 characteristic, for A_2 weights.

²It is worth noting that the A_p condition that follows had been previously studied by M. Rosenblum in [Ros62], even if in a different form.

³The characterization of A_1 weights, which we do not focus on for now, is due to B. Muckenhoupt ([Muc72]), who developed a good amount of the classical theory on A_p weights.

⁴These references come from [Duo01] and [Per19], where one can find more information and references on the history of weighted inequalities.

⁵Both these references are cited in [CMP12], whose introduction also contains pieces of the historical context.

Through the years, results for specific operators have been proved, a survey of which can be found in sources like the introduction of [Hyt12] or [Ler13a]. In the end, as seen in the timeline in [Ler13a], the ideas in these partial proofs culminated in a proof of the A_2 Conjecture - now A_2 Theorem - by Hytönen in his paper [Hyt12]. Interestingly, the story does not end there: a couple of years later, there were enough approaches to the A_2 theorem that Hytönen published a survey [Hyt14] going over all the theory studied and created in view of a proof of the A_2 conjecture, summed up in the following diagram:

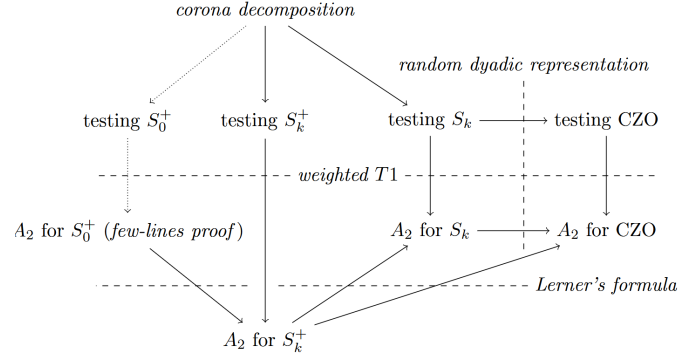


Figure 1: Approaches to the A_2 theorem by 2012, published in 2014 by Hytönen in [Hyt14]

Up to this point, the proof of the A_2 theorem was, even with all the simplifications, quite technical, and relied in a Haar Shift operator decomposition technique originally used to prove the result for the Hilbert transform (by S. Petermichl in [Pet07]) and then generalized to an arbitrary Calderón-Zygmund operator. However, a quick succession of papers gradually simplified the proof:

- Lerner proves that Calderón-Zygmund operators can be bounded by sparse forms in [Ler13b] (in a general Banach norm sense), which can be used to prove a good number of results, notably yielding a simplified proof of the A_2 conjecture ([Moe12]). However, this bound still relied on the representation of the operator in term of the Haar shift operators.
- In Lerner's "A Simple Proof of the A_2 Conjecture" ([Ler13a]), the role of the Haar Shift operator is replaced by the "local mean oscillation decomposition", a first step towards a more elementary proof.
- In [CR16], a duality condition crucial in [Ler13b] is lifted by a sparse bound (of dyadic shifts), which remarkably is a pointwise bound instead of a bound in norm. Such pointwise bounds were independently established in [LN19] as well.
- In Lacey's "An elementary proof of the A_2 bound" ([Lac17]), inspired by the pointwise bounds, the mean oscillation inequality used in [Ler13a] is replaced by exploiting the behavior of dyadic grids, as well as weak bounds for the maximal truncations of the operators.
- Finally, in [Ler16] Lerner modified Lacey's proof from [Lac17] into a shorter and elementary proof of A_2 via pointwise sparse domination. This proof, with some of the improvements from the later paper [LO20] by Lerner and Ombrosi ("Some remarks on the pointwise sparse domination"), is arguably the simplest proof of the A_2 conjecture to date, and will be the main focus of this thesis.

Even more remarkable than the body of work produced in search of a proof of the A_2 conjecture, it turns out that the most elementary proof via Sparse Domination is extremely powerful and versatile. In the following years, varied types of Sparse Domination were used to prove weighted inequalities for a very diverse collection of operators, in what Pereyra coined in [Per19] as the Sparse Revolution. In this document, Pereyra surveys the role of dyadic harmonic analysis in the proof of weighted inequalities, detailing the path outlined above and referencing some of the later work it inspired.

Sparse Domination

The Sparse Domination technique, used in [Ler16] to prove the A_2 conjecture, consists on bounding varied operators, notably Calderón-Zygmund operators, by Sparse Forms,

$$\mathcal{A}_S f(x) := \sum_{Q \in S} \langle f \rangle_Q \chi_Q(x), \quad \text{where } \langle f \rangle_Q = \frac{1}{|Q|} \int_Q |f|.$$

with S a sparse (that is, in a sense to be formalized later, “almost disjoint”) family of cubes, in order to prove certain boundedness properties of our operators of interest.

These sparse operators are quite simple, as they are just (infinite) linear combinations of characteristic functions, indexed in a sparse family of sets. The reason why they are so helpful to prove sparse bounds is that not only do they have the strong weighted bounds we intend to prove, but we can also prove an explicit dependence on the A_p constant, which turns out to be sharp for many of the operators we are interesting in bounding:

Proposition: For every sparse family S , every $\omega \in A_p$ and every $f \in L^p(\omega)$, we have the bound

$$\|\mathcal{A}_S f\|_{L^p(\omega)} \lesssim [\omega]_p^{\max(1, \frac{1}{p-1})} \|f\|_{L^p(\omega)}$$

Essentially, the fact that we have (infinite) linear combinations of characteristic functions allows us to use duality and introduce integrals over these Q of our test functions. Then, a clever use of bounds for the weighted maximal function and the Hölder inequality makes integrals of ω over the cubes Q pop up, which - almost by sorcery - can be rearranged into exactly the A_p characteristic. That is, the dependence on the A_p characteristic that we so desire can be found for these operators using truly elementary methods, which is the big advantage of Sparse Forms.

As a result of these weighted bounds for Sparse Forms with explicit dependence on the A_p characteristic, in order to prove (sharp) weighted bounds for other operators it is enough to show that these are bounded in some way by Sparse Forms. In the case of the A_2 conjecture, this is done by proving the following pointwise domination result:

Proposition: Let T be a Calderón-Zygmund operator and f be a compactly supported L^1 function. Then, there is a sparse family $S = S(f)$ such that

$$|Tf| \leq C \sum_{\tilde{Q} \in S} \langle f \rangle_{\tilde{Q}, p} \chi_{\tilde{Q}} \quad \text{a.e.}$$

The proof of this proposition is the main content of this thesis, and consists of a clever use of Calderón-Zygmund decomposition methods and weak inequalities for certain operators. The more general Sparse Domination results arise from proving similar results for different operators, generally with different types of domination, such as dominating the bilinear form $\langle Tf, g \rangle$ used in proofs by duality rather than proving a pointwise bound.

Outline

In the first chapter of the thesis, we present the classical theory of A_p weights, including some of the important results referenced above. We lay the groundwork for the work done with A_p weights in the following chapters, as well as highlight the challenges of attempting to apply classical methods to keep track of the dependence on the A_p characteristic in weighted inequalities.

In the second chapter, we begin by introducing the notion of sparse families and forms, and we prove strong weighted results for these operators, with explicit dependence on the A_p constant. We then prove that the maximal function is pointwise bounded by sparse forms, to showcase the main ideas of the proof of the A_2 conjecture in a simpler case, so that the technical details do not obscure the essence of the proof. Finally, we prove the A_2 theorem, in a very similar way to Lerner's proof from [Ler16].

In the third chapter, we show some of the versatility of Sparse Domination, both via the generalized pointwise theorem in [LO20] and the very fruitful bilinear formulation of Sparse Bounds. We also address endpoint and weak bounds and discuss a negative result for multiparameter Sparse Domination.

Throughout this work, we assume familiarity with maximal operators and their bounds, as well as with Calderón-Zygmund operators and Calderón-Zygmund decomposition. The reader can find such material and all results used in textbooks such as [Duo01] or [Gra14]. We also work with dyadic grids beyond the standard dyadic grid, and technical proofs about them are omitted. We refer the reader to [LN19] for such proofs and an overview of dyadic calculus, including dyadic grids and the associated notation.

Chapter 1

A_p Weights: a Classical Approach

1.1 Introduction

The theory of A_p weights deals with the study of integrals over weighted measures,

$$\int f(x)\omega(x)dx,$$

where ω is a non-negative measurable function, which we call a weight. In particular, A_p weights are, in a specific harmonic analytical sense, the class of “well-behaved” weights.

There are a few things that can be said about why one would be interested in weights. The first is just conceptual: we often like to stretch out our theories to the most general context possible, without losing the essential properties at play. Measures defined by weights are more general than the Lebesgue measure while still being much more concrete than arbitrary measures, so they seem like a good enough space to test out possible generalizations.

The second one is much more practical: weights are useful. PDE theory is a prime example, in which weights arise naturally. For instance, Sobolev spaces, defined via Fourier transform, heavily depend on weights. Indeed, we say that $f \in H^s$ if $f \in S'$ and $\hat{f}(\xi)(1 + |\xi|^2)^{\frac{s}{2}} \in L^2$. This means that the Sobolev norm of f can be seen as a weighted L^2 norm of its Fourier transform, with $\omega = (1 + |\xi|^2)^s$. As another example, in quantum mechanics we work with states described by L^2 functions such that the position and moment operators are well-defined. These are given respectively by $xf \in L^2$ and $\partial f = (i\xi\hat{f})^\vee \in L^2$, and can similarly be seen as weighted integrals. If we hope to establish bounds for Sobolev norms or for the position and moment operators, a reasonable way to go at it is simply to try to generalize the theory we have already established over the Lebesgue measure into the more general setting of weighted integrals.

1.2 The A_p Condition

Before introducing A_p weights, we begin by defining weights and weighted integrals:

Definition 1.2.1: A weight ω is a non-negative measurable function. The ω -weighted integral of f is defined as $\int f\omega := \int f(x)\omega(x)dx$.

In Harmonic Analysis, if we want to build a theory on “nice” weights, ensuring maximal operators have the expected $L^p(\omega)$ boundedness properties is a natural place to start. Indeed, the Hardy-Littlewood maximal operator, while being a relatively simple but fundamental operator, mimics the bounds we have come to expect from plenty of operators of interest, such as Calderón-Zygmund operators, namely weak L^1 bounds and strong L^p bounds for $p > 1$. For the purposes of this chapter, we will be using the Hardy-Littlewood maximal operator over uncentered cubes (with sides parallel to the axis). That is,

$$Mf(x) := \sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f(y)| dy.$$

With this in mind, let us then assume ω is a weight such that for any $f \in L^p(\omega)$, we have

$$\|Mf\|_{L^p(\omega)} \leq C\|f\|_{L^p(\omega)}, \quad 1 < p < \infty,$$

$$\omega(\{x : Mf(x) > \lambda\}) \leq \frac{\|f\|_{L^1(\omega)}}{\lambda}, \quad p = 1.$$

that is, such that the maximal function is bounded from $L^p(\omega)$ to $L^p(\omega)$, for $1 < p < \infty$ or from $L^1(\omega)$ to $L^{1,\infty}(\omega)$. In this subsection, we will deduce a nice characterization for all such weights, which constitutes the definition of A_p (or Muckenhoupt) weights.

1.2.1 Motivation for the characterization of A_p weights

Let us assume we have a strong (p, p) estimate, that is, let us assume $1 < p < \infty$ and that for every measurable f we have:

$$\|Mf\|_{L^p(\omega)} \lesssim \|f\|_{L^p(\omega)} \iff \left(\int |Mf|^p \omega \right)^{\frac{1}{p}} \lesssim \left(\int |f|^p \omega \right)^{\frac{1}{p}} \iff \int |Mf|^p \omega \lesssim \int |f|^p \omega.$$

We wish to characterize the ω for which we have such an estimate. For now, let us assume that everything is locally integrable, in order to simplify the deduction. Let $f \geq 0$. If we fix a cube Q , then $Mf(x) \geq \frac{1}{|Q|} \int_Q f$, for every $x \in Q$, which implies

$$\int f^p \omega \gtrsim \int Mf^p \omega \geq \int_Q Mf^p \omega \geq \int_Q \left(\frac{1}{|Q|} \int_Q f \right)^p \omega = \left(\frac{1}{|Q|} \int_Q f \right)^p \left(\int_Q \omega \right).$$

It would be convenient for the integral of $|f|^p \omega$ to also be taken over Q , for the sake of consistency. This is easily achievable: note that none of the bounds above would change in the slightest if we took $f\chi_Q$ instead, since $Mf\chi_Q(x)$ is still greater or equal to $\frac{1}{|Q|} \int_Q f$ for $x \in Q$:

$$\begin{aligned} \int_Q f^p \omega &= \int (f\chi_Q)^p \omega \gtrsim \int M(f\chi_Q)^p \omega \geq \int_Q M(f\chi_Q)^p \omega \geq \int_Q \left(\frac{1}{|Q|} \int_Q f \right)^p \omega \\ &= \left(\frac{1}{|Q|} \int_Q f \right)^p \left(\int_Q \omega \right). \end{aligned} \tag{1.1}$$

At this point, we can choose a convenient f in order to simplify the expression. In fact, note that if we choose $f = \omega^\alpha$, then the expression above becomes dependent on only Q and ω . Moreover, note that a possible way to simplify the expression, asking that $f^p \omega$ and f be the same, gets us a function of this form, as $f^p \omega = f \iff \omega = f^{1-p} \iff f = \omega^{\frac{1}{1-p}} = \omega^{1-p'}$. This gives us, distributing the $\frac{1}{|Q|}$ factors in the last step,

$$\left(\frac{1}{|Q|} \int_Q \omega^{1-p'} \right)^p \left(\int_Q \omega \right) \leq C \int_Q \omega^{1-p'} \iff \left(\frac{1}{|Q|} \int_Q \omega^{1-p'} \right)^{p-1} \left(\frac{1}{|Q|} \int_Q \omega \right) \leq C.$$

Since Q is arbitrary, we have in fact proven (for locally integrable $\omega, \omega^{1-p'}$) that

$$\sup_Q \left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{1}{|Q|} \int_Q \omega^{1-p'} \right)^{p-1} \leq C. \quad (1.2)$$

1.2.2 A hint we are on the right track

Before proving that (1.2) holds for all weights (through an approximation argument), we prove here that we can recover the *weak* boundedness of the maximal operator from (1.2). This is the first indicator that this property is exactly what we are looking for.

We begin with a lemma which exploits the Hölder numerology of condition (1.2). Among other results, it allows us to control the ω -measure of dilations of cubes, which will be very important in what follows.

Lemma 1.2.2: Let ω satisfy (1.2), for $1 < p < \infty$. Then, we have, for every cube Q and every measurable f ,

$$\left(\frac{1}{|Q|} \int_Q |f| \right)^p \leq C \frac{1}{\omega(Q)} \left(\int_Q |f|^p \omega \right), \quad (1.3)$$

with C independent of Q . In particular, ω satisfies, for every measurable $S \subseteq Q$ and the same constant C

$$\left(\frac{|S|}{|Q|} \right)^p \leq C \frac{\omega(S)}{\omega(Q)} \quad \text{and} \quad \omega(\alpha Q) \leq \alpha^{np} C \omega(Q), \text{ for } \alpha \geq 1. \quad (1.4)$$

Proof: The first part of the lemma can be proved by multiplying and dividing by $\omega^{\frac{1}{p}}$ in order to make the weighted L^p average appear, followed by applying Hölder's inequality ($p, p' = \frac{p}{p-1}$) and using the (1.2) bound as follows:

$$\begin{aligned} \left(\frac{1}{|Q|} \int_Q |f| \right)^p &= \frac{1}{|Q|^p} \left(\int_Q |f| \omega^{\frac{1}{p}} \omega^{-\frac{1}{p}} \right)^p \leq \frac{1}{|Q|^p} \left(\int_Q |f|^p \omega \right) \left(\int_Q \omega^{-\frac{1}{p} \frac{p}{p-1}} \right)^{p \frac{p-1}{p}} \\ &= \left(\frac{1}{|Q|} \int_Q |f|^p \omega \right) \left(\frac{1}{|Q|} \int_Q \omega^{1-p'} \right)^{p-1} \\ &\leq C \left(\frac{1}{|Q|} \int_Q |f|^p \omega \right) \frac{|Q|}{\omega(Q)} \\ &= C \frac{1}{\omega(Q)} \left(\int_Q |f|^p \omega \right). \end{aligned}$$

From here, we can take $f = \chi_S$, with $S \subseteq Q$. Then, the left hand side becomes $\left(\frac{|S|}{|Q|} \right)^p$ and the right

hand side becomes $C \frac{\omega(S)}{\omega(Q)}$ which gives us

$$\left(\frac{|S|}{|Q|} \right)^p \leq C \frac{\omega(S)}{\omega(Q)}.$$

Finally, by taking this formula and replacing Q by αQ (which is also a cube) and S by Q , we get

$$\left(\frac{|Q|}{|\alpha Q|} \right)^p \leq C \frac{\omega(Q)}{\omega(\alpha Q)} \implies \omega(\alpha Q) \leq C \alpha^{np} \omega(Q),$$

as intended. □

Proposition 1.2.3: Let ω be a weight such that $\sup_Q \left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{1}{|Q|} \int_Q \omega^{1-p'} \right)^{p-1} \leq C$, with $1 < p < \infty$. Then, the maximal operator is bounded $L^p(\omega) \rightarrow L^{p,\infty}(\omega)$.

Proof: This proof is an adaptation of standard arguments. Firstly, we reduce this case to the one dealing with the dyadic maximal function: this allows us to write our superlevel set as a disjoint union of cubes. This argument carries over from the Lebesgue measure to the ω measure essentially unchanged (we will nevertheless see it shortly). Then, where we would typically use that the averages over these cubes are greater than λ to bound $|Q|$ by an integral of f (which doesn't work in this case because the maximal function uses the Lebesgue measure rather than ω), we instead use (1.3). In the process, the second property in (1.4) will be key in recovering a disjoint collection of cubes from their dilations, as it often is in these arguments.

So, we begin by showing that we can reduce the problem to using the dyadic maximal function. Let x be a point in $\{x : Mf(x) > \lambda\}$ and let Q be a cube containing x over which the average is greater than λ . Then, Q intersects at most 2^n dyadic cubes Q_j such that $\ell(Q) \leq \ell(Q_j) < 2\ell(Q)$, where $\ell(Q)$ is the length of the side of a cube Q (see figure 1.1). We have that $\int_Q |f| = \sum_j \int_{Q_j \cap Q} |f|$, which means that if $\frac{1}{|Q|} \int_Q |f| > \lambda$, then there is some dyadic Q' such that $\frac{1}{|Q'|} \int_{Q' \cap Q} |f| > 2^{-n} \lambda$, which in particular means Q and Q' intersect. Then, we have

$$\frac{1}{|Q'|} \int_{Q'} |f| \geq \frac{1}{|Q'|} \int_{Q' \cap Q} |f| = \frac{|Q|}{|Q'|} \frac{1}{|Q|} \int_{Q' \cap Q} |f| > 2^{-n} 2^{-n} \lambda = 4^{-n} \lambda,$$

where we used $2\ell(Q) > \ell(Q')$ to conclude $|Q| > 2^{-n} |Q'|$.

So, we were able to show that we can always find a close dyadic cube with average greater than $4^{-n} \lambda$. This is not immediately useful, as Q' might not contain x . However, we can dilate Q' in order to always contain x . Indeed, since Q' intersects Q and $\ell(Q) < \ell(Q')$, extending Q' by $\ell(Q')$ in every axial direction ensures it contains Q , and so we will be considering this new cube $3Q'$ instead:

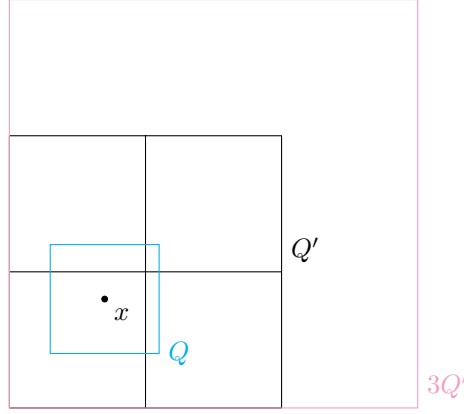


Figure 1.1: In blue, the cube Q which contains x and has large average. In black, the 2^n relevant dyadic cubes which intersect Q . In pink, an example of a dilation of a dyadic cube which contains x .

Let us use a Calderón-Zygmund decomposition at height $4^{-n}\lambda$ to write $\{x : M_d f > 4^{-n}\lambda\} = \cup Q_i$, with Q_i disjoint dyadic cubes over which the average of f is greater than $4^{-n}\lambda$. Then, the argument above showed

$$\{x : Mf > \lambda\} \subseteq 3\{x : M_d f > 4^{-n}\lambda\},$$

where $3\{x : M_d f > 4^{-n}\lambda\}$ is understood to mean $\cup 3Q_i$. Using this, we have

$$\omega(\{x : Mf > \lambda\}) \leq \omega(\cup 3Q_i) \leq \sum_i \omega(3Q_i) \leq C_{n,p} \sum_i \omega(Q_i),$$

where we used (1.4) to get rid of the dilations. Now, using (1.3), we get (using our assumption on the averages of f over Q_j)

$$\begin{aligned} \sum_i \omega(Q_i) &\lesssim \sum_i \left(\int_{Q_i} |f|^p \omega \right) \left(\frac{1}{|Q_i|} \int_{Q_i} |f| \right)^{-p} \\ &\lesssim \sum_i \left(\int_{Q_i} |f|^p \omega \right) (4^{-n}\lambda)^{-p} \\ &\lesssim_{n,p} \sum_i \frac{\int_{Q_i} |f|^p \omega}{\lambda^p} \\ &\lesssim_{n,p} \frac{\int_{\mathbb{R}^n} |f|^p \omega}{\lambda^p}, \end{aligned}$$

□

and so we have established the weak bound.

1.2.3 Definition of A_p Weights

In section 1.2.2, we proved we can recover weak bounds for the maximal operator from condition (1.2). As it turns out, this can be bumped up to strong bounds, as we will see in the next section. That leads us to define A_p weights as follows:

Definition 1.2.4 (A_p weight): A weight ω is called an A_p weight, for $1 < p < \infty$, if it satisfies

$$[\omega]_p := \sup_Q \left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{1}{|Q|} \int_Q \omega^{1-p'} \right)^{p-1} < \infty.$$

The quantity $[\omega]_p$ is referred to as the A_p weight characteristic of ω .

Before proceeding, we will prove that the strong bound for the maximal operator implies the A_p condition, for general ω . It follows exactly like the deduction above, we just need to be careful with local integrability.

Proposition 1.2.5: Let ω be a weight such that the maximal operator is bounded $L^p(\omega) \rightarrow L^p(\omega)$. Then, ω is an A_p weight.

Proof: Throughout this proof, we assume $f \geq 0$, as the proof for general f consists of replacing f by $|f|$ everywhere. Now, we can again use the fact that $Mf\chi_Q(x) \geq \frac{1}{|Q|} \int_Q f$ for $x \in Q$ to deduce (as in (1.1))

$$C \int_Q f^p \omega \geq \left(\frac{1}{|Q|} \int_Q f \right)^p \left(\int_Q \omega \right).$$

Ideally, we would choose $f = \omega^{1-p'}$, but we have no reason to believe it is locally integrable. So, we do the next best thing and take f to be a locally integrable version of $\omega^{1-p'}$. Indeed, let $Q_N = \{x : \omega^{1-p'} \leq N\}$. Then, $\omega^{1-p'} \chi_{Q_N}$ is locally integrable (its integral is bounded by $N|Q_N| \leq N|Q|$). Choosing $f = \omega^{1-p'} \chi_{Q_N}$, we get

$$\int_{Q_N} \omega^{1-p'} \gtrsim \left(\frac{1}{|Q|} \int_{Q_N} \omega^{1-p'} \right)^p \left(\int_Q \omega \right) \iff C \gtrsim \left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{1}{|Q|} \int_{Q_N} \omega^{1-p'} \right)^{p-1}.$$

Now, we can take the limit in N using the monotone convergence theorem on $\omega^{1-p'} \chi_{Q_N}$, recovering the A_p condition. $\square$

1.2.4 Some Important Properties

In the general setting, a weight ω is defined as an arbitrary non-negative measurable function, which means it can be 0 in a significant region, or fail to be locally integrable. As it turns out, the A_p condition heavily restricts the behavior of ω and prohibits both these cases from happening:

Proposition 1.2.6: Let ω be a weight which is not identically 0 nor identically ∞ and satisfies, for every cube Q and measurable $S \subseteq Q$, $\omega(Q) \cdot \left(\frac{|S|}{|Q|} \right)^p \leq C\omega(S)$, for some $1 < p < \infty$. Then

- a) The weight ω is non-zero almost everywhere in the Lebesgue measure;
- b) The weight ω is locally integrable.

Proof:

a) Let $\tilde{S} = \{x : \omega(x) > 0\}$ and let $S \subseteq \tilde{S}$ have positive measure and be contained in a cube $\tilde{Q}$. We can assume S exists because $\tilde{S} \cap [-N, N]^n$ is a sequence of sets whose measures converge to the (positive) measure of $\tilde{S}$, and therefore there must be at least one $\tilde{S} \cap [-N, N]^n$ with positive measure.

Then, for every $Q \supseteq S$, we would have $\frac{\omega(Q)}{|Q|^p} \leq C \frac{\omega(S)}{|S|^p} = 0$, and so we would have $\omega(Q) = 0$ for arbitrarily large cubes (for example $[-M, M]^n$ for $M > N$), which would contradict ω not being identically zero.

b) Similarly, let $\tilde{S} = \{x : \omega(x) < \infty\}$ have positive measure. By the same reasoning as above, there is some $S' = \{x : \omega(x) < C\}$ with positive measure. Let $S \subseteq S'$ have positive measure and be contained in some $[-N, N]^n$. Note that we have $\omega(S) \leq C(2N)^n < \infty$.

Then, if there was a cube Q such that $\omega(Q) = \infty$, then any $\tilde{Q} \supseteq Q$ would satisfy $\omega(\tilde{Q}) = \infty$ as well. Let M be such that $Q \subseteq [-M, M]^n$ and let us consider $\tilde{Q} = [-\max(M, N), \max(M, N)]^n$. Then (since $Q \subseteq \tilde{Q}$), $\omega(\tilde{Q}) = \infty$ and (since $S \subseteq \tilde{Q}$) we would have

$$\infty = \frac{\omega(\tilde{Q})}{|\tilde{Q}|^p} \leq C \frac{\omega(S)}{|S|^p} < \infty,$$

which is a contradiction. □

Corollary 1.2.7: A_p weights are locally integrable and non-zero almost everywhere in the Lebesgue measure.

Proof: Lemma 1.2.2 ensures that the hypothesis of proposition 1.2.6 holds. □

Another very important property of A_p weights is how they relate to each other. On the one hand, we are able to use L^p duality to relate an A_p to $A_{p'}$, with p and p' dual exponents. On the other hand, we can prove that an A_p weight is also an A_q weight, for $q > p$. That is our next proposition.

Proposition 1.2.8: Let $1 < p < \infty$.

1. We have that $\omega \in A_p$ if and only $\omega^{1-p'} \in A_{p'}$.
2. Let ω be an A_p weight and $q > p$. Then, ω is an A_q weight.

Proof:

1. Since ω is an A_p weight, we have that

$$\sup_Q \left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{1}{|Q|} \int_Q \omega^{1-p'} \right)^{p-1} < \infty.$$

In order to prove that $\omega^{1-p'}$ is in $A_{p'}$, our first step is to take this condition and make $\omega^{1-p'}$ appear:

$$\left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{1}{|Q|} \int_Q \omega^{1-p'} \right)^{p-1} = \left(\frac{1}{|Q|} \int_Q (\omega^{1-p'})^{\frac{1}{1-p'}} \right) \left(\frac{1}{|Q|} \int_Q (\omega^{1-p'}) \right)^{p-1}.$$

Now, recall that we have $\frac{1}{p} + \frac{1}{p'} = 1 \iff 1 + \frac{p}{p'} = p \iff 1 - p = -\frac{p}{p'}$ and similarly $1 - p' = -\frac{p'}{p}$. So, $\frac{1}{1-p'} = 1 - p$. That means that the first integrand reduces to $(\omega^{1-p'})^{1-p}$. So, we get (using $(p-1)(p'-1) = 1$):

$$\begin{aligned} \left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{1}{|Q|} \int_Q \omega^{1-p'} \right)^{p-1} &= \left(\frac{1}{|Q|} \int_Q (\omega^{1-p'})^{1-p} \right)^{(p-1)(p'-1)} \left(\frac{1}{|Q|} \int_Q (\omega^{1-p'}) \right)^{p-1} \iff \\ &\iff \left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{1}{|Q|} \int_Q \omega^{1-p'} \right)^{p-1} = \left(\left(\frac{1}{|Q|} \int_Q (\omega^{1-p'})^{1-p} \right)^{p'-1} \left(\frac{1}{|Q|} \int_Q (\omega^{1-p'}) \right) \right)^{p-1}. \end{aligned}$$

From here, the result follows: the supremum of the left hand side is finite (A_p condition for ω) if and only if the supremum of the right hand side is finite, which happens if and only if the supremum of the $\frac{1}{p-1}$ power of the right hand side is finite ($A_{p'}$ condition for $\omega^{1-p'}$).

2. Since $\omega \in A_p$, we know that

$$\sup_Q \left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{1}{|Q|} \int_Q \omega^{1-p'} \right)^{p-1} < \infty.$$

We want to show that given any $q > p$ we have

$$\sup_Q \left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{1}{|Q|} \int_Q \omega^{1-q'} \right)^{q-1} < \infty.$$

The first term is already on the right track. For the second one, we want to use Hölder, raising the integrand to the power $\frac{1-p'}{1-q'}$ and χ_Q to the dual power. We get:

$$\begin{aligned} \left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{1}{|Q|} \int_Q \omega^{1-q'} \right)^{q-1} &\leq \left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{1}{|Q|} \left(\int_Q \omega^{1-p'} \right)^{\frac{1-q'}{1-p'}} \left(|Q|^{1-\frac{1-q'}{1-p'}} \right) \right)^{q-1} = \\ &= \left(\frac{1}{|Q|} \int_Q \omega \right) \left(\left(\frac{1}{|Q|} \int_Q \omega^{1-p'} \right)^{\frac{1-q'}{1-p'}} \right)^{q-1} = \left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{1}{|Q|} \int_Q \omega^{1-p'} \right)^{p-1}, \end{aligned}$$

where we used $(q-1)(1-q') = -1$ and $p-1 = -\frac{1}{1-p'}$. Taking the supremum over all Q and using the fact that $\omega \in A_p$ now proves that ω satisfies the A_q condition as well and the proof is done. $\square$

Remark 1.2.9: Note that one can extract quantitative bounds from the proofs above:

1. $[\omega]_p = [\omega^{1-p'}]_{p'}^{p-1}$;
2. $[\omega]_q \leq [\omega]_p$.

In the previous section, it was established (in lemma 1.2.2) that A_p weights satisfy the property $\omega(\alpha Q) \leq C\alpha^{np}\omega(Q)$, for any $\alpha \geq 1$, which is typically very useful in arguments which require doubling or tripling cubes, such as in proposition 1.2.3. As it turns out, that can be improved to cover cases in which $|S| \leq \alpha|Q|$, with $\alpha \geq 1$, as long we accept losing some control over the constant:

Lemma 1.2.10: Let $\omega \in A_p$, with $1 < p < \infty$. Then, given any $0 < \alpha < 1$, there exists some $0 < \beta < 1$ such that whenever $|S| \leq \alpha|Q|$ for $S \subseteq Q$ we have $\omega(S) \leq \beta\omega(Q)$.

Proof: Recall that by lemma 1.2.2 we have an inequality involving the quantities at play, namely

$$\left(\frac{|S|}{|Q|}\right)^p \leq C \frac{\omega(S)}{\omega(Q)}.$$

However, it seems to be working in the opposite direction, which motivates us to look at $Q \setminus S$, as we have $\frac{|Q \setminus S|}{|Q|} = \frac{|Q| - |S|}{|Q|} = 1 - \frac{|S|}{|Q|}$ and $\frac{\omega(Q \setminus S)}{\omega(Q)} = 1 - \frac{\omega(S)}{\omega(Q)}$, which flips the sign. So, using the above inequality for $Q \setminus S$, we get

$$\left(1 - \frac{|S|}{|Q|}\right)^p \leq C \left(1 - \frac{\omega(S)}{\omega(Q)}\right).$$

Since $|S| \leq \alpha|Q| \implies \frac{|S|}{|Q|} \leq \alpha$, we have $(1 - \alpha)^p \leq \left(1 - \frac{|S|}{|Q|}\right)^p$ and so

$$(1 - \alpha)^p \leq C \left(1 - \frac{\omega(S)}{\omega(Q)}\right) \implies \omega(S) \leq \frac{C - (1 - \alpha)^p}{C} \omega(Q),$$

which means that the result follows by taking $\beta = \frac{C - (1 - \alpha)^p}{C}$, as long as this number is in $(0, 1)$. Since $\alpha < 1$, $C - (1 - \alpha)^p < C$ and so $\beta < 1$. The fact that $\beta > 0$ follows from the inequality it satisfies, as $\beta \leq 0$ would imply $\omega(S) \leq 0$ (note that we can assume without loss of generality that $\omega(S) > 0$, since otherwise the inequality $\omega(S) \leq \beta\omega(Q)$ always holds). $\square$

1.3 The A_1 case

When $p = 1$, weak bounds are typically all we can prove. In this section, we treat this endpoint case separately in order to reach our definition of an A_1 weight. In the process, we will adapt the proof to show that weak bounds are enough to imply the A_p condition, even for $1 < p < \infty$.

So, let ω be a weight such that the maximal function is bounded from $L^1(\omega)$ to $L^{1,\infty}(\omega)$ and let us fix a function $f \geq 0$ (we can assume this without loss of generality, as neither M or $\|\cdot\|_{L^1(\omega)}$ see the sign of f). By hypothesis, we have

$$\omega(\{x : Mf > \lambda\}) \leq C \frac{\|f\|_{L^1(\omega)}}{\lambda}.$$

Our gameplay here will be to fix a cube Q over which $\int_Q f \neq 0$ and try to extract information from the weak bound. The three most important ingredients are the following:

- For any $\lambda < \frac{1}{|Q|} \int_Q f$, we have that $Mf > \lambda$ for every point in Q , which allows us to bound $\omega(Q) \leq \omega(\{x : Mf > \lambda\})$ using the weak bound.
- As long as we ensure that $\lambda < \frac{1}{|Q|} \int_Q f$, we have the freedom to choose λ . In particular, we can fix an arbitrary $\lambda < \frac{1}{|Q|} \int_Q f$ and “in the limit” replace λ by this known quantity.
- Just like in the $1 < p < \infty$ case, our focus on a specific Q allows us to consider $f\chi_Q$ instead of f , while keeping the two other ingredients essentially unchanged. This has the advantage of forcing another average over Q on the right-hand side, as long as we compensate the Q term.

Putting these together, let us take $f\chi_Q$, for Q such that $\int_Q f > 0$, and $\lambda < \frac{1}{|Q|} \int_Q f$. Then, since $Q \subseteq \{x : Mf\chi_Q > \lambda\}$, we have

$$\omega(Q) \leq C \frac{\int_Q f \omega}{\lambda} \iff \lambda \omega(Q) \leq C \int_Q f \omega \xrightarrow{\lambda \rightarrow \frac{1}{|Q|} \int_Q f} \frac{1}{|Q|} \left(\int_Q f \right) \omega(Q) \leq C \int_Q f \omega.$$

Choosing $f = \chi_S$, with $S \subseteq Q$ of positive Lebesgue measure, allows us to get an expression depending only on ω , relating the ω measure and the Lebesgue measure of these sets:

$$\frac{\omega(Q)}{|Q|} |S| \leq C \omega(S) \iff \frac{\omega(Q)}{|Q|} \leq C \frac{\omega(S)}{|S|}. \quad (1.5)$$

Note that this expression is equivalent (when $\omega(Q) \neq 0$) to $\frac{|S|}{|Q|} \leq C \frac{\omega(S)}{\omega(Q)}$, which allows us to use Lemma 1.2.6 to deduce that ω is non-zero almost-everywhere, as well as locally integrable.

From expression 1.5, since S is an arbitrary set, we should be able to recover $\omega(x)$ on the right-hand side for almost everywhere x , due to the Lebesgue Differentiation Theorem. Then, we would take the sup on the left side and get $M\omega \leq C\omega$. Indeed, given any point x , for every $Q \ni x$ we can find a sequence of cubes S_n such that $x \in S_n \subseteq Q$ and that their diameter goes to zero (see figure 1.3).

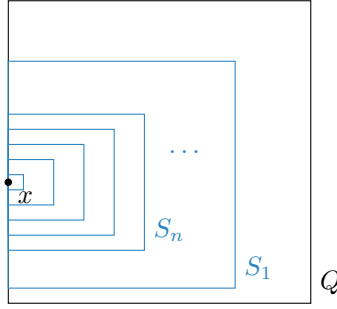


Figure 1.2: An example of a sequence S_n of cubes containing x and contained in Q with diameter converging to zero.

Using the fact that ω is locally integrable, we have that Lebesgue Differentiation Theorem applies to ω , so we can apply the reasoning above, and we get (almost everywhere in $x \in Q$)

$$\forall_{Q \ni x} \frac{\omega(Q)}{|Q|} \leq C \omega(x) \iff M\omega(x) \leq C \omega(x) \text{ a.e.}, \quad (1.6)$$

which we will refer to as the A_1 condition.

Definition 1.3.1: We say that a non-negative weight function ω is an A_1 weight if it satisfies the A_1 condition $M\omega(x) \leq C\omega(x)$ a.e..

We define the A_1 characteristic of ω as $[\omega]_1 := \inf\{C : M\omega(x) \leq C\omega(x) \text{ a.e.}\}$.

An interesting fact about this deduction is that it can be used to prove that weak L^p bounds are enough to imply that a weight is A_p . That is our next proposition:

Proposition 1.3.2: Let ω be a weight such that the maximal operator is bounded $L^p(\omega) \rightarrow L^{p,\infty}(\omega)$. Then, it is an A_p weight.

Proof: By hypothesis, ω satisfies, for every $f \in L^p(\omega)$,

$$\omega(\{x : Mf(x) > \lambda\}) \leq C \frac{\|f\|_{L^p(\omega)}^p}{\lambda^p}.$$

Proceeding as above, we fix Q such that $\int_Q f \neq 0$ and consider an arbitrary $\lambda < \frac{1}{|Q|} \int_Q f$. Then, we have, taking $f\chi_Q$ instead of f

$$\omega(Q) \leq \omega(\{x : M(f\chi_Q)(x) > \lambda\}) \leq C \frac{\int_Q |f|^p \omega}{\lambda^p} \implies \omega(Q)\lambda^p \leq C \int_Q |f|^p \omega$$

Taking the limit as $\lambda \rightarrow \frac{1}{|Q|} \int_Q f$, we get

$$\omega(Q) \left(\frac{1}{|Q|} \int_Q f \right)^p \leq C \int_Q |f|^p \omega.$$

From here, we can deduce the A_p condition exactly like we did in Proposition 1.2.5. $\square$

In the $p = 1$ case, it is also true that we can extract weak bounds from the A_1 condition, and it is a simple consequence of the following theorem.

Theorem 1.3.3: Given any non-negative, measurable function ω there is a constant C_p such that

$$\|Mf\|_{L^p(\omega)} \leq C_p \|f\|_{L^p(M\omega)}, \quad 1 < p \leq \infty,$$

$$\|Mf\|_{L^{1,\infty}(\omega)} \leq C_1 \|f\|_{L^1(M\omega)}.$$

Proof: The proof for $1 < p < \infty$ follows from the endpoint cases by Marcinkiewicz interpolation, so we will proceed to proving these directly.

In the L^∞ case, it is important to look at $M\omega$ -null sets. A set E having $M\omega$ null measure is the same as it satisfying $\int_E M\omega = 0$, which (since $M\omega \geq 0$) implies $M\omega$ is zero in E , possibly except for a Lebesgue-null set. Now, one of two things happen:

- If $M\omega$ is zero in at least a point, it means that all averages around that point must be zero as well, which forces ω to be identically zero. In this case, there is nothing to prove.
- If $M\omega$ is never zero, then E must have Lebesgue measure 0. That means that taking the essential supremum with respect to $M\omega$ is greater or equal to taking the essential supremum with respect to the Lebesgue measure.

Assuming without loss of generality we are in the second case, it is now enough to show $\|Mf\|_{L^\infty(\omega)} \leq \|f\|_{L^\infty}$, as we will then have

$$\|Mf\|_{L^\infty(\omega)} \leq \|f\|_{L^\infty} \leq \|f\|_{L^\infty(M\omega)},$$

concluding the $p = \infty$ case. But now $\|Mf\|_{L^\infty(\omega)} \leq \|f\|_{L^\infty}$ is just a consequence of Hölder's inequality:

$$Mf(x) \leq \frac{1}{|Q|} \int_Q |f| \leq \frac{1}{|Q|} \|\chi_Q\|_1 \|f\|_\infty = \|f\|_\infty,$$

and now taking the supremum over any x we get (since the essential supremum is lesser or equal to the supremum)

$$\|Mf\|_{L^\infty(\omega)} \leq \sup_x Mf(x) \leq \|f\|_\infty.$$

For the $p = 1$ case, we use the trick we proved in proposition 1.2.3. Namely, we will use the fact that $\{Mf > \lambda\} \subseteq \cup_j 3Q_j$, with Q_j disjoint dyadic cubes of average greater than $4^{-n}\lambda$. From there, the trick will be to multiply and divide by $|3Q_j|$: one of the terms turns $\omega(3Q_j)$ into an average, which can be bounded by $M\omega$, while the other makes an integral of f appear in the usual way:

$$\begin{aligned} \omega(\{Mf > \lambda\}) &\leq \sum_j \omega(3Q_j) = \sum_j \int_{3Q_j} \omega = 3^n \sum_j |Q_j| \frac{1}{|3Q_j|} \int_{3Q_j} \omega(y) dy \\ &\leq 3^n \sum_j \left(4^n \lambda^{-1} \int_{Q_j} |f(x)| dx \right) \frac{1}{|3Q_j|} \int_{3Q_j} \omega(y) dy \\ &= \frac{12^n}{\lambda} \sum_j \int_{Q_j} |f(x)| \cdot \left(\frac{1}{|3Q_j|} \int_{3Q_j} \omega(y) dy \right) dx \\ &\lesssim \frac{1}{\lambda} \sum_j \int_{Q_j} |f(x)| M\omega(x) dx \lesssim \frac{\int_{\mathbb{R}^n} |f| M\omega}{\lambda}, \end{aligned}$$

which establishes the $L^1(M\omega) \rightarrow L^{1,\infty}(\omega)$ bound. $\square$

Proposition 1.3.4: Let ω be a weight which satisfies (1.6). Then, the maximal function is a bounded operator $L^1(\omega) \rightarrow L^{1,\infty}(\omega)$.

Proof: From theorem 1.3.3 we have the first inequality, and from (1.6) we have the second:

$$\|Mf\|_{L^{1,\infty}(\omega)} \leq \frac{C_1}{\lambda} \|f\|_{L^1(M\omega)} = \frac{C_1}{\lambda} \int |f| M\omega \leq \frac{C_1}{\lambda} \int |f| C\omega = \frac{CC_1}{\lambda} \int |f| \omega = \frac{\tilde{C}}{\lambda} \|f\|_{L^1(\omega)},$$

and so the maximal operator is bounded $L^1(\omega) \rightarrow L^{1,\infty}(\omega)$. $\square$

1.4 A_p Weights Exist

Given the definitions above, it is not immediately clear that interesting A_p weights even exist. It is immediate that $1 \in A_p$ with A_p constant one¹, but it is natural to wonder if this definition encompasses other weights we care about.

In this section, we will introduce a way to construct A_1 weights, as well as a way to construct more general A_p weights from A_1 weights. In the process, we will use these indirect methods to prove that $|x|^a$ is in A_p whenever $-n < a < n(p-1)$.

¹This is in fact the smallest possible A_p constant, as we have (using Hölder's inequality),

$$1 = \frac{1}{|Q|} \int_Q 1 = \frac{1}{|Q|} \int_Q \omega^{\frac{1}{p}} \omega^{\frac{1-p'}{p}} \leq \left(\frac{1}{|Q|} \int_Q \omega \right)^{\frac{1}{p}} \left(\frac{1}{|Q|} \int_Q \omega^{1-p'} \right)^{\frac{p-1}{p}} = \left(\left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{1}{|Q|} \int_Q \omega^{1-p'} \right)^{p-1} \right)^{\frac{1}{p}}$$

1.4.1 Constructing A_1 weights

Proposition 1.4.1: Let μ be a Borel measure which is finite in sets of finite Lebesgue measure², and $0 \leq \delta < 1$ be a real number. Then, assuming that $M\mu(x) < \infty$ almost everywhere, we have that $M\mu(x)^\delta$ is an A_1 weight, with A_1 characteristic depending only on δ and n .

Proof: Note that we require $M\mu(x) < \infty$ almost everywhere so that $M\mu(x)^\delta < \infty$ almost everywhere as well and $M\mu^\delta$ is a weight. From here, we wish to show that the weight $M\mu^\delta$ is in A_1 , that is that

$$M((M\mu)^\delta)(x) \leq CM\mu(x)^\delta$$

almost everywhere in x . It is therefore enough to show, for almost everywhere x and for every $Q \ni x$.

$$\frac{1}{|Q|} \int_Q M\mu^\delta \leq CM\mu(x)^\delta.$$

Let us define $\mu_1 = \mu\chi_{2Q}$ and $\mu_2 = \mu - \mu_1$. Then, we have $M\mu \leq M\mu_1 + M\mu_2$, and since $0 \leq \delta < 1$ we have

$$M\mu^\delta \leq (M\mu_1 + M\mu_2)^\delta \leq M\mu_1^\delta + M\mu_2^\delta,$$

which allows us to prove the bound for $M\mu_1^\delta$ and $M\mu_2^\delta$ separately. The idea of this decomposition is the following:

- For μ_1 , the χ_{2Q} term makes μ -averages appear from the μ_1 -measure of the whole space, and we will exploit this to prove the intended bound;
- For μ_2 , the fact that we removed all the measure in $2Q$ leads any non-zero averages of $\mu - \mu\chi_{2Q}$ around $y \in Q$ to necessarily be taken over big cubes, in the sense that they are comparable to Q . In particular, we will find that there is a small dilation of it which contains all elements of Q , in particular x , which will allow us to prove $M\mu_2(y) \leq cM\mu(x)$ for every $y \in Q$, from which the inequality above follows.

For $M\mu_1^\delta$, we can prove and use the so-called Kolmogorov's inequality,

$$\int_Q M\mu^\delta \lesssim |Q|^{1-\delta} \|\mu\|^\delta,$$

using the layer-cake representation:

$$\int_Q M\mu_1^\delta = \delta \int_0^\infty \alpha^{\delta-1} |\{x \in Q \mid M\mu_1(x)^\delta > \alpha\}| d\alpha \leq \delta \int_0^\infty \alpha^{\delta-1} \min(|Q|, \frac{C}{\alpha} \|\mu_1\|) d\alpha,$$

where $\|\mu_1\|$ denotes $\mu_1(\mathbb{R}^n)$ and we used the fact

$$\lambda |\{M\mu_1 > \lambda\}| \leq \frac{C}{\lambda} \|\mu_1\|,$$

²In particular, we can take $f \in L^1_{\text{loc}}$.

which can be proved in exactly the same way one proves the weak $(1, 1)$ bound for the maximal operator. The desired Kolmogorov inequality now follows from splitting the integral in the different regions associated to the minimum and computing each sub-integral:

$$\delta \int_0^{\frac{C\|\mu_1\|}{|Q|}} \alpha^{\delta-1} |Q| d\alpha + \delta \int_{\frac{C\|\mu_1\|}{|Q|}}^{\infty} \alpha^{\delta-1} \frac{C}{\alpha} |\mu_1| d\alpha = \left(\frac{C\|\mu_1\|}{|Q|} \right)^{\delta} |Q| - \frac{\delta}{\delta-1} \left(\frac{C\|\mu_1\|}{|Q|} \right)^{\delta-1} C \|\mu_1\| = C_{\delta} \frac{\|\mu_1\|^{\delta}}{|Q|^{\delta-1}}.$$

Using this last expression to bound the average of $M\mu_1^{\delta}$, we get:

$$\frac{1}{|Q|} \int_Q M\mu_1^{\delta} \leq C_{\delta} \left(\frac{\|\mu_1\|}{|Q|} \right)^{\delta} = C_{\delta,n} \left(\frac{\|\mu_1\|}{|2Q|} \right)^{\delta} = C_{\delta,n} \left(\frac{\|\mu\chi_{2Q}\|}{|2Q|} \right)^{\delta} = C_{\delta,n} \left(\frac{\mu(2Q)}{|2Q|} \right)^{\delta} \leq C_{\delta,n} M\mu^{\delta},$$

since we had assumed $x \in Q$ and therefore $x \in 2Q$.

For μ_2 , let us show that for any $y \in Q$ we have $M\mu_2(y) \leq M\mu(x)$. So, fix an arbitrary $y \in Q$ and let $\tilde{Q}$ be a cube which contains y and such that $\frac{\mu_2(\tilde{Q})}{|\tilde{Q}|} \neq 0$. Then, since $\mu_2(\chi_{2Q}) = 0$, $\tilde{Q}$ must contain points outside $2Q$ (as well as y), which means that we must have $\ell(\tilde{Q}) > \frac{\ell(Q)}{2}$ (see figure 2.9). Then, considering that $\tilde{Q}$ intersects Q , we have that $5\tilde{Q}$ covers the whole cube Q , as we are extending $\tilde{Q}$ in each direction by $4\frac{\ell(\tilde{Q})}{2} > 4\frac{\frac{1}{2}\ell(Q)}{2} = \ell(Q)$.

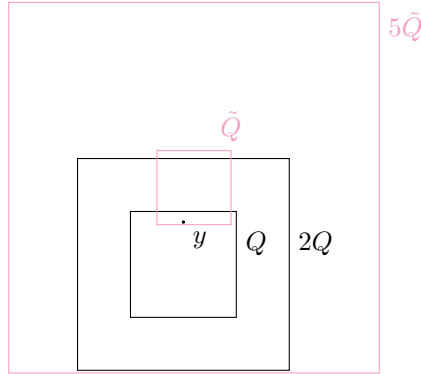


Figure 1.3: In black, the cubes Q and $2Q$. In pink, a cube $\tilde{Q}$ which contains the point $y \in Q$ and has positive μ_2 -average, as well as its 5-dilation, which contains Q .

Then, we have

$$\frac{\mu_2(\tilde{Q})}{|\tilde{Q}|} \leq 5^n \frac{\mu_2(5\tilde{Q})}{|5\tilde{Q}|} \leq 5^n \frac{\mu(5\tilde{Q})}{|5\tilde{Q}|} \leq 5^n M\mu(x),$$

since $Q \subseteq 5\tilde{Q}$ and therefore $x \in 5\tilde{Q}$. Taking the supremum over the different $\tilde{Q}$, we get $M\mu_2(y) \leq 5^n M\mu(x)$, from which we get

$$\frac{1}{|Q|} \int_Q M\mu_2(y)^{\delta} dy \leq 5^n \frac{1}{|Q|} \int_Q M\mu(x)^{\delta} dy = 5^n M\mu(x)^{\delta},$$

which concludes the proof. $\square$

Remark 1.4.2: It is possible to prove a converse of proposition 1.4.1. Indeed, given any $\omega \in A_1$ we can

write it as $c(x)Mf(x)$ for some $f \in L^1_{\text{loc}}$ and some $c \in L^\infty$ such that $c^{-1} \in L^\infty$. The proof can be found in [Duo01] (Theorem 7.7 - (2)) and requires the Reverse Hölder Inequality, an important technical lemma we prove in section 1.5.

Example 1.4.3: Using this result, we can construct A_1 weights. In particular, by considering the dirac delta δ , we get that $M\delta^\delta = |x|^{-n\delta} \in A_1$. Indeed, $\frac{\delta(Q)}{|Q|}$ is only non-zero when $0 \in Q$, in which case $\frac{\delta(Q)}{|Q|} = |Q|^{-1}$. This value is maximized by $|x|^{-n}$, since $0, x \in Q$, which establishes $M\delta^\delta = |x|^{n\delta}$. Since $0 \leq \delta < 1$, we get $|x|^{-\alpha} \in A_1$ for $0 \leq \alpha < n$.

Note that, for $\alpha \geq n$, $|x|^{-\alpha}$ is not locally integrable, and therefore not in A_1 , and for $\alpha < 0$ we have that $M(|x|^{-\alpha})$ is not finite almost everywhere, which means the A_1 condition cannot be satisfied. This can be seen by considering balls around the origin (since cubes and balls are comparable) with increasing radii and noting that the average is $\propto R^\alpha$, which explodes when $R \rightarrow \infty$.

1.4.2 Constructing A_p weights

Proposition 1.4.4: Let $\omega_0, \omega_1 \in A_1$. Then, $\omega_0\omega_1^{1-p} \in A_p$.

Proof: We want to show that $\omega_0\omega_1^{1-p} \in A_p$. That is, that

$$\sup_Q \left(\frac{1}{|Q|} \int_Q \omega_0\omega_1^{1-p} \right) \left(\frac{1}{|Q|} \int_Q (\omega_0\omega_1^{1-p})^{1-p'} \right)^{p-1} = \sup_Q \left(\frac{1}{|Q|} \int_Q \omega_0\omega_1^{1-p} \right) \left(\frac{1}{|Q|} \int_Q \omega_0^{1-p'} \omega_1 \right)^{p-1} < \infty.$$

By definition, $\omega_0, \omega_1 \in A_1$ means that $M\omega_i(x) \leq C_i\omega_i(x)$ almost everywhere, which implies that $\omega_i(x)^{-1} \leq C_i M\omega_i(x)^{-1} \leq C_i \left(\frac{\omega_i(Q)}{|Q|} \right)^{-1}$ almost everywhere. This allows us to bound ω_1^{1-p} and $\omega_0^{1-p'}$ in the expression above:

$$\begin{aligned} \left(\frac{1}{|Q|} \int_Q \omega_0\omega_1^{1-p} \right) \left(\frac{1}{|Q|} \int_Q \omega_0^{1-p'} \omega_1 \right)^{p-1} &\lesssim_{C_0, C_1} \left(\frac{1}{|Q|} \int_Q \omega_0 \left(\frac{\omega_1(Q)}{|Q|} \right)^{1-p} \right) \left(\frac{1}{|Q|} \int_Q \left(\frac{\omega_0(Q)}{|Q|} \right)^{1-p'} \omega_1 \right)^{p-1} = \\ &= \left(\frac{\omega_0(Q)}{|Q|} \right) \left(\frac{\omega_1(Q)}{|Q|} \right)^{1-p} \left(\frac{\omega_0(Q)}{|Q|} \right)^{(1-p')(p-1)} \left(\frac{\omega_1(Q)}{|Q|} \right)^{p-1} = 1, \end{aligned}$$

where we used $(1-p')(p-1) = -1$. Now taking the supremum over all Q gives us the A_p condition on $\omega_0\omega_1^{1-p}$, with a constant depending only on C_0 and C_1 , the A_1 characteristics of ω_0 and ω_1 . $\square$

Remark 1.4.5: It can be proven that any A_p weight can be factorized as above. That is, given $\omega \in A_p$, there are ω_0, ω_1 such that $\omega = \omega_0\omega_1^{1-p}$. The proof can be found in [Gra14] section 7.5.1³.

Example 1.4.6: Using this proposition, we can extend example 1.4.3 to the A_p case. Indeed, since $|x|^\alpha \in A_1$ for $0 \leq \alpha < n$, we can take $\omega_0 = |x|^\alpha$ and $\omega_1 = 1$ to get $\omega_0\omega_1^{1-p} = |x|^\alpha \in A_p$ for $0 \leq \alpha < n$ and we can take $\omega_0 = 1$ and $\omega_1 = |x|^\alpha$ to get $\omega_0\omega_1^{1-p} = |x|^{(1-p)\alpha} \in A_p$. Combining these, we get that $|x|^{-\alpha} \in A_p$ for $n(1-p) < \alpha < n$.

³In editions before the third, it can be found in section 9.5.1 of [Gra09] instead.

Note that for $\alpha \geq n$ we have that $|x|^{-\alpha}$ is not locally integrable and for $\alpha \leq n(1-p)$ we have $\alpha = m(1-p)$ with $m > n$ and so $(|x|^{m(1-p)})^{p'-1} = |x|^{-m}$ is not locally integrable. Since $\omega \in A_p$ requires both ω and $\omega^{p'-1}$ to be locally integrable, this shows that the range $n(1-p) < \alpha < n$ is sharp.

1.5 Strong Boundedness for A_p weights

In proposition 1.2.3, we saw that the maximal operator is bounded $L^p(\omega) \rightarrow L^{p,\infty}(\omega)$. In this section, we will bump up this result, and show that it is bounded $L^p(\omega) \rightarrow L^p(\omega)$. The first step is noting that, as usual, the L^∞ case works rather well:

Lemma 1.5.1: Let ω be an A_p weight. Then, $\|Mf\|_{L^\infty(\omega)} \leq \|f\|_{L^\infty(\omega)}$

Proof: The key observation is that since ω is non-zero almost everywhere, we have that $\omega(E) = 0 \implies |E| = 0$. Since the converse holds for all functions, we have $\omega(E) = 0 \iff |E| = 0$ and so the essential supremum is taken over the same collection of sets as the Lebesgue case. So, we have

$$\|Mf\|_{L^\infty(\omega)} = \|Mf\|_{L^\infty} \leq \|f\|_{L^\infty} = \|f\|_{L^\infty(\omega)},$$

where the middle inequality is a well-known fact whose proof can be found in the proof of theorem 1.3.3. □

Using the results from lemma 1.5.1 and 1.2.3, we can use Marcinkiewicz interpolation to interpolate between (∞, ∞) and weak (p, p) and conclude that the maximal function is bounded $L^q(\omega) \rightarrow L^q(\omega)$ for $p < q < \infty$. If by any chance an A_p weight was also an $A_{\tilde{p}}$ weight, for a slightly smaller $\tilde{p} < p$, then we could use interpolation between weak $(\tilde{p}, \tilde{p})$ and strong (∞, ∞) in order to get the strong L^p bound.

As it turns out, it is possible to decrease the p a little bit. Doing so requires a technical but important lemma, typically called the Reverse Hölder Inequality:

Lemma 1.5.2 (Reverse Hölder Inequality): Given any weight $\omega \in L^p$, with $1 \leq p < \infty$, we have that for any cube Q there is an ε such that

$$\left(\frac{1}{|Q|} \int_Q \omega^{1+\varepsilon} \right)^{\frac{1}{1+\varepsilon}} \leq \frac{C}{|Q|} \int_Q \omega,$$

where C is a constant which depends only on p and the A_p constant of ω .

Before proceeding with a proof, let us show that this lemma does what it sets out to do, by using it to prove that an $\omega \in A_p$ is also in A_q for a smaller q , hence proving the strong L^p bound for $\omega \in A_p$. In the process, we prove a couple related consequences of the Reverse Hölder Inequality:

Corollary 1.5.3: Let $1 \leq p < \infty$. Then,

1. We have that $A_p = \cup_{q < p} A_q$ ($p \neq 1$). In particular, any $\omega \in A_p$ is an A_q weight, for some $q < p$.
2. If $\omega \in A_p$, then there is an $\varepsilon > 0$ such that $\omega^{1+\varepsilon}$ is still in A_p .

3. There is a constant C and a real number $\delta > 0$ only depending on p and the A_p constant of ω such that for $\omega \in A_p$ we have

$$\omega(S) \leq \omega(Q) \left(\frac{|S|}{|Q|} \right)^\delta$$

Proof:

1. The fact that $\cup_{q < p} A_q \subseteq A_p$ is a direct consequence of the second result of proposition 1.2.8, which tells us that an $\omega \in A_p$ is in any A_q with $q > p$. So, all we need to prove is that given any $\omega \in A_p$ there is a $q < p$ such that $\omega \in A_q$. Note that we want to decrease p a bit. The reverse Hölder inequality seems promising, but it seems to do the opposite: it allows us to increase the power a little. This motivates us to try to use duality and consider $\omega^{1-p'}$ instead. By the reverse Hölder inequality, there is some ε such that

$$\left(\frac{1}{|Q|} \int_Q \omega^{(1-p')(1+\varepsilon)} \right)^{\frac{1}{1+\varepsilon}} \leq \frac{C}{|Q|} \int_Q \omega^{1-p'}.$$

Now this is exactly what we needed! If we raise both sides to the power of $\frac{1}{1-p'}$ and multiply them both by the average of ω in Q , we get

$$\left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{1}{|Q|} \int_Q \omega^{(1-p')(1+\varepsilon)} \right)^{\frac{1}{(1+\varepsilon)(1-p')}} \leq \left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{C}{|Q|} \int_Q \omega^{1-p'} \right)^{\frac{1}{1-p'}}.$$

On the right-hand side, we have (minus the supremum) the A_p condition for ω and on the left-hand side, we have the A_q condition for ω , as long as $1 - q' = (1 - p')(1 + \varepsilon)$. Note that $q' = p' - 1 - \varepsilon + p'\varepsilon + 1 = p' + (p' - 1)\varepsilon$. So, since $p' - 1 > 0$, we have $q' > p' \implies p > q$. Taking the supremum over all the cubes on both sides proves that $\omega \in A_q$ ($q < p$), as intended.

2. The key idea in this proof is recalling that by proposition 1.2.8 we have that $\omega^{1-p'} \in A_{p'}$, which allows us to apply the reverse Hölder to both ω and $\omega^{1-p'}$ and get that there are $\varepsilon_1, \varepsilon_2$ such that

$$\left(\frac{1}{|Q|} \int_Q \omega^{1+\varepsilon_1} \right)^{\frac{1}{1+\varepsilon_1}} \leq \frac{C_1}{|Q|} \int_Q \omega \quad \text{and} \quad \left(\frac{1}{|Q|} \int_Q (\omega^{1-p'})^{1+\varepsilon_2} \right)^{\frac{1}{1+\varepsilon_2}} \leq \frac{C_2}{|Q|} \int_Q \omega^{1-p'}.$$

Using Hölder's inequality, we it is easy to check that the reverse Hölder inequality applies to any $\varepsilon' < \varepsilon$. So, taking $\varepsilon = \min(\varepsilon_1, \varepsilon_2)$, we can use both these inequalities to get:

$$\left(\frac{1}{|Q|} \int_Q \omega^{1+\varepsilon} \right) \left(\frac{1}{|Q|} \int_Q (\omega^{1+\varepsilon})^{1-p'} \right)^{p-1} \leq C_1 C_2 \left(\frac{1}{|Q|} \int_Q \omega \right)^{1+\varepsilon} \left(\frac{1}{|Q|} \int_Q \omega^{1-p'} \right)^{(p-1)(1+\varepsilon)},$$

which means the A_p condition for $\omega^{1+\varepsilon}$ follows from the A_p condition for ω (by taking the supremum over all Q).

3. Let ε be the constant given by the reverse Hölder inequality lemma. That is, such that ω satisfies

$$\left(\frac{1}{|Q|} \int_Q \omega^{1+\varepsilon} \right)^{\frac{1}{1+\varepsilon}} \leq \frac{C}{|Q|} \int_Q \omega.$$

Then, we can use Hölder's inequality as follows:

$$\omega(S) = \int_Q \omega \chi_S \leq \left(\int_Q \omega^{1+\varepsilon} \right)^{\frac{1}{1+\varepsilon}} |S|^{\frac{\varepsilon}{1+\varepsilon}}$$

□

Now that we have shown the importance of this inequality, we proceed to proving it:

Proof (Reverse Hölder Inequality): In this proof we will be trying to prove

$$\frac{1}{|Q|} \int_Q \omega^{1+\varepsilon} \lesssim \left(\frac{\omega(Q)}{|Q|} \right)^{1+\varepsilon},$$

by performing a sequence of Calderón-Zygmund decompositions of Q at heights $\frac{1}{|Q|} \int_Q \omega =: \lambda_0 < \lambda_1 < \lambda_2 < \dots < \lambda_m < \dots$. We will prove that the “bad parts” of the decomposition cover the space (almost everywhere). This will allow us to split the integral into a collection of “bad parts” and use the corresponding bounds of ω by λ_m on the integral to get the intended inequality. The key will be to consider $\lambda_m = c\lambda_{m-1}$, as this will allow us to recover λ_0 terms from the λ_m .

For each λ_m , let us then Calderón-Zygmund decompose Q^4 into $G_m = \cup Q_j^m$ and B_m . These sets, as usual, satisfy $\lambda_m < \frac{1}{|Q_j^m|} \int_{Q_j^m} \omega \leq 2^n \lambda_m$ and $\omega(x) \leq \lambda_m$, $x \in B_m$. Recall that by construction the G_m contain all the points such that $M_d \omega > \lambda_m$. This implies that $G_{m+1} \subseteq G_m$ as $M_d \omega > \lambda_{m+1} \implies M_d \omega > \lambda_m$ and so

$$B_0 \subseteq B_1 \subseteq B_2 \subseteq \dots \subseteq B_m \subseteq B_{m+1} \subseteq \dots$$

We wish to show that these sets cover the whole space (almost everywhere). In order to do that, we will show that $|\cap_m G_m| = 0$, by estimating $|G_{m+1}| = |G_m \cap G_{m+1}|$. Recall that $G_m = \cup Q_j^m$, with Q_j^m disjoint dyadic cubes, and that $G_{m+1} \subseteq G_m$. Since G_{m+1} also consists of a union of dyadic cubes in the same dyadic grid $\mathcal{D}(Q)$, we have that the intersection of each Q_j^m with G_{m+1} will be a union $\cup_{\alpha \in A_j} Q_\alpha^{m+1}$ of dyadic cubes in G_{m+1} .

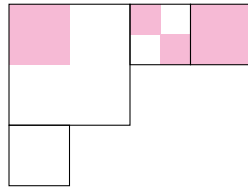


Figure 1.4: A visual representation of G_m (in drawn black) and G_{m+1} (in filled pink).

⁴This decomposition works in a similarly to one using the standard dyadic grid, except that we are considering the dyadic grid $\mathcal{D}(Q)$. The construction is the same, we add to G_m the maximal dyadic cubes such that the average is greater than λ_m and the remaining points are in B_m and satisfy $\omega(x) \leq \lambda_m$ by Lebesgue Differentiation Theorem. Note that by construction of λ_m , $\frac{1}{|Q|} \int_Q \omega \leq \lambda_m$, which means the stopping time argument applies.

So, we have

$$\begin{aligned} |G_m \cap G_{m+1}| &= \sum_j |Q_j^m \cap G_{m+1}| = \sum_j \sum_{\alpha \in A_j} |Q_\alpha^{m+1}| \leq \sum_j \sum_{\alpha \in A_j} \frac{1}{\lambda_{m+1}} \int_{Q_\alpha^{m+1}} \omega \\ &\leq \sum_j \frac{1}{\lambda_{m+1}} \int_{Q_j^m} \omega \leq \sum_j \frac{2^n \lambda_m}{\lambda_{m+1}} |Q_j^m| = \frac{2^n \lambda_m}{\lambda_{m+1}} |G_m| \end{aligned}$$

If we choose $\lambda_{m+1} = \alpha^{-1} 2^n \lambda_m$ (for a fixed $0 < \alpha < 1$), we get that $|G_{m+1}| = |G_m \cap G_{m+1}| \leq \alpha |G_m|$, for every m , which we can iterate to get $|G_m| \leq \alpha^m |G_0|$. This means that $|G_m| \rightarrow 0$, which in turn implies $|\cap_m G_m| = 0$.

We are now ready to bound the integral of $\omega^{1+\varepsilon}$ in Q :

$$\begin{aligned} \frac{1}{|Q|} \int_Q \omega^{1+\varepsilon} &= \frac{1}{|Q|} \int_{B_0} \omega^{1+\varepsilon} + \sum_{m=0}^{\infty} \frac{1}{|Q|} \int_{B_{m+1} \setminus B_m} \omega^{1+\varepsilon} \\ &\leq \frac{1}{|Q|} \omega(B_0) \lambda_0^\varepsilon + \frac{1}{|Q|} \sum_{m=0}^{\infty} \omega(B_{m+1} \setminus B_m) \lambda_{m+1}^\varepsilon, \end{aligned}$$

where we used the fact that in B_m we have $\omega(x) \leq \lambda_m$ to bound ω^ε .

The first term can be easily bounded by $\frac{\omega(Q)}{|Q|} \lambda_0^\varepsilon = \frac{\omega(Q)}{|Q|}^{1+\varepsilon}$. Moreover, by definition of the λ_m , we have $\lambda_m = 2^{mn} \alpha^{-m} \lambda_0$, which means that we have a geometric sum multiplied by λ_0^ε , and an $\frac{\omega(B_{m+1} \setminus B_m)}{|Q|}$ term which could be bounded by $\frac{\omega(Q)}{|Q|}$, so we seem to be in the right track. However, since $\alpha < 1$, the ratio $2^{n\varepsilon} \alpha^{-\varepsilon}$ can't be taken to be smaller than 1. This can be fixed by exploiting the size of $\omega(B_{m+1} \setminus B_m)$ in order to introduce a smaller factor.

Note that we have $B_{m+1} \setminus B_m = B_{m+1} \cap G_m$, which allows us to bound $\omega(B_{m+1} \setminus B_m)$ by $\omega(G_m)$. Now, we can use the bounds on $|G_m|$ to extract a bound for $\omega(G_m)$: recall that (by lemma 1.2.10) $|G_{m+1}| \leq \alpha |G_m|$ implies that there is a $0 < \beta < 1$ such that $\omega(G_{m+1}) \leq \beta \omega(G_m)$, and so $\omega(G_m) \leq \beta^m \omega(G_0)$.

Tying these all together, we have

$$\begin{aligned} \frac{1}{|Q|} \int_Q \omega^{1+\varepsilon} &\leq \left(\frac{\omega(Q)}{|Q|} \right)^{1+\varepsilon} + \sum_{m=0}^{\infty} (2^{(m+1)n} \alpha^{-(m+1)})^\varepsilon \beta^m \frac{\omega(G_0)}{|Q|} \lambda_0^\varepsilon \\ &\leq \left(\frac{\omega(Q)}{|Q|} \right)^{1+\varepsilon} + 2^{n\varepsilon} \alpha^{-\varepsilon} \sum_{m=0}^{\infty} \gamma^m \left(\frac{\omega(Q)}{|Q|} \right)^{1+\varepsilon}, \end{aligned}$$

with $\gamma = (2^n \alpha^{-1})^\varepsilon \beta$, which can now be taken to be less than 1 by choosing small enough ε . Once we do, the series converges, and we get

$$\frac{1}{|Q|} \int_Q \omega^{1+\varepsilon} \lesssim \left(\frac{\omega(Q)}{|Q|} \right)^{1+\varepsilon},$$

as intended. □

We are now ready to prove the strong boundedness:

Theorem 1.5.4: Let $\omega \in A_p$. Then, the maximal operator is bounded in $L^p(\omega)$.

Proof: By corollary 1.5.3 (item 1), we have that since $\omega \in A_p$, there is some $q < p$ such that $\omega \in A_q$. So, by proposition 1.2.3, we have that the maximal operator is bounded $L^q(\omega) \rightarrow L^{q,\infty}(\omega)$. Moreover, lemma 1.5.1 establishes that the maximal operator is bounded $L^\infty(\omega) \rightarrow L^\infty(\omega)$.

Using the Marcinkiewicz interpolation theorem, one gets from these endpoint cases that the maximal operator is bounded $L^{\tilde{q}}(\omega) \rightarrow L^{\tilde{q}}(\omega)$ for any $q < \tilde{q} \leq \infty$, p in particular. $\square$

1.5.1 A_∞ Weights

Up to this point, we have not considered the $p = \infty$ case. Even though this case will not be important in what follows, the A_∞ class ties up the A_p classes very nicely, and it can be proven to be quite robust, in the sense that many natural alternative definitions end up being the same. A careful exposition on the A_∞ weights can be found in [Gra14] section 7.3 (including the proof of the equivalence of the different definitions, omitted in this section).

The first step in defining A_∞ spaces is perhaps motivated by the first item in Corollary 1.5.3, which states that each A_p can be obtained from the union of the lower A_q . On the other hand, the second item in proposition 1.2.8 tells us that these classes are all increasingly chained. It is therefore reasonably natural to think of their limit when p goes to ∞ , which will then correspond to $\cup_{1 \leq p} A_p$. This is one of many possible definitions for the A_∞ class, but it doesn't tell us anything about $[\omega]_\infty$.

Alternatively, but still in the same line of thinking, one could try to define $[\omega]_\infty$, by taking the limit as $p \rightarrow \infty$ of

$$\left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{1}{|Q|} \int_Q \omega^{1-p'} \right)^{p-1}.$$

There is a technical obstacle to overcome here, as it is not necessarily obvious what the limit on p of the second term would be. This is solved via a consequence of the Jensen inequality, which states that

$$\left(\int |f|^q \mu \right)^{\frac{1}{q}} \geq \exp \left(\int \log(|f|) \mu \right),$$

and it can be proven that the right-hand expression is the limit of the left-hand side. Applying this to the A_p condition, one gets

$$[\omega]_\infty := \sup_Q \left(\frac{1}{|Q|} \int_Q \omega \right) \exp \left(\int_Q \log(\omega^{-1}) \right),$$

and one can define $A_\infty := \{\omega : [\omega]_\infty < \infty\}$. Note that by construction, we have $\cup_{1 \leq p} A_p \subseteq \{\omega : [\omega]_\infty < \infty\}$, but the converse doesn't necessarily hold.

The remarkable fact about the A_∞ class is that not only these two definitions are equivalent, but they are also equivalent to many other characterizations, including some important properties of the A_p weights, such as:

- The weight ω satisfying the Reverse Hölder Inequality (Lemma 1.5.2)
- The weight ω satisfying property 3 of Corollary 1.5.3. That is, there existing δ, C such that whenever $S \subseteq Q$, we have $\frac{\omega(S)}{\omega(Q)} \leq C \left(\frac{|S|}{|Q|} \right)^\delta$. Note that in particular, this means this property also implies the Reverse Hölder Inequality.

It is important to point out that since we have $A_\infty = \cup_{1 \leq p} A_p$, any $\omega \in A_\infty$ is in some A_p , and so Lemma 1.5.1 gives us that the maximal operator is bounded in $L^\infty(\omega)$. However, the converse is far from holding. Given any weight ν which is positive almost everywhere in the Lebesgue measure, we have that

$$\|Tf\|_{L^\infty(\nu)} \leq \|f\|_{L^\infty(\nu)},$$

by a similar reasoning to the one in the proof of Lemma 1.5.1: since $\nu(E) = 0 \iff |E| = 0$, the essential suprema are taken over the same sets regardless of which of the measures we choose, and so the bound follows from the standard L^∞ bound. So, we can choose any such ν which is not locally integrable as a counterexample: the maximal is bounded in $L^\infty(\nu)$ but it definitely is not in any A_p and is therefore is not in $\cup_{1 \leq p} A_p = A_\infty$ either. As a consequence, we have that unlike in the remaining cases, there is no equivalence between $\omega \in L^\infty$ and $L^\infty(\omega)$ bounds for the maximal operator, which justifies treating this endpoint case separately and differently.

1.6 Important Weighted Results

In the previous sections, we have defined and studied the weights over which the maximal operator has the natural boundedness properties, with the hope that these would translate well into more general settings. In this section, we present a couple of important results in this direction. The first, Rubio de Francia's extrapolation, which we will not prove at all, allows us to extrapolate boundedness in one A_p into other A_q s, as long as the bound depends only on the A_p constant. The second, which we will not prove just yet (that is where Sparse Domination will come in!), establishes L^p boundedness of general Calderón-Zygmund operators in the setting of A_p weights.

1.6.1 Extrapolation

The following proposition, although generally very useful, will not be needed in what follows, since Sparse Domination techniques will allow us to prove the results for all A_p directly. However, the proof for A_2 weights is oftentimes simpler, in which case it can be quicker to just use extrapolation to extend the result to the remaining A_p . In particular, one big advantage of using extrapolation is that it keeps track of the dependence on the A_p characteristic:

Theorem 1.6.1: Let T be an operator defined on every $L^p(\omega)$, for every $1 \leq p < \infty$ and every $\omega \in A_p$ and assume that T is bounded in $L^q(\omega)$ for a fixed q and any $\omega \in A_q$, such that

$$\|Tf\|_{L^q(\omega)} \leq N([\omega]_q) \|f\|_{L^q(\omega)},$$

with N a positive increasing function. Then, for any $1 < p < \infty$ and any $\omega \in A_p$, we have

$$\|Tf\|_{L^p(\omega)} \leq C_{p,q} N \left(c_{n,p,q} [\omega]_p^{\max(\frac{q-1}{p-1}, 1)} \right) \|f\|_{L^p(\omega)},$$

with $C_{p,q}$ and $c_{n,p,q}$ constants depending on respectively p, q and n, p, q .

The full proof of this theorem can be found in section 7.5.2 of [Gra14]. It is a technical proof, which uses a trick called Rubio de Francia algorithm, which hinges in the study of the operators

$$Ru(x) = \sum_{k=0}^{\infty} \frac{T^k u(x)}{(2\|T\|_{L^p(\omega) \rightarrow L^p(\omega)})^k},$$

in the particular cases of $T = M$ and $T = \frac{M(f\omega)}{\omega}$.

An alternative and more elementary proof can be found in [Duo01] (Theorem 7.8), although the theorem is stated and proved without considering the dependence on the A_p characteristic.⁵

1.6.2 Calderón-Zygmund Operators

Recall that a Calderón-Zygmund Operator T is an operator which is bounded in L^2 and such that there is a standard kernel K such that for every L^2 function of compact support f we have

$$Tf(x) = \int_{\mathbb{R}^n} K(x, y) f(y) dy, \quad x \notin \text{supp}(f).$$

Recall also that a standard kernel $K(x, y)$ is a function defined in $\mathbb{R}^n \times \mathbb{R}^n \setminus \{(x, x) : x \in \mathbb{R}^n\}$ which satisfies, for some $\delta > 0$

- $|K(x, y)| \leq \frac{C}{|x-y|^n}$
- $|K(x, y) - K(x, z)| \leq C \frac{|y-z|^\delta}{|x-y|^{n+\delta}}, \text{ if } |x-y| > 2|y-z|$
- $|K(x, y) - K(w, y)| \leq C \frac{|x-w|^\delta}{|x-y|^{n+\delta}}, \text{ if } |x-y| > 2|x-w|$

It has been long known that Calderón-Zygmund have nice boundedness properties in A_p -weighted L^p spaces:

Proposition 1.6.2: Let T be a Calderón-Zygmund operator and $\omega \in A_p$. Then,

$$\|Tf\|_{L^p(\omega)} \lesssim \|f\|_{L^p(\omega)}, \quad 1 < p < \infty$$

Moreover,

⁵It is worth pointing out that this proof has an implicit step: showing that $M((uw)^s) < \infty$ almost everywhere, in order to use Proposition 1.4.1. This follows from the following equality near the end of the proof

$$C \int_{\mathbb{R}^n} (wu)^{(p/q)'} w^{1-(p/q)'} = C,$$

since $(wu)^{(p/q)'} = ((wu)^s)^{\frac{(p/q)'}{s}}$ and s was chosen such that $w^{1-(p/q)'} \in A_{(p/q)'/s}$, which means that

$$(wu)^s \in L^{\frac{(p/q)'}{s}} \left(w^{1-(p/q)'} \right) \implies M((wu)^s) \in L^{\frac{(p/q)'}{s}} \left(w^{1-(p/q)'} \right),$$

which implies that $M((wu)^s)$ is finite ω -almost everywhere, and since ω is non-zero almost everywhere this implies that $M((wu)^s)$ is also finite almost everywhere in the Lebesgue measure.

Proposition 1.6.3: Let T be a Calderón-Zygmund operator and $\omega \in A_1$. Then,

$$\omega(\{x : Tf(x) > \lambda\}) \lesssim \frac{1}{\lambda} \int f \omega.$$

In this section, we will discuss the classical methods of proving these theorems.

The weak (1,1) Inequality

The proof of proposition 1.6.3 follows from proposition 1.6.2 in the style of usual Calderón-Zygmund Operators weak (1,1) proofs, with small modifications to account for the extra ω term:

Proof (Proposition 1.6.3): We will prove the weak bound for $f \in L^2 \cap L^1 \cap L^1(\omega)$ ⁶, from which we know that T can be extended to $L^1(\omega)$ by density. Recall that T is originally only defined in L^2 , so the only concern would be if this extension agrees with the original operator in $L^2 \cap L^1(\omega)$. One can prove that it does by considering simple functions $s_n \nearrow f$ (the usual construction works well). Then, we have that the functions s_n are in $L^2 \cap L^1(\omega)$ and converge to f in both these topologies. As a result, Ts_n converges to both the original Tf and the extension respectively in L^2 and $L^1(\omega)$, from which we can extract almost everywhere convergence along a subsequence. Indeed, by taking a subsequence of Ts_n which converges to the original Tf almost everywhere and then taking a subsequence of this subsequence which converges to the extension almost everywhere⁷, we ensure that the original Tf and the extension are the same function, as intended.

We begin by decomposing f , using the Calderón-Zygmund decomposition at height λ of the space into $G = \cup Q_j$ and $B = G^c$, such that

$$\lambda < \frac{1}{|Q_j|} \int_{Q_j} f \leq 2^n \lambda \quad \text{and} \quad f(x) \leq \lambda, \quad x \in B.$$

From here, we define

$$g(x) = \begin{cases} \frac{1}{|Q_j|} \int_{Q_j} f, & x \in Q_j \\ f(x) & x \in B \end{cases}; \quad b(x) = f(x) - g(x).$$

These two functions each have an important property: g is bounded by $2^n \lambda$, whereas b has integral zero. Since we have

$$\{M(g+b) > \lambda\} \subseteq \left\{Mg > \frac{\lambda}{2}\right\} + \left\{Mb > \frac{\lambda}{2}\right\} \implies \omega(\{Mf > \lambda\}) \leq \omega\left(\left\{Mg > \frac{\lambda}{2}\right\}\right) + \omega\left(\left\{Mb > \frac{\lambda}{2}\right\}\right),$$

⁶This might seem like an odd choice, but it comes from the fact that we need the function to be

- $L^1(\omega)$, since it is the space for which we want to prove boundedness;
- L^2 , since this is where T is originally defined and bounded;
- L^1 for technical reasons (we will use this to prove that the function b we will define is an L^2 function).

Moreover, choosing this space instead of $\mathcal{S}$ makes the extension argument a bit more direct.

⁷Recall that since $\omega \in A_1$ we have $|E| = 0 \iff \omega(E) = 0$, so a sequence converges ω -almost everywhere if and only if it converges almost everywhere.

it is enough to bound each of the terms. We will begin to show that $\|g\|_{L^1(\omega)} \lesssim \|f\|_{L^1(\omega)}$ (which in turn implies $\|b\|_{L^1(\omega)} = \|f - g\|_{L^1(\omega)} \leq \|g\|_{L^1(\omega)} + \|f\|_{L^1(\omega)} \lesssim \|f\|_{L^1(\omega)}$), in order to have the freedom to bound the superlevel sets by the norms of b and g and have that translate into bounds by the norm of f . In B , we have $f = g$. In each Q_j , we have

$$\int_{Q_j} |g(x)|\omega(x)dx \leq \int_{Q_j} \left(\frac{1}{|Q_j|} \int_{Q_j} |f(y)|d(y) \right) \omega(x)dx = \int_{Q_j} \frac{\omega(Q_j)}{|Q_j|} |f(y)|d(y).$$

Since $\omega \in A_1$, we have $\frac{\omega(Q_j)}{|Q_j|} \leq M\omega(y) \leq C\omega(y)$, which gives us

$$\int_{Q_j} |g(x)|\omega(x)dx \leq \int_{Q_j} \frac{\omega(Q_j)}{|Q_j|} |f(y)|d(y) \leq C \int_{Q_j} |f(y)|\omega(y)dy,$$

and so indeed $\|g\|_{L^1(\omega)} \lesssim \|f\|_{L^1(\omega)}$.

For the good part g , we use the fact that $g \leq 2^n \lambda$, along with the fact that by proposition 1.6.2 the operator T is bounded in $L^2(\omega)$ and therefore (by the Chebyshev inequality) satisfies the weak $(2, 2)$ bound:

$$\omega \left(\left\{ x : Tg > \frac{\lambda}{2} \right\} \right) \lesssim \int \frac{|g|^2 \omega}{\lambda^2} \leq \int \frac{2^n \lambda |g| \omega}{\lambda^2} \lesssim \frac{\int |g| \omega}{\lambda} \lesssim \frac{\|f\|_{L^1(\omega)}}{\lambda}.$$

For the bad part b , we will consider separately $\cup_j 2Q_j$ and its complement. For the bound of $\omega \left(\left\{ x \in \cup_j 2Q_j : Mb > \frac{\lambda}{2} \right\} \right)$, it is enough to simply estimate $\omega(\cup_j 2Q_j)$, using Lemma 1.2.2 to control the dilation as well as the bound on the average of f over Q_j :

$$\omega(\cup_j 2Q_j) \leq \sum_j \omega(2Q_j) \lesssim \sum_j \omega(Q_j) = \sum_j \frac{\omega(Q_j)}{|Q_j|} |Q_j| \lesssim \sum_j \int_{Q_j} f \frac{\omega(Q_j)}{|Q_j|} \lesssim \sum_j \int_{Q_j} f \omega \leq \|f\|_{L^1(\omega)},$$

where we used $\omega \in A_1 \implies \frac{\omega(Q_j)}{|Q_j|} \leq \omega(x)$ for $x \in Q_j$ again.

On the complement, we can use Chebyshev's inequality

$$\omega \left(\left\{ x \in (\cup_j 2Q_j)^c : |Tb| > \frac{\lambda}{2} \right\} \right) \leq \frac{2}{\lambda} \int_{(\cup_j 2Q_j)^c} |Tb| \omega$$

and try to bound this integral.

In order to do that, it is important to separate b into parts $b_i := b\chi_{Q_i}$. Note that since there is a countable number of dyadic squares, we can write $b = \sum_{i=0}^{\infty} b_i$. Our goal is to show that $|Tb| \leq \sum_{i=0}^{\infty} |Tb_i|$, which would then allow us to deal with each b_i separately. By the dominated convergence theorem, we have that $b(x) = \sum_{i=0}^{\infty} b_i(x)$ holds in L^2 , as $|b|^2$ is an L^2 function which dominates $\sum_i b_i$: we have that $g \in L^2$ (it is in L^∞ and in L^1 by construction), which means $b \in L^2$, and we also have $|\sum b_i| \leq \sum |b_i| = |b|$, which means that $|b|^2$ dominates $|\sum b_i|^2$. As a consequence, since T is bounded in L^2 , Tb is the L^2 limit of $T(\sum_{i=0}^n b_i) = \sum_{i=0}^n T(b_i)$.

As a result, we can prove that $|Tb(x)| \leq \sum_{i=0}^{\infty} |Tb_i(x)|$ holds almost everywhere. Indeed, in the case $\sum_{i=0}^{\infty} |Tb_i(x)| = \infty$, there is nothing to prove. On the other hand, if the series converges, we have, for

every $n_j, k > 0$:

$$\left| \sum_{i=1}^{n_j+k} Tb_i(x) - Tb(x) \right| \leq \left| \sum_{i=1}^{n_j} Tb_i(x) - Tb(x) \right| + \left| \sum_{i=n_j+k}^{\infty} Tb_i(x) \right| \leq \left| \sum_{i=1}^{n_j} Tb_i(x) - Tb(x) \right| + \sum_{i=n_j}^{\infty} |Tb_i(x)|.$$

From here, we can see that $\sum_{i=0}^{\infty} T(b_i) = Tb$ holds almost everywhere: since the series converges to Tb in L^2 , there is a subsequence which converges almost everywhere. For every such x choose n_j such that the first term above is less than $\frac{\varepsilon}{2}$ while the tail of the series is also less than $\frac{\varepsilon}{2}$ and we get that (since any $m > n_j$ can be written as $n_j + k$) the series converges to $Tb(x)$, as intended. This means we have, almost everywhere,

$$|Tb(x)| = \left| \sum_{i=0}^{\infty} Tb_i(x) \right| \leq \sum_{i=0}^{\infty} |Tb_i(x)|,$$

as intended. This allows us to bound the integral of each Tb_j separately, as it implies

$$\frac{2}{\lambda} \int_{(\cup_j 2Q_j)^c} |Tb| \leq \frac{2}{\lambda} \int_{(\cup_j 2Q_j)^c} \sum_i |Tb_i| \omega = \frac{2}{\lambda} \sum_i \int_{(\cup_j 2Q_j)^c} |Tb_i| \omega,$$

by the monotone convergence theorem. We are now ready to bound each of these integrals. Since b_i is supported in Q_j , it is an L^2 function of compact support, which means th compact support, and points in $(\cup_j 2Q_j)^c$ are not in the support of b_i , which means we have (by definition of Calderón-Zygmund operators):

$$\frac{1}{\lambda} \int_{(\cup_j 2Q_j)^c} |Tb_i(x)| \omega(x) dx = \frac{1}{\lambda} \int_{(\cup_j 2Q_j)^c} \left| \int_{Q_i} K(x, y) b_i(y) dy \right| \omega(x) dx.$$

At this point, we want to use the fact that b_i has integral 0 (by the same reasoning b does) to introduce a $K(x, c_i)$ term on the integral, with c_i the centre of Q_i . The advantage of this is that we have $|x - c_i| > 2|y - c_i|$, since $2|y - c_i| < 2\frac{\ell(Q_i)}{2} = \ell(Q_i) < |x - c_i|$ (see picture 1.5), which means we can use the property

$$|K(x, y) - K(x, z)| \leq C \frac{|y - z|^\delta}{|x - y|^{n+\delta}}, \quad \text{with } x, y, z = x, c_i, y.$$

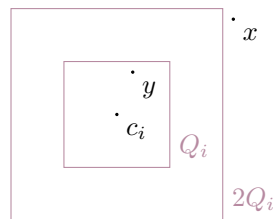


Figure 1.5: Example of a cube Q_i of centre c_i , with a point $x \in (2Q_i)^c$ and a point $y \in Q_i$.

So, we have, using the fact that b_i has integral 0 and $K(x, c_i)$ does not depend on y :

$$\begin{aligned}
\frac{1}{\lambda} \int_{(\cup_j 2Q_j)^c} \left| \int_{Q_i} K(x, y) b_i(y) dy \right| \omega(x) dx &= \frac{1}{\lambda} \int_{(\cup_j 2Q_j)^c} \left| \int_{Q_i} (K(x, y) - K(x, c_i)) b_i(y) dy \right| \omega(x) dx \\
&\leq \frac{1}{\lambda} \int_{(\cup_j 2Q_j)^c} \left(\int_{Q_i} |K(x, y) - K(x, c_i)| |b_i(y)| dy \right) \omega(x) dx \\
&\lesssim \frac{1}{\lambda} \int_{(\cup_j 2Q_j)^c} \left(\int_{Q_i} \frac{|y - c_i|^\delta}{|x - c_i|^{n+\delta}} |b_i(y)| dy \right) \omega(x) dx \\
&= \frac{1}{\lambda} \int_{Q_i} \left(\int_{(\cup_j 2Q_j)^c} \frac{|y - c_i|^\delta}{|x - c_i|^{n+\delta}} \omega(x) dx \right) |b_i(y)| dy,
\end{aligned}$$

where we used Fubini in the last step. Now, we want to bound the inner integral, and to do that we use a very common Calderón-Zygmund operator trick, and split the domain into the regions $2^k \ell(Q_i) < |x - c_i| \leq 2^{k+1} \ell(Q_i)$ (note that $(2Q_i)^c$ is contained in the region $|x - c_i| > \ell(Q_i)$):

$$\begin{aligned}
\int_{(\cup_j 2Q_j)^c} \frac{|y - c_i|^\delta}{|x - c_i|^{n+\delta}} \omega(x) dx &\leq \int_{(2Q_i)^c} \frac{|y - c_i|^\delta}{|x - c_i|^{n+\delta}} \omega(x) dx \\
&\leq \sum_{k=0}^{\infty} \int_{2^k \ell(Q_i) < |x - c_i| < 2^{k+1} \ell(Q_i)} \frac{|y - c_i|^\delta}{|x - c_i|^{n+\delta}} \omega(x) dx \\
&\leq \sum_k \ell(Q_i)^\delta \left(\frac{1}{2^k \ell(Q_i)} \right)^{n+\delta} \int_{2^k \ell(Q_i) < |x - c_i| \leq 2^{k+1} \ell(Q_i)} \omega(x) dx \\
&\leq \sum_k 2^{-k\delta} \left(\frac{1}{2^k \ell(Q_j)} \right)^n \int_{|x - c_i| \leq 2^{k+1} \ell(Q_i)} \omega(x) dx.
\end{aligned}$$

Each of the terms above is now almost an average of ω . Indeed, we have that $\{x : |x - c_i| \leq 2^{k+1} \ell(Q_i)\}$ is contained in the square $2^{k+2} Q_i$ (since we have $\|x - c_i\|_\infty \leq |x - c_i| \leq 2^{k+1} \ell(Q_i) = \frac{2^{k+2}}{2} \ell(Q_i)$, which means $x \in 2^{k+2} Q_i$) and we can make the area of Q_i appear from $(2^k \ell(Q_i))^n$. This allows us to bound each term by $M\omega(y)$ (since $y \in Q_i$ and therefore in the dilations as well), which in turn can be bounded by $\omega(y)$, since $\omega \in A_1$:

$$\begin{aligned}
\sum_k 2^{-k\delta} \left(\frac{1}{2^k \ell(Q_j)} \right)^n \int_{|x - c_i| < 2^{k+1} \ell(Q_i)} \omega(x) dx &= \sum_k 2^{-k\delta} 2^n \left(\frac{1}{2^{k+2} \ell(Q_j)} \right)^n \int_{|x - c_i| < 2^{k+1} \ell(Q_i)} \omega(x) dx \\
&\leq \sum_k 2^{-k\delta} 2^{2n} \frac{1}{|2^{k+2} Q_j|} \int_{2^{k+2} Q_j} \omega(x) dx \\
&\leq \sum_k 2^{-k\delta} 2^{2n} M\omega(y) \lesssim M\omega(y) \lesssim \omega(y).
\end{aligned}$$

Now, using this bound, we can bound the inner integral

$$\frac{1}{\lambda} \int_{Q_i} \left(\int_{(\cup_j 2Q_j)^c} \frac{|y - c_i|^\delta}{|x - c_i|^{n+\delta}} \omega(x) dx \right) |b_i(y)| dy \leq \frac{1}{\lambda} \int_{Q_i} |b_i(y)| \omega(y) dy.$$

Summing this over i , we get

$$\frac{2}{\lambda} \sum_i \int_{(\cup_j 2Q_j)^c} |Tb_i| \omega \leq \frac{2}{\lambda} \sum_i \int_{Q_i} |b_i(y)| \omega(y) dy = \frac{2}{\lambda} \int |b(y)| \omega(y) dy \lesssim \frac{\|f\|_{L^1(\omega)}}{\lambda},$$

which concludes the proof. $\square$

The A_2 Conjecture

After the boundedness in $L^p(\omega)$ was established, an effort started being made to understand its dependence on the A_p characteristic. This remained an open problem until recently, and it was conjectured - the A_2 conjecture - that the following bound held and was sharp in the exponent of $[\omega]_2$:

$$\|Tf\|_{L^2(\omega)} \lesssim [\omega]_2 \|f\|_{L^2(\omega)}.$$

By extrapolation, this result would imply the similar conjecture for A_p ($1 < p < \infty$)

$$\|Tf\|_{L^p(\omega)} \lesssim [\omega]_p^{\max(1, p-1)} \|f\|_{L^p(\omega)}.$$

In later sections, we will use Sparse Domination to prove the A_2 conjecture, as well as the general conjecture for A_p . For now, we will only sketch the classical qualitative proof of this result, in order to showcase the challenges around the dependence of the constant. The full proof can be found in [Duo01], theorem 7.11.

Sketch of Proof (Proposition 1.6.2): Let $\omega \in A_p$. Then, by Corollary 1.5.3 (item 1), we have that $\omega \in A_{\frac{p}{s}}$ for some s (take $s = \frac{p}{q}$, where q is given by the corollary). Assuming without loss of generality that f is a bounded function of compact support, we have the following chain of inequalities

$$\int_{\mathbb{R}^n} |Tf|^p \omega \stackrel{(1)}{\leq} \int_{\mathbb{R}^n} M_d(Tf)^p \omega \stackrel{(2)}{\lesssim} \int_{\mathbb{R}^n} M^\#(Tf)^p \omega \stackrel{(3)}{\lesssim} \int_{\mathbb{R}^n} M(|f|^s)^{\frac{p}{s}} \omega \stackrel{(4)}{\lesssim} \int_{\mathbb{R}^n} |f|^p \omega, \quad \text{q.e.d.} \quad \square$$

Here, $M^\#(f) := \sup_Q \frac{1}{|Q|} \int_Q (f - f_Q)$ (where $f_Q = |Q|^{-1} \int_Q f$) is the mean oscillation of f .

We proceed to say some words about each of these inequalities, specially in regards to the constant:

- (1) This is a known pointwise property of the dyadic maximal function (it is a consequence of the Lebesgue Differentiation Theorem), and therefore has no influence in the constant;
- (2) This inequality is Lemma 7.10 in [Duo01], and requires proving that $M_d(Tf) \in L^{p_0}(\omega)$ with $1 \leq p_0 \leq p$, which is omitted here, but a significant part of [Duo01]'s proof. The lemma is proved using a good- λ inequality between superlevel sets to get an inequality between suitable truncations of the layer-cake representations and then taking the limit. Analysing the proof, one can see that the constant does not depend on ω , only the dimension.
- (3) This inequality is a consequence of the pointwise estimate $M^\#(Tf)(x) \leq C_s M(|f|^s)(x)^{1/s}$ (Lemma 7.9 in [Duo01]). The constant C_s depends on the s for which $\omega \in A_{\frac{p}{s}}$.
- (4) This step follows directly from the $L^{\frac{p}{s}}(\omega)$ boundedness of the maximal function. The constant depends on $[\omega]_{\frac{p}{s}}$ in the same way as the boundedness of the maximal function does.

Looking at the sketch of proof above, we can see that the constant depends on s and $[\omega]_{\frac{p}{s}}$, which is not at all promising if we want to study the dependence on $[\omega]_p$. Firstly, the Reverse Hölder Inequality does not give us enough information on s , nor any indication that it would depend on $[\omega]_p$. Secondly, even though we can bound $[\omega]_p$ by a power of $[\omega]_{\frac{p}{s}}$ (see Proposition 1.2.8, item 2), there is no obvious way to do the converse. These obstacles seem to indicate that new ideas are needed if one wants to prove the A_2 conjecture.

Chapter 2

A Proof of the A_2 Conjecture via Pointwise Sparse Domination

2.1 Sparse and Dyadic Families of Sets

Dyadic grids are not new in harmonic analysis, and have proven to be very helpful, for example in Chapter 1. A big advantage of these grids is that they allow us to easily write sets of interest as disjoint unions of controlled sets. Sometimes, as is the case of what follows, expecting relevant collections of sets to be pairwise disjoint turns out to be too tall an order. This is where the notion of sparseness, a weakening of pairwise disjointness, comes in.

Definition 2.1.1 (Sparse Collection): A family of sets $\mathcal{S} = \{S_\alpha, \alpha \in A\}$ is η -Sparse if for every α there is some set $E(S_\alpha) \subseteq S_\alpha$ such that $|E(S_\alpha)| \geq \eta|S_\alpha|$ and the $E(S_\alpha)$ are pairwise disjoint.

We say that a family of sets is Sparse if it is η -Sparse, for some $0 < \eta \leq 1$.

We sometimes refer to $E(S_\alpha)$ as the E -set of S_α .

It is clear from the definition above that any pairwise disjoint family is sparse, with $\eta = 1$ and $E(S_\alpha) = S_\alpha$. We now show a couple of examples of more interesting sparse families:

Example 2.1.2: Let $\mathcal{S} = \{[0, 2^n[, n \in \mathbb{N}^+\}$. This is a $\frac{1}{2}$ sparse family, with $E([0, 2^n[) = [2^{n-1}, 2^n[$. It is clear that $|E([0, 2^n[)| = 2^n - 2^{n-1} = 2^{n-1} = \frac{1}{2}|[0, 2^n[|$ and that the sets in $\{E([0, 2^n[), n \in \mathbb{N}^+\} = \{[2^{n-1}, 2^n[, n \in \mathbb{N}^+\}$ are pairwise disjoint.

The next few examples will focus on dyadic grids, which will end up being very relevant in our proof of the A_2 conjecture, especially for the simpler case of the Dyadic Maximal Operator. Recall that we are following the notation on [LN19] for dyadic grids.

Example 2.1.3: Generalizing the example above, given a dyadic grid $\mathcal{D}$, any ascending chain $Q_0 \subseteq Q_1 \subseteq \dots \subseteq Q_n \subseteq \dots$ such that $\{Q_n\} \subseteq \mathcal{D}$ is $\frac{2^d-1}{2^d}$ -Sparse, where d is the dimension of the dyadic grid $\mathcal{D}$. Indeed, we can take $E(Q_n) = Q_n \setminus Q_{n-1}$. It is clear that these sets are disjoint and that if k_1, k_2 are the

generations of Q_n, Q_{n-1} , we have $|E(Q_n)| = |Q_n| - |Q_{n-1}| = |Q_n| - 2^{(k_2-k_1)d}|Q_n| \geq |Q_n| - 2^{-d}|Q_n| = \frac{2^d-1}{2^d}|Q_n|$.

Example 2.1.4: Given a dyadic grid $\mathcal{D}$, any collection of dyadic cubes $\mathcal{S}$ with bounded generation is M^{-1} -sparse, with M the number of generations.

Indeed, let k be the minimal generation in $\mathcal{S}$. Divide each cube in $\mathcal{D}_k$ into M disjoint subsets of equal size and assign each one to a generation (see figure 2.1). Then, for each cube Q in $\mathcal{S}$ let $E(Q)$ be the union of these chosen subsets which are contained in Q . Then $|E(Q)| = M^{-1}|Q|$ and the $E(Q)$ are disjoint by construction, since two distinct cubes of the same generation are disjoint and given two cubes of different generations, the intersections of their $E(Q)$ with each cube in $\mathcal{D}_k$ will be disjoint as different generations select different subsets.

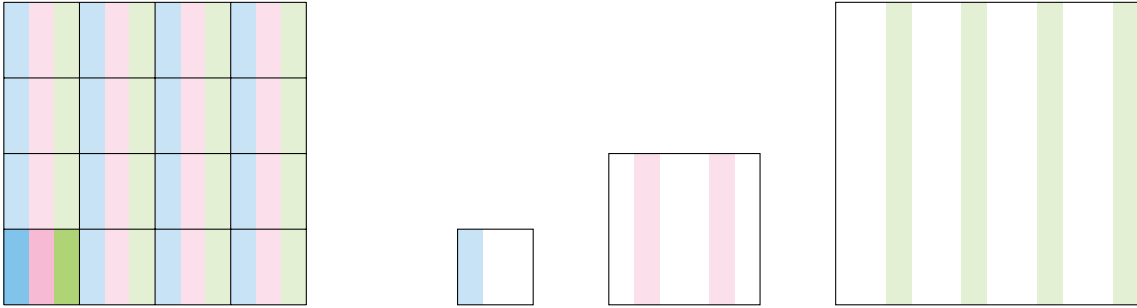


Figure 2.1: Division of $\mathcal{D}_1$ into 3 disjoint sets of equal size (blue, pink and green), followed by $E(Q)$ for a cube Q of generation 1 in blue, $E(Q)$ for a cube of generation 2 in pink and $E(Q)$ for a cube of generation 3 in green.

Before we proceed, it is worthwhile to look at an example of a family which is not Sparse for any $0 < \eta \leq 1$, which motivates the introduction of an equivalent concept which will be useful in proving results on sparse families later in this section.

Example 2.1.5: Let $\mathcal{D}$ be a dyadic grid. Then, $\mathcal{D}$ is not η -Sparse for any $0 < \eta \leq 1$. In order to see it, let us assume $\mathcal{D}$ is η -Sparse and fix a square $Q \in \mathcal{D}$. The main observation is that all of the $E(S)$ with S in smaller generations must be disjoint, have considerable measure and be contained in Q , which we will exploit to get a contradiction. Indeed, for each generation we have $\sum_{S \in \mathcal{D}_k, S \subseteq Q} |E(S)| \geq \eta \sum_{S \in \mathcal{D}_k, S \subseteq Q} |S| = \eta|Q|$. This is a contradiction, as we can't have an infinite amount of disjoint sets (the union of $E(S)$ for each generation) each of size at least $\eta|Q|$ contained in Q , unless $\eta = 0$.

From this example, we can see that if we have Sparse family $\mathcal{S}$ and any element $S \in \mathcal{S}$, we can't have too many sets inside S , as otherwise there would be no room for all the disjoint parts in there. This is the notion of a Carleson family of sets, formalized as follows:

Definition 2.1.6: A collection of sets $\mathcal{S}$ is said to be C -Carleson if for every $S \in \mathcal{S}$, we have

$$\sum_{\tilde{S} \in \mathcal{S}, \tilde{S} \subseteq S} |\tilde{S}| \leq C|S|.$$

The fact that sparse families must be Carleson follows from an argument very similar to the one above. The surprising fact is that the converse also holds, as proven in [Hän18]. We present here a proof for collections of sets in a bounded dyadic grid $\mathcal{D}(\mathcal{Q})$ instead, which will be enough for our purposes.

Lemma 2.1.7: A collection of sets $\mathcal{S} \subseteq \mathcal{D}(\mathcal{Q})$ is η -sparse if and only if it is $\frac{1}{\eta}$ -Carleson.

Proof: For the first direction, let us assume that $\mathcal{S} = \{S_\alpha, \alpha \in A\}$ is an η -sparse collection. Then,

$$\sum_{\tilde{S} \in \mathcal{S}} |\tilde{S}| \leq \frac{1}{\eta} \sum_{\tilde{S} \in \mathcal{S}} |E(\tilde{S})| \leq \frac{1}{\eta} |\mathcal{S}|,$$

since the $E(\tilde{S}) \subseteq \tilde{S} \subseteq S$ and are disjoint.

For the opposite direction, let us first assume that $\mathcal{S}$ is a C -Carleson family of rectangles in a dyadic grid $\mathcal{D}$ with generation bounded below (let k be the minimal generation) and proceed to constructing $\{E(Q) : Q \in \mathcal{S}\}$ generation to generation, increasingly.

Let us consider the cubes in $\mathcal{D}_k \cap \mathcal{S}$. They are pairwise disjoint and since it is the minimal generation they don't contain any other cubes in $\mathcal{D} \cap \mathcal{S}$, so we can safely choose for each $E(Q)$ any set contained in Q such that $|E(Q)| = \frac{1}{C}|Q|$. Now, for each generation, we want to find for each $\tilde{Q}$ a $E(\tilde{Q})$ which is disjoint from all other $E(Q)$ (with Q in any lower generation) and satisfies $|E(\tilde{Q})| = \frac{1}{C}|\tilde{Q}|$. All we need to do is make sure there is enough space left in $\tilde{Q}$ that is not in each of the previously defined $E(Q)$. Note that any $S \subseteq \tilde{Q}$ is of lower generation and therefore has well defined $E(S)$ with the size above and that any two dyadic cubes of the same generation are disjoint, so it is fine to define the $E(\tilde{Q})$ separately. Note also that

$$\sum_{S \subseteq \tilde{Q}, S \in \mathcal{S}} |S| \leq C|\tilde{Q}| \iff \sum_{S \subseteq \tilde{Q}, S \in \mathcal{S}, S \neq \tilde{Q}} |S| \leq (C-1)|\tilde{Q}|,$$

which means that we have

$$|\tilde{Q}| - \sum_{S \subseteq \tilde{Q}, S \in \mathcal{S}, S \neq \tilde{Q}} |E(S)| = |\tilde{Q}| - \frac{1}{C} \sum_{S \subseteq \tilde{Q}, S \in \mathcal{S}, S \neq \tilde{Q}} |S| \geq |\tilde{Q}| - \frac{C-1}{C}|\tilde{Q}| = \frac{1}{C}|\tilde{Q}|,$$

as intended. So, proceeding inductively, we construct $E(S)$ for each $S \in \mathcal{S}$ which satisfy by construction both $|E(S)| \geq \frac{1}{C}|S|$ and that the $E(S)$ are pairwise disjoint, since two dyadic cubes in the same generation were already disjoint, and if two dyadic cubes $Q \subseteq \tilde{Q}$, have different generations the $E(\tilde{Q})$ was constructed as a subset of $\tilde{Q} \setminus E(Q)$. This shows that $\mathcal{S}$ is $\frac{1}{C}$ -Sparse.

In the case where we have infinite cubes, this approach of starting with the smallest cubes and building our way up with the remaining space does not work, and we have no other choice but to start at the biggest generation if we want to iteratively construct the $E(Q)$. Of course, it could happen that our choices in higher generations make it impossible to find $E(Q)$ s for lower generations. We deal with that by going back and changing our $E(Q)$ for higher generations in order to accomodate the lower generations, but we do it in a systematic way which allows us to prove in the end we get good sets. This means that at the step k , we will create sets $E^k(Q)$ for every cube $Q \in \mathcal{D}_j$ with $k \leq j$ and in the end we

will define $E(Q) = \lim_{k \rightarrow \infty} E^k(Q)$ ¹.

Given a cube Q , let $Q^\top$ be the vertex of Q with maximal value in every component², and let Q_t be the cube of size t such that $Q^\top = Q_t^\top$.

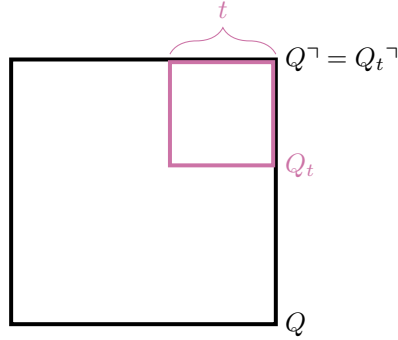


Figure 2.2: Definition of Q_t .

In step k , we define for each $Q \in \mathcal{S} \cap \mathcal{D}_{-k}$ the set $E^k(Q)$ as Q_t , with t_k such that $|Q_t| = \frac{1}{C}|Q|$. Then, from $j = k$ to $j = 0$, we define $E^k(Q)$, for $Q \in \mathcal{D}_j$, as $Q_{t_k} \setminus S$, with

$$S = \cup_{k \geq \ell > j} \cup_{R \in \mathcal{S} \cap \mathcal{D}_{-\ell}, R \subseteq Q} E^k(R)$$

and t such that $|Q_{t_k} \setminus S| = \frac{1}{C}|Q|$. Note that for $t = 0$ we have $|Q_t \setminus S| = 0$ and that $|Q_{\ell(Q)} \setminus S| = |Q \setminus S| \geq \frac{1}{C}|Q|$, by the same reasoning as in the finite case. Moreover, the function $t \mapsto |Q_t \setminus S|$ is continuous by the dominated convergence theorem (using χ_Q as the dominating L^1 function, for example), so we can be sure there exists such a t and in particular that $E^k(Q) \subseteq Q$. Note that t_k is a nondecreasing sequence and that (fixed a k) the $E^k(Q)$ are all pairwise disjoint by construction. Note that we consider $E^k(Q) = \emptyset$ for Q in smaller generations.

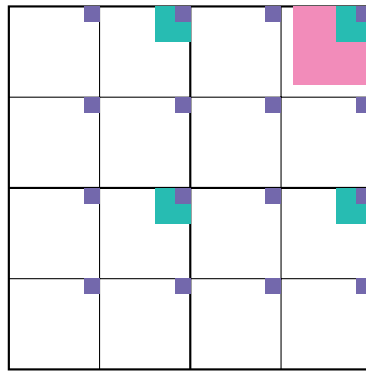


Figure 2.3: Example of a third step of the iteration. The purple sets correspond exactly to Q_t , whereas the green $E^3(Q)$ were extended to account for the blue $E^3(Q)$, and then the pink $E^3(Q)$ were extended to account for the blue and green $E^3(Q)$.

Let $f_k^Q := \chi_{E^k(Q)}$. Then, $f_k^Q(x)$ changes at most twice: from 0 to 1 when t_k becomes large enough

¹We will see that this limit makes sense as the pointwise limit of characteristic functions.

²In dimension 2, $Q^\top$ corresponds to the upper right vertex, as the notation suggests.

so that $Q_{t_k} \ni x$ and from 0 to 1 if x becomes contained in some other $E^\ell(R)$ (from this point on, the other cube R might change to another one in a smaller generation, but by construction x will never be in a $E^{k'}(Q)$ again, for $k' > k$). That means that the sequence f_k^Q converges pointwise, to either 0 or 1, and is dominated by χ_Q . As a result, there is a characteristic function $f^Q = \lim_{k \rightarrow \infty} f_k^Q$ to which the sequence of functions converges in L^1 , whose corresponding set we define as $E(Q)$. Note that we then have $|E(Q)| = \|f^Q\|_1 = \lim_{k \rightarrow \infty} \|f_k^Q\|_1 = \lim_{k \rightarrow \infty} \frac{1}{C}|Q| = \frac{1}{C}|Q|$, as intended. Moreover, since for any point outside Q we have $f_k^Q(x) = 0$, we have $f^Q(x) = 0$ and so $x \notin E(Q)$ and $E(Q) \subseteq Q$ as a result. Finally, we need to check that the $E(Q)$ are disjoint. Indeed, let $x \in E(Q) \cap E(R)$. Then, $f^Q(x) = f^R(x) = 1$, which means that by definition of limit there must be N_Q and N_R such that $f_n^Q(x) = 1$ and $f_n^R(x) = 1$ for $n \geq \max(N_Q, N_R)$ and so $x \in E^n(Q) \cap E^n(R)$, which is impossible, by construction. As a result, we have that $\mathcal{S}$ is $\frac{1}{C}$ -Sparse, with these $E(Q)$. $\square$

We conclude this section with a few lemmas that show that sparseness is reasonably robust when it comes to common transformations of collections of sets.

Lemma 2.1.8 (Dilations of Sparse Families): Let $\mathcal{S} = \{S_\alpha, \alpha \in A\}$ be an η -sparse family in which dilating the sets makes sense³ and $\lambda > 1$ a real number. Then, $\lambda\mathcal{S} := \{\lambda S_\alpha, \alpha \in A\}$ is a $\eta\lambda^{-n}$ -sparse family.

Proof: Since $\mathcal{S}$ is η -sparse, we have $E(S_\alpha)$ such that these are disjoint and $\frac{|E(S_\alpha)|}{|S_\alpha|} \geq \eta$. Let us define $E(\lambda S_\alpha) = E(S_\alpha)$, which makes sense since $E(S_\alpha) \subseteq S_\alpha \subseteq \lambda S_\alpha$. Then, we have $\frac{|E(\lambda S_\alpha)|}{|\lambda S_\alpha|} = \frac{|E(S_\alpha)|}{\lambda^n |S_\alpha|} \geq \eta\lambda^{-n}$ and since the $E(S_\alpha)$ are disjoint the $E(\lambda S_\alpha)$ are too. So, $\lambda\mathcal{S}$ is $\eta\lambda^{-n}$ -sparse. $\square$

Lemma 2.1.9 (Concatenation of Sparse Families): Let $\{S_\alpha\}_\alpha$ be a collection of η_α -Sparse families, each contained in some R_α , with $\{R_\alpha\}_\alpha$ pairwise disjoint. Then, $\cup_{\alpha \in A} S_\alpha$ is $\inf_{\alpha \in A} \eta_\alpha$ -Sparse.

Proof: Let $S_\alpha = \{S_{\alpha,\beta}, \beta \in B_\alpha\}$. We claim that $E(S_{\alpha,\beta})$ does the trick. Indeed, we have that by hypothesis $|E(S_{\alpha,\beta})| \geq \eta_\alpha |S_{\alpha,\beta}| \geq \inf_{\alpha \in A} \eta_\alpha |S_{\alpha,\beta}|$. Moreover, given S_{α_1,β_1} and S_{α_2,β_2} , either $\alpha_1 = \alpha_2$, in which case $E(S_{\alpha_1,\beta_1})$ and $E(S_{\alpha_2,\beta_2})$ are disjoint due to the sparseness of S_{α_1} , or $\alpha_1 \neq \alpha_2$ and $E(S_{\alpha_1,\beta_1})$ and $E(S_{\alpha_2,\beta_2})$ are disjoint since they are contained in respectively R_{α_1} and R_{α_2} , which are disjoint. $\square$

Lemma 2.1.10 (Finite Union of Sparse Families): Let $\mathcal{S}_1, \mathcal{S}_2$ be η_1, η_2 -sparse families. Then, $\mathcal{S}_1 \cup \mathcal{S}_2$ is a $(\eta_1^{-1} + \eta_2^{-1})^{-1}$ -sparse family. As a consequence, finite unions of Sparse families are Sparse.

Proof: Since $\mathcal{S}_1$ is η_1 -sparse, by Lemma 2.1.7 it is $\frac{1}{\eta_1}$ -Carleson. Similarly, $\mathcal{S}_2$ is $\frac{1}{\eta_2}$ -Carleson. That means that

$$\sum_{\tilde{S} \in \mathcal{S}_1 \cup \mathcal{S}_2, \tilde{S} \subseteq S} |\tilde{S}| \leq \sum_{\tilde{S} \in \mathcal{S}_1, \tilde{S} \subseteq S} |\tilde{S}| + \sum_{\tilde{S} \in \mathcal{S}_2, \tilde{S} \subseteq S} |\tilde{S}| \leq \frac{1}{\eta_1}|S| + \frac{1}{\eta_2}|S| = \left(\frac{1}{\eta_1} + \frac{1}{\eta_2}\right)|S|,$$

which means that $\mathcal{S}_1 \cup \mathcal{S}_2$ is $\frac{1}{\eta_1} + \frac{1}{\eta_2}$ -Carleson, which by Lemma 2.1.7 again means that $\mathcal{S}_1 \cup \mathcal{S}_2$ is $(\eta_1^{-1} + \eta_2^{-1})^{-1}$ -sparse, as intended. It is easy to check that this extends to finite unions by standard induction arguments. $\square$

³The fundamental properties of the dilation are $|\lambda S| = \lambda^n |S|$ and $S \subseteq \lambda S$. For our purposes, this will be applied to cubes, where dilating them by λ maintains the centre and dilates the side.

Remark 2.1.11: We can see that finiteness is important in Lemma 2.1.10 by considering a dyadic grid. Each generation is 1-Sparse, since it consists of pairwise disjoint squares, whereas the union of all generations is not Sparse, as seen in Example 2.1.5.

2.2 Strong Weighted Bounds for Sparse Forms

The Sparse Domination technique consists on bounding varied operators by Sparse Forms:

$$\mathcal{A}_S f(x) := \sum_{Q \in S} \langle f \rangle_Q \chi_Q(x),$$

with S a sparse family of cubes, in order to prove certain boundedness properties of our operators of interest.

In this section, we will prove, with explicit dependence on the A_p constant, several boundedness properties of $\mathcal{A}_S$ and with it justify the usefulness of Sparse Domination, as we can then use it to translate these boundedness properties onto a wide range of different operators. We note that the series of lectures [BC20b] by D. Beltran and J. M. Conde Alonso and Rodrigo Duarte's informal lecture notes [Dua22] were very helpful in understanding the content of this section, as well as the remainder of chapter 2.

Before proceeding to prove our main result, we show the unweighted proof of boundedness, which serves as a simpler motivating example for the main proof.

Lemma 2.2.1: For every sparse family S , the operator $\mathcal{A}_S$ is bounded from L^p to L^p , for $1 < p < \infty$, with a constant which depends only on the η constant of sparseness.

Proof: We will argue by duality, that is, we will use

$$\|\mathcal{A}_S f\|_p = \sup_{g \in L^{p'}; \|g\|_{p'}=1} \left| \int_{\mathbb{R}^n} \mathcal{A}_S f \cdot g \right|.$$

Recall that if a function is positive, the maximizer is also positive, which allows us to assume without loss of generality that g is positive and therefore use the monotone convergence theorem:

$$\int \left(\sum_{Q \in S} \langle f \rangle_Q \chi_Q(x) \right) g(x) dx = \sum_{Q \in S} \langle f \rangle_Q \int \chi_Q(x) g(x) dx = \sum_{Q \in S} \langle f \rangle_Q \langle g \rangle_Q |Q|.$$

This last term will be important in what follows and was written like this in order to be symmetric in f and g . In the process, we have made a term $|Q|$ appear, on which we can apply the hypothesis that S is sparse:

$$\sum_{Q \in S} \langle f \rangle_Q \langle g \rangle_Q |Q| \leq \frac{1}{\eta} \sum_{Q \in S} \langle f \rangle_Q \langle g \rangle_Q |E(Q)| = \frac{1}{\eta} \sum_{Q \in S} \int_{E(Q)} \langle f \rangle_Q \langle g \rangle_Q dx.$$

At this point, we can use the fact that $\langle f \rangle_Q \leq Mf(x)$ for every $x \in Q$, that $E(Q) \subseteq Q$ and that the integrand is non-negative as follows:

$$\frac{1}{\eta} \sum_{Q \in \mathcal{S}} \int_{E(Q)} \langle f \rangle_Q \langle g \rangle_Q dx \leq \frac{1}{\eta} \sum_{Q \in \mathcal{S}} \int_{E(Q)} Mf(x) Mg(x) dx \leq \frac{1}{\eta} \int_{\mathbb{R}^n} Mf(x) Mg(x) dx,$$

where we used the fact that the $E(Q)$ are disjoint, as well as the fact that the integrand is non-negative, in the last step. Finally, we can use Hölder and the boundedness of the maximal operator to get

$$\frac{1}{\eta} \int_{\mathbb{R}^n} Mf(x) Mg(x) dx \leq \frac{1}{\eta} \|Mf\|_p \|Mg\|_{p'} \lesssim_\eta \|f\|_p \|g\|_{p'} = \|f\|_p.$$

Taking the sup now gives us the desired bound. $\square$

One of the most important aspects of $\mathcal{A}_\mathcal{S}$ is that a small adaptation of the proof above allows us to prove weighted bounds, with an explicit (oftentimes sharp) dependency on the A_p constant. Before we give the proof for general p , we show here the simpler proof for $\omega \in A_2$ ⁴. In all these weighted proofs, the main ingredient is the following L^p bound of the weighted maximal function:

Lemma 2.2.2: Let $M_\sigma f := \sup_Q \frac{1}{\sigma(Q)} \int_Q f \sigma$, where σ satisfies a doubling condition. Then, we have

$$\|M_\sigma f\|_{L^p(\sigma)} \leq \|f\|_{L^p(\sigma)}.$$

Proof: A direct translation of the proof for unweighted maximal function proves this lemma. $\square$

Recall that all A_p weights satisfy a doubling condition, as a particular case of the second part of Lemma 1.2.2, so the previous lemma applies for A_p weights. This will allow us to manipulate the weights from $L^p(\omega)$ and turn them into A_p characteristics.

Proposition 2.2.3: For every sparse family $\mathcal{S}$, every $\omega \in A_2$ and every $f \in L^2(\omega)$, we have the bound

$$\|\mathcal{A}_\mathcal{S} f\|_{L^2(\omega)} \lesssim [\omega]_2 \|f\|_{L^2(\omega)}.$$

Proof: We begin by following the proof of lemma 2.2.1 exactly, but adding the weight, which is non-negative, so we can still assume $g \geq 0$. That is, we start by using duality to write

$$\|\mathcal{A}_\mathcal{S} f\|_{L^2(\omega)} = \sup_{\|g\|_{L^2(\omega)}=1; g \geq 0} \int \mathcal{A}_\mathcal{S} f \cdot g \cdot \omega,$$

and then simplify the inner expression to get

$$\sum_{Q \in \mathcal{S}} \langle f \rangle_Q \langle g \omega \rangle_Q |Q| = \sum_{Q \in \mathcal{S}} \left(\frac{1}{|Q|} \int_Q f \right) \left(\frac{1}{|Q|} \int_Q g \omega \right) |Q|.$$

In the $p = 2$ case, we have the advantage that $p' = 2$ and so $p' - 1 = 1$ and the A_2 characteristic is just the supremum of $\left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{1}{|Q|} \int_Q \omega^{-1} \right) = \frac{\omega(Q)}{|Q|} \frac{\omega^{-1}(Q)}{|Q|}$. In particular, by proposition 1.2.8 we have

⁴Note that technically this is enough to prove the A_2 conjecture, and also that the dependence on $[\omega]_p$ for more general p would then follow by Rubio de Francia's extrapolation.

that $\omega^{-1} \in A_2$. So, we can just multiply and divide by $\omega(Q)\omega^{-1}(Q)$:

$$\sum_{Q \in \mathcal{S}} \left(\frac{1}{|Q|} \int_Q f \right) \left(\frac{1}{|Q|} \int_Q g\omega \right) |Q| = \sum_{Q \in \mathcal{S}} \left(\frac{\omega^{-1}(Q)}{|Q|} \right) \left(\frac{1}{\omega^{-1}(Q)} \int_Q f\omega\omega^{-1} \right) \left(\frac{\omega(Q)}{|Q|} \right) \left(\frac{1}{\omega(Q)} \int_Q g\omega \right) |Q|.$$

The product of the quotients is by definition bounded by $[\omega]_2$, and the remaining terms are weighted averages (with $\sigma = \omega^{-1}, \omega$) so they are bounded by the maximal function at any point in Q . So, we proceed as before, using sparseness to make an integral over each $E(Q)$ appear and disjointness of $E(Q)$ to bound the sum by an integral on $\mathbb{R}^n$. Then, bounding the averages by the respective weighted maximal functions (and multiplying and dividing by $\omega^{\frac{1}{2}}$ to introduce the right weights in the resulting norms) gives the bound

$$\leq [\omega]_2 \int_{\mathbb{R}^n} M_{\omega^{-1}}(f\omega)(x) \omega^{-\frac{1}{2}}(x) M_{\omega}(g)(x) \omega^{\frac{1}{2}}(x) dx \leq [\omega]_2 \cdot \|M_{\omega^{-1}}(f\omega)\|_{L^2(\omega^{-1})} \cdot \|M_{\omega}(g)\|_{L^2(\omega)}.$$

Using the weighted bound for the maximal function (lemma 2.2.2), this is bounded by:

$$\leq [\omega]_2 \cdot \|f\omega\|_{L^2(\omega^{-1})} \cdot \|g\|_{L^2(\omega)} = [\omega]_2 \cdot \|f\|_{L^2(\omega)} \cdot \|g\|_{L^2(\omega)} = [\omega]_2 \cdot \|f\|_{L^2(\omega)}, \quad \square$$

as intended.

We are finally in place to prove the result for general p , by adapting the $p = 2$ case.

Proposition 2.2.4: For every sparse family $\mathcal{S}$, every $\omega \in A_p$ and every $f \in L^p(\omega)$, we have the bound

$$\|\mathcal{A}_{\mathcal{S}} f\|_{L^p(\omega)} \lesssim [\omega]_p^{\max(1, \frac{1}{p-1})} \|f\|_{L^p(\omega)}.$$

Proof: Restricting to $p > 2$:

Let us start by noting that the dependence on the A_p constant in the proposition, which is the one we expect (and can in fact prove) via Rubio de Francia extrapolation, is $[\omega]_p^{\max(1, \frac{1}{p-1})}$. This means that while for $p > 2$ this is still $[\omega]_p$, for $p < 2$ it becomes $[\omega]_p^{\frac{1}{p-1}}$. If we recall proposition 1.2.8 and in particular remark 1.2.9, this is to be expected, since we have that if $p < 2$ and $\omega \in A_p$, then $\omega^{1-p'} \in A_{p'}$ and we have $[\omega]_p = [\omega^{1-p'}]_{p'}^{p-1}$, so if were to try a duality argument we could obtain $[\omega]_p^{\frac{1}{p-1}}$ as $[\omega^{1-p'}]_{p'}$, which has the “right exponent” since $p' > 2$.

So, we begin the proof exactly with a duality argument which proves the result for $p < 2$ from the result for $p > 2$. Let us then fix some $p < 2$ and assume the result for every p greater than 2. We again use L^p duality, so it is enough to bound, for $\|g\|_{L^{p'}(\omega)} = 1$, $\int_{\mathbb{R}^n} \mathcal{A}_{\mathcal{S}} f g \omega$ by $[\omega]_p^{\frac{1}{p-1}} \|f\|_{L^p(\omega)}$. The idea will be to rewrite the integrand, so that $\mathcal{A}_{\mathcal{S}}$ is applied to $g\omega$, and then use the fact that we are assuming the result holds for $\omega^{1-p'} = \omega^{-\frac{p'}{p}} \in A_{p'}$:

$$\begin{aligned} \int_{\mathbb{R}^n} \mathcal{A}_{\mathcal{S}} f g \omega &= \sum_{Q \in \mathcal{S}} \langle f \rangle_Q \langle g\omega \rangle_Q |Q| = \int_{\mathbb{R}^n} f \mathcal{A}_{\mathcal{S}}(g\omega) = \int_{\mathbb{R}^n} f \omega^{\frac{1}{p}} \mathcal{A}_{\mathcal{S}}(g\omega) \omega^{-\frac{1}{p}} \leq \|f \omega^{\frac{1}{p}}\|_{L^p} \|\mathcal{A}_{\mathcal{S}}(g\omega) \omega^{-\frac{1}{p}}\|_{L^{p'}} \leq \\ &\leq \|f\|_{L^p(\omega)} \|\mathcal{A}_{\mathcal{S}}(g\omega)\|_{L^{p'}(\omega^{-\frac{p'}{p}})} \leq \|f\|_{L^p(\omega)} [\omega^{-\frac{p'}{p}}]_{p'} \|g\omega\|_{L^{p'}(\omega^{-\frac{p'}{p}})} = [\omega]_p^{\frac{1}{p-1}} \|f\|_{L^p(\omega)} \|g\|_{L^{p'}(\omega)} = [\omega]_p^{\frac{1}{p-1}} \|f\|_{L^p(\omega)}, \end{aligned}$$

proving the result holds for $\omega \in A_p$ as well.

Numerology for the weighted averages:

So, from now on we can assume that $p > 2$. Similarly to the proofs above we use duality to reduce the proof to bounding the following expression:

$$\sum_{Q \in \mathcal{S}} \left(\frac{1}{|Q|} \int_Q f \right) \left(\frac{1}{|Q|} \int_Q g \omega \right) |Q| = \sum_{Q \in \mathcal{S}} \frac{1}{|Q|} \left(\int_Q f \right) \left(\int_Q g \omega \right).$$

Before proceeding, it is worth working out the numerology that allows us to apply a similar reasoning to the $p = 2$ case. Since every weight involved is a power of ω , we investigate the powers of ω which makes our weighted maximal functions work well:

$$\frac{1}{\omega^\alpha(Q)} \int_Q f = \frac{1}{\omega^\alpha(Q)} \int_Q f \omega^{-\alpha} \omega^\alpha \leq M_{\omega^\alpha}(f \omega^{-\alpha}); \quad \frac{1}{\omega^\beta(Q)} \int_Q g \omega = \frac{1}{\omega^\beta(Q)} \int_Q g \omega \omega^{-\beta} \omega^\beta \leq M_{\omega^\beta}(g \omega \omega^{-\beta});$$

If we are to follow our previous strategy (without worrying about whether we can introduce these weights for now), we will need a ω^α weight to apply the weighted maximal bound to $M_{\omega^\alpha}(f \omega^{-\alpha})$ and similarly a ω^β for $M_{\omega^\beta}(g \omega \omega^{-\beta})$. These terms will still be weighing the norm after using the bound, so we will be getting $\|f \omega^{-\alpha}\|_{L^p(\omega^\alpha)}$ and $\|g \omega \omega^{-\beta}\|_{L^{p'}(\omega^\beta)}$, which means that if we want to obtain $\|f\|_{L^p(\omega)}$ and $\|g\|_{L^{p'}(\omega)}$ we need to choose α, β such that $-\alpha p + \alpha = 1 \iff \alpha = 1 - p'$ and $p' - \beta p' + \beta = 1 \iff \beta = 1$. This means that we are interested in the weights ω and $\omega^{1-p'}$, which is promising given the definition of $[\omega]_p$.

A small adaptation for $p \neq 2$:

Now, if we try to follow the proof for $p = 2$ exactly, we'd quickly run into issues: in order to introduce the weights necessary to apply the weighted maximal bound, we would need to introduce $\omega^{\frac{\alpha}{p}} \omega^{\frac{\beta}{p'}} = \omega^{\frac{1-p'}{p} + \frac{1}{p'}} = \omega^{1-\frac{p'}{p}}$. In the case $p = 2$, this is 1, which allowed us to simply add this term without issue. When $p > 2$, this is not 1, so we would have an extra weight (its inverse) when trying to apply Hölder to the maximal function. We will be fixing this by reversing the order: instead of bounding the weighted averages by the maximal function and then using Hölder, we first use Hölder (for series!), and only then use this bound. The main advantage is that this allows us to introduce the weight before we introduce a dependence on x (which comes from bounding the (constant) average by the maximal function on x), so the compensating term will just be a (weighted) measure of a set, instead of an inconvenient weight:

$$\begin{aligned} \frac{1}{\omega(Q)} \int_Q g \omega \leq M_\omega g(x) &\implies \left(\frac{1}{\omega(Q)} \int_Q g \omega \right)^{p'} \omega(x) \leq (M_\omega g(x))^{p'} \cdot \omega(x) \implies \\ &\implies \int_S \left(\frac{1}{\omega(Q)} \int_Q g \omega \right)^{p'} \omega(x) dx \leq \int_S M_\omega g(x)^{p'} \omega(x) dx, \end{aligned} \quad (2.1)$$

where the right-hand side is the weighted average we were hoping to get, and the left-hand side is $\omega(S)$ times $\left(\frac{1}{\omega(Q)} \int_Q g \omega \right)^{p'}$, which shows that we can bound the latter term by $(\omega(S))^{-1} \|M_\omega g\|_{L^{p'}(\omega)}$, as intended.

Collecting the measure terms:

Before proceeding, note that the role of $|E(Q)|$ in the $p = 2$ in introducing the integral and ensuring the pieces are disjoint is in this case replaced by making sure that the integrals we introduce in (2.1) are each over $E(Q)$ instead of Q (that is, $S = E(Q)$), so that their sum can be bounded by a single integral in $\mathbb{R}^n$. In particular, this means that the $|E(Q)|$ term does not need to be introduced just yet, so the measure terms so far are just $\frac{1}{|Q|}\omega(Q)\omega^{1-p'}(Q)$, in which $\frac{1}{|Q|}$ is introduced by the sparse forms average and the remaining appear when turning our integrals into weighted averages:

$$\sum_{Q \in S} \frac{1}{|Q|} \left(\int_Q f \right) \left(\int_Q g\omega \right) = \sum_{Q \in S} \frac{1}{|Q|} \omega(Q) \omega^{1-p'}(Q) \left(\frac{1}{\omega^{1-p'}(Q)} \int_Q f \right) \left(\frac{1}{\omega(Q)} \int_Q g\omega \right).$$

Finally, if we apply Hölder to the sum of the weighted averages (disregarding for now the extra measure factors), we get

$$\sum_{Q \in S} \left(\frac{1}{\omega^{1-p'}(Q)} \int_Q f \right) \left(\frac{1}{\omega(Q)} \int_Q g\omega \right) \leq \left(\sum_{Q \in S} \left(\frac{1}{\omega^{1-p'}(Q)} \int_Q f \right)^p \right)^{\frac{1}{p}} \left(\sum_{Q \in S} \left(\frac{1}{\omega(Q)} \int_Q g\omega \right)^{p'} \right)^{\frac{1}{p'}}, \quad (2.2)$$

which means that due to the $\frac{1}{p}, \frac{1}{p'}$ powers we will want to introduce the terms $\omega(E(Q))^{-\frac{1}{p}} \omega(E(Q))^{\frac{1}{p'}}$ and $(\omega^{1-p'}(E(Q)))^{-\frac{1}{p}} (\omega^{1-p'}(E(Q)))^{\frac{1}{p'}}$, the second factor of which will be used to introduce the maximal function.

To sum up, what we need to do now is make sure that

$$|Q|^{-1} \cdot \omega(Q) \cdot \omega^{1-p'}(Q) \cdot \omega(E(Q))^{-\frac{1}{p}} \cdot (\omega^{1-p'}(E(Q)))^{-\frac{1}{p'}} \lesssim_{\phi, Q} [\omega]_p, \quad (2.3)$$

where the notation represents a lack of dependence on both ω and Q .

The proof:

If we prove (2.3), the proof follows:

$$\begin{aligned} \int_{\mathbb{R}^n} \mathcal{A}_S f g \omega &= \sum_{Q \in S} \frac{1}{|Q|} \left(\int_Q f \right) \left(\int_Q g\omega \right) = \\ &= \sum_{Q \in S} \frac{\omega(Q)}{|Q|} \omega^{1-p'}(Q) \omega(E(Q))^{-\frac{1}{p}} (\omega^{1-p'}(E(Q)))^{-\frac{1}{p'}} \left(\frac{\int_Q f}{\omega^{1-p'}(Q)} (\omega^{1-p'}(E(Q)))^{\frac{1}{p'}} \right) \left(\frac{\int_Q g\omega}{\omega(Q)} \omega(E(Q))^{\frac{1}{p}} \right) \stackrel{(2.3)}{\lesssim_{\phi, Q}} \\ &\lesssim_{\phi, Q} [\omega]_p \sum_{Q \in S} \left(\frac{\int_Q f}{\omega^{1-p'}(Q)} (\omega^{1-p'}(E(Q)))^{\frac{1}{p'}} \right) \left(\frac{\int_Q g\omega}{\omega(Q)} \omega(E(Q))^{\frac{1}{p}} \right) \stackrel{(2.2)}{\leq} \\ &\leq [\omega]_p \left(\sum_{Q \in S} \left(\frac{\int_Q f}{\omega^{1-p'}(Q)} (\omega^{1-p'}(E(Q)))^{\frac{1}{p'}} \right)^p \right)^{\frac{1}{p}} \left(\sum_{Q \in S} \left(\frac{\int_Q g\omega}{\omega(Q)} \omega(E(Q))^{\frac{1}{p}} \right)^{p'} \right)^{\frac{1}{p'}} \stackrel{(2.1)}{\leq} \\ &\leq [\omega]_p \left(\sum_{Q \in S} \int_{E(Q)} M_{\omega^{1-p'}}(f \omega^{p'-1})^{p'} \omega^{1-p'} \right)^{\frac{1}{p}} \left(\sum_{Q \in S} \int_{E(Q)} M_{\omega} g^{p'} \omega \right)^{\frac{1}{p'}}. \end{aligned}$$

At this point, the numerology works out as intended due to our choice of α, β above. Using the fact

that the $E(Q)$ are disjoint, we have the bound

$$[\omega]_p \|M_{\omega^{1-p'}}(f\omega^{p'-1})\|_{L^p(\omega^{1-p'})} \|M_\omega g\|_{L^{p'}(\omega)} \stackrel{2.2.2}{\leq} [\omega]_p \|f\omega^{p'-1}\|_{L^p(\omega^{1-p'})} \|g\|_{L^{p'}(\omega)} = [\omega]_p \|f\|_{L^p(\omega)}.$$

So, all that we need to do now is prove (2.3), and the result follows.

Proving (2.3):

Recall that

$$[\omega]_p := \sup_Q \left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{1}{|Q|} \int_Q \omega^{1-p'} \right)^{p-1} = \sup_Q \left(\frac{1}{|Q|} \int_Q \omega \right) \left(\frac{1}{|Q|} \int_Q \omega^{-\frac{p'}{p}} \right)^{\frac{p}{p'}}.$$

and that we want to prove

$$|Q|^{-1} \omega(Q) \omega^{1-p'}(Q) \omega(E(Q))^{-\frac{1}{p'}} \left(\omega^{1-p'}(E(Q)) \right)^{-\frac{1}{p}} \lesssim_{\phi, \mathcal{Q}} [\omega]_p.$$

Let us use the notation $\tilde{\omega} := \omega^{1-p'}$. If we had Q instead of $E(Q)$ in the last two terms, the left-hand-side would simplify to

$$\frac{\omega(Q)^{\frac{1}{p}} \tilde{\omega}(Q)^{\frac{1}{p'}}}{|Q|} = \left(\frac{\omega(Q)}{|Q|} \left(\frac{\tilde{\omega}(Q)}{|Q|} \right)^{\frac{p}{p'}} \right)^{\frac{1}{p}} \leq [\omega]_p^{\frac{1}{p}},$$

which indicates that we should expect a term bounded by $[\omega]_p^{\frac{1}{p'}}$ to appear in the process of removing the dependence on the E -sets. In particular, we would like to prove

$$\omega(E(Q))^{-\frac{1}{p'}} \tilde{\omega}(E(Q))^{-\frac{1}{p}} \lesssim_{\phi, \mathcal{Q}} \frac{\omega(Q)^{\frac{1}{p'}} \tilde{\omega}(Q)^{\frac{p}{p'} - \frac{1}{p'}}}{|Q|^{\frac{p}{p'}}} \omega(Q)^{-\frac{1}{p'}} \tilde{\omega}(Q)^{-\frac{1}{p}},$$

since this removes the $E(Q)$ and introduces something bounded by $[\omega]_p^{\frac{1}{p'}}$. Simplifying the expression above gives us the endgoal

$$|Q|^{\frac{p}{p'}} \lesssim_{\phi, \mathcal{Q}} \omega(E(Q))^{\frac{1}{p'}} \tilde{\omega}(E(Q))^{\frac{1}{p}} \tilde{\omega}(Q)^{\frac{p}{p'} - 1}, \quad (2.4)$$

which we can prove using the fact that Hölder allows us to write

$$|E(Q)| = \int_{E(Q)} \omega^{\frac{1}{p}} \omega^{-\frac{1}{p}} \leq \omega(E(Q))^{\frac{1}{p}} \left(\omega^{-\frac{p}{p'}}(E(Q)) \right)^{\frac{1}{p'}} = \omega(E(Q))^{\frac{1}{p}} \tilde{\omega}(E(Q))^{\frac{1}{p'}},$$

which means that using sparseness we have

$$|Q|^{\frac{p}{p'}} \leq \frac{1}{\eta^{\frac{p}{p'}}} |E(Q)|^{\frac{p}{p'}} \lesssim_{\phi, \mathcal{Q}} \omega(E(Q))^{\frac{1}{p'}} \tilde{\omega}(E(Q))^{\frac{p}{p'} - \frac{1}{p'}} = \omega(E(Q))^{\frac{1}{p'}} \tilde{\omega}(E(Q))^{\frac{p}{p'} - 1} \tilde{\omega}(E(Q))^{\frac{1}{p}}.$$

Now, since $p > 2$, we have that $\frac{p}{p'} - 1 > 0$, which means that we can use the fact that $\tilde{\omega}(E(Q)) \leq \tilde{\omega}(Q) \implies \tilde{\omega}(E(Q))^{\frac{p}{p'} - 1} \leq \tilde{\omega}(Q)^{\frac{p}{p'} - 1}$, which finishes the proof since it establishes (2.4). $\square$

2.3 The Maximal Operator

Before seeing useful sparse domination results, it is worthwhile to study the sparse domination of the maximal operator, as it captures the essence of the technique without the difficulties which arise when dealing with more complicated operators.

The proof follows by an iteration argument. At each step, we add a relevant cube to our sparse collection, which restricts the set of points we still need to concern ourselves with, as all the points on which Mf is bounded by the average over that cube become taken care of. In the process, we use the Calderón-Zygmund decomposition to both control the measures of the sets (which will be key in ensuring our collection will be sparse) and neatly organize the leftover space into cubes, on each of which we can apply the same trick, and so on.

2.3.1 From the Dyadic to the Continuous

Even with its high applicability, Calderón-Zygmund decomposition shines its brightest in the dyadic context, since it exploits the dyadic grid and the properties of the associated dyadic maximal operator. As a result, our first step will be to show that in order to bound the maximal operator it is enough to bound the dyadic maximal operator. Note that even though the converse is obvious, as we have $M_{\mathcal{D}}f \leq CMf$, this direction requires more work. An easy way to see that is recalling that the maximal function is only 0 when f is 0. However, it is possible that the standard dyadic maximal function is 0 in a point. For instance, if f is the characteristic function of the first quadrant, all points in the remaining quadrants satisfy $M_{\mathcal{D}}f(x) = 0$, so there is no hope for such a bound in the opposite direction. Even though this problem is specific of multiquadrant dyadic grids, it is nonetheless easy to construct arbitrarily small cubes or balls with a common smallest dyadic cube which contains them, like in the following picture:

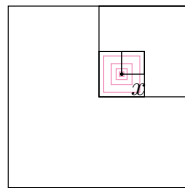


Figure 2.4: In black the dyadic cubes containing x . In pink arbitrarily small cubes all such that the smallest cube which contains them is the second biggest black square.

As a result, relating these maximal functions requires the so-called Three Lattice Lemma, which hinges on the following observation:

Lemma 2.3.1: Given any cube Q , let $\tilde{Q}$ be the dyadic cube $\tilde{Q}$ which contains $c(Q)$ and satisfies $l(\tilde{Q}) < l(Q) \leq 2l(\tilde{Q})$. Then, $3\tilde{Q}$ contains Q .

Proof: Note that $Q = \{x : \|x - c(Q)\|_{\infty} \leq \frac{l(Q)}{2}\}$. We wish to show that Q is contained in $3\tilde{Q} = \{x :$

$\|x - c(3\tilde{Q})\|_\infty \leq \frac{3l(\tilde{Q})}{2}$, and so we want to bound $\|x - c(3\tilde{Q})\|_\infty$ for $x \in Q$:

$$\|x - c(3\tilde{Q})\|_\infty = \|x - c(\tilde{Q})\|_\infty \leq \|x - c(Q)\|_\infty + \|c(Q) - c(\tilde{Q})\|_\infty \leq \frac{l(Q)}{2} + \frac{l(\tilde{Q})}{2} \leq \frac{2l(\tilde{Q})}{2} + \frac{l(\tilde{Q})}{2} = \frac{3l(\tilde{Q})}{2} = \frac{l(3\tilde{Q})}{2},$$

where we used the fact that $c(Q) \in \tilde{Q} = \{x : \|x - c(\tilde{Q})\|_\infty \leq \frac{l(\tilde{Q})}{2}\}$, the assumption that $l(Q) \leq 2l(\tilde{Q})$ and the definition of $3\tilde{Q}$. $\square$

So, given any cube Q , we can find a cube $3\tilde{Q}$ which contains it and such that $\frac{3|Q|}{|Q|}$ is bounded (by 3^n). Recall that if $3\tilde{Q}$ was a dyadic cube, this would be enough to relate the maximal functions, as

$$\frac{1}{|Q|} \int_Q |f| = \frac{|3\tilde{Q}|}{|Q|} \frac{1}{|3\tilde{Q}|} \int_Q |f| \leq 3^n \frac{1}{|3\tilde{Q}|} \int_{3\tilde{Q}} |f|.$$

It is clear that, given a dyadic grid $\mathcal{D}$, the collection $\tilde{\mathcal{D}} := \{3Q : Q \in \mathcal{D}\}$ is not a dyadic grid: there are too many weird intersections. However, $\tilde{\mathcal{D}}$ is the next best thing, as it is the union of 3^n dyadic grids, which will allow us to compare the maximal function with not one but 3^n dyadic maximal functions. We will see this is nonetheless enough for all our purposes.

Proposition 2.3.2: Given a dyadic grid $\mathcal{D}$, the collection $\tilde{\mathcal{D}} := \{3Q : Q \in \mathcal{D}\}$ is the union of 3^n dyadic grids such that each $3Q$ is in exactly one of the dyadic grids.

Sketch of Proof: The proof of this result can be found for example in [LN19], Theorem 3.1. The following picture should hopefully help illustrate why it should be true:

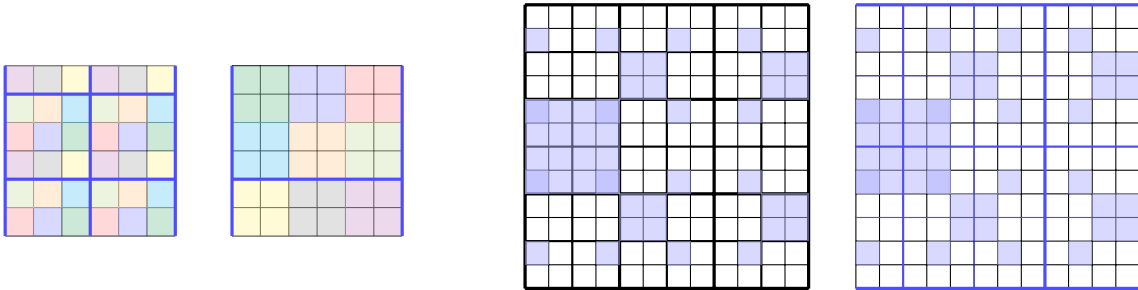


Figure 2.5: In the first and second picture we have, side by side, the same section of a grid, with focus on two consecutive generations. Each color is associated to a different tripled dyadic grid: tripling every square of a given color gives us one of the 9 dyadic grids, with the one associated to the darker blue squares drawn in dark blue for reference. In the third and fourth picture, we focus on only one color, but add one more generation. In black is the original grid, and in blue the tripled grid associated to the blue squares.

$\square$

This is enough to show that proving Sparse Domination of the dyadic maximal function implies the same result for the maximal function. Indeed, we have that since each average $\langle f \rangle_Q$ is bounded by some $3^n \langle f \rangle_{\tilde{Q}}$, with $\tilde{Q}$ in one of the 3^n dyadic grids, and this quantity is in turn bounded by $3^n M_{\mathcal{D}_i} f$ and

by $3^n \sum_{0 \leq i < 3^n} M_{\mathcal{D}_i} f$, so if we take the supremum we get

$$Mf(x) \leq 3^n \sum_{0 \leq i < 3^n} M_{\mathcal{D}_i} f(x) \lesssim \sum_{0 \leq i < 3^n} \sum_{Q \in \mathcal{S}_i} \langle f \rangle_Q \chi_Q(x) = \sum_{Q \in \cup_{0 \leq i < 3^n} \mathcal{S}_i} \langle f \rangle_Q \chi_Q(x),$$

where we used the fact that the sum is finite to take the maximum of the associated constants, and Lemma 2.1.10 ensures that $\cup_{0 \leq i < 3^n} \mathcal{S}_i$ is still sparse. Note that the last step is not undercounting since the different $\mathcal{S}_i$ don't have elements in common.

2.3.2 Sparse Domination of the Dyadic Maximal Function

In this subsection, we prove the main proposition of this section, the pointwise sparse domination of $M_{\mathcal{D}} f$ (and consequently Mf).

Proposition 2.3.4: Let f be a finitely supported locally integrable function and $\mathcal{D}$ a one quadrant dyadic grid $\mathcal{D} = \mathcal{D}(Q_0)$, with Q_0 containing the support of f . Then, there is a sparse family $\mathcal{S} \subseteq \mathcal{D}$ such that

$$M_{\mathcal{D}} f \leq C \sum_{Q \in \mathcal{S}} \langle f \rangle_Q.$$

Proof: Since f is compactly supported and $\mathcal{D}$ is a one quadrant dyadic grid, we can find $Q_0 \in \mathcal{D}$ which contains $\text{supp } f$. In order to prove the result, we will be adding strategic cubes to $\mathcal{S}$ until for almost every point there is a cube whose associated average bounds the maximal function at that point, ensuring in the end that $\mathcal{S}$ is still sparse. Our first step of the proof will be to prove that it is enough to worry about the points inside Q_0 .

Reduction to a cube:

Let $Q_0 \subseteq Q_1 \subseteq \dots \subseteq Q_n \subseteq \dots$ be the predecessors of Q_0 . We will show that as long as we add Q_i to $\mathcal{S}$ then $M_{\mathcal{D}} f$ at every point in Q_0^c is bounded by one of the associated sparse forms. Note that by Example 2.1.3 this collection is $\frac{2^d-1}{2^d}$ -sparse.

Recall that in dyadic grids if two distinct cubes intersect then one is contained in the other. For a cube to contain a point in Q_0^c , it can't be contained in Q_0 and therefore must contain it. So, for a point x in Q_0^c , the only dyadic cubes which contain it and have non-zero average are the ones that also contain x , that is, whichever Q_i contain x . Moreover, since f is zero outside Q_0 , the greatest average will be the associated to the smallest such Q_j containing x (see picture 2.6), so we have $M_{\mathcal{D}} f(x) = \langle f \rangle_{Q_j}$. As a result, $M_{\mathcal{D}} f(x)$ (for $x \in Q_0^c$) is bounded by our sparse form as long as $C \geq 1$.

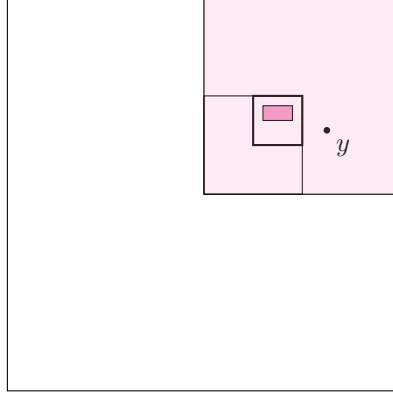


Figure 2.6: A section of one-quadrant dyadic grid. In dark pink, the support of f , in thick black Q_0 and in thin black its ancestors. For the point y outside, the smallest dyadic square containing x and intersecting the support of f is indicated in light pink.

Step 0:

Now that we have restricted our attention to Q_0 , our next step is to add it to $\mathcal{S}$ and investigate the points which are still not bounded by our sparse form, choosing an appropriate constant so that this set has a reasonable size. In particular, we have that the set $G_0 := \{M_{\mathcal{D}}f(x) \leq C\langle f \rangle_{Q_0}\}$ consists by definition of “good points”, in which the maximal function bounded by C times the sparse form, so we are looking to study the size of the set of “bad points” $B_0 := \{M_{\mathcal{D}}f(x) > C\langle f \rangle_{Q_0}\}$, which can easily be bounded using the weak $(1, 1)$ bound of the maximal function:

$$|\{x \in Q_0 : M_{\mathcal{D}}f(x) > C\langle f \rangle_{Q_0}\}| \leq C_{M_{\mathcal{D}}} \frac{\int_{Q_0} f}{C\langle f \rangle_{Q_0}} = \frac{C_{M_{\mathcal{D}}}|Q_0|}{C}.$$

If we choose $C = 2C_{M_{\mathcal{D}}}$, then this set has measure at most $\frac{1}{2}|Q_0|$, which means that if we take $E(Q_0) = G_0 = \{x \in Q_0 : M_{\mathcal{D}}f(x) \leq C\langle f \rangle_{Q_0}\}$ and choose the remaining elements of $\mathcal{S}$ such that they are contained in $B_0 = Q_0 \setminus E(Q_0)$, we are on our way to get a $\frac{1}{2}$ -sparse family, as $|E(Q_0)| \geq \frac{1}{2}|Q_0|$ and it will be disjoint from whatever other E -sets we choose.

Step 1:

The goal now is to find a way to proceed iteratively, starting from the observation that the set B_0 is exactly the “bad set” of the Calderón-Zygmund decomposition and therefore can be written as a (disjoint) union of squares $Q_{1,i}$ in $\mathcal{D}$, in each of which we can hope to apply step 0. Note that if we were to repeat the weak-bound argument, we would run into $\frac{\int_{Q_0} f}{\int_{Q_{1,i}} f}$, which doesn’t quite simplify. We will see that this is easy to fix, by exploiting the stopping-time argument within the Calderón-Zygmund decomposition to show that $M_{\mathcal{D}}f(x) = M_{\mathcal{D}}(f\chi_{Q_{1,i}})(x)$ for points in $Q_{1,i}$, in which case the $\chi_{Q_{1,i}}$ factor restricts the integral in the numerator into the integral in the denominator and we can proceed as in step 0.

Indeed, recall that by construction, $Q_{1,i}$ are the maximal cubes which contain each $x \in B_0$ and have average greater than $C\langle f \rangle_Q$. This means that in particular any $\tilde{Q} \supseteq Q_0$ must satisfy $\langle f \rangle_{\tilde{Q}} \leq C\langle f \rangle_Q < \langle f \rangle_{Q_0}$ and so given a point x in $Q_{1,i}$ (i arbitrary) in order to compute $M_{\mathcal{D}}f(x)$ it is enough to consider cubes contained in $Q_{1,i}$. Indeed, since the intersection of any cube containing x and $Q_{1,i}$

would contain x , and dyadic cubes which intersect must be contained in each other, any relevant cube is either contained in $Q_{1,i}$ or it contains $Q_{1,i}$, in which case its average is smaller than the average over $Q_{1,i}$, making it redundant for the supremum. In particular, this shows that for every point in $Q_{1,i}$ we do have $M_{\mathcal{D}}f(x) = M_{\mathcal{D}}(f\chi_{Q_{1,i}})(x)$, which means that using the weak bound we have, defining $E(Q_{1,i}) = \{x \in Q_{1,i} : M_{\mathcal{D}}f(x) \leq C\langle f \rangle_{Q_{1,i}}\}$,

$$\begin{aligned} |Q_{1,i} \setminus E(Q_{1,i})| &= |\{x \in Q_{1,i} : M_{\mathcal{D}}f(x) > C\langle f \rangle_{Q_{1,i}}\}| = \\ &= |\{x \in Q_{1,i} : M_{\mathcal{D}}(f\chi_{Q_{1,i}})(x) > C\langle f \rangle_{Q_{1,i}}\}| \leq C_{M_{\mathcal{D}}} \frac{\int_{Q_{1,i}} f}{C\langle f \rangle_{Q_{1,i}}} = \frac{C_{M_{\mathcal{D}}} |Q_{1,i}|}{C}, \end{aligned}$$

which by choosing $C = 2C_{M_{\mathcal{D}}}$ gives us as in step 0 that the measure of this set is at most $\frac{1}{2}|Q_{1,i}|$ and so the measure of $E(Q_{1,i})$ is at least $\frac{1}{2}|Q_{1,i}|$. As a result, our set of “new good points” G_1 will equal $\dot{\cup}_i E(Q_{1,i})$, and B_1 , the set of points we have yet to bound, will be $B_0 \setminus G_1 = \dot{\cup}_i E(Q_{1,i})^c$.

In short, we have decomposed $Q_0 = G_0 \dot{\cup} G_1 \dot{\cup} B_1$, with $|B_1| = \sum_i |E(Q_{1,i})^c| \leq \sum_i \frac{1}{2}|Q_{1,i}| = \frac{1}{2}|B_0| \leq \frac{1}{4}|Q_0|$, and added $\{Q_{1,i}\}_i$, which is $\frac{1}{2}$ -sparse and contained in B_1^c , to $\mathcal{S}$.

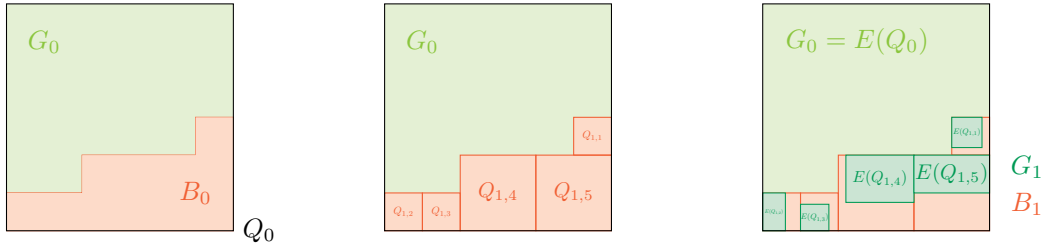


Figure 2.7: Illustration of the first two steps. In light green G_0 as defined above, the set of “good points”. The bad set B_1 in orange is then written as a disjoint union of dyadic squares. For each of which, we define $E(Q_{1,i})$ consisting of the “good points” in that square. We then define G_1 as the union of these new good points, and $B_1 = B_0 \setminus G_1$. In these steps, we added Q_0 and $Q_{1,1}, \dots, Q_{1,5}$ to $\mathcal{S}$, with the corresponding E -sets represented in green. Note that these need not be rectangles and will usually in fact be more complicated sets.

The full iteration:

By this point, we have used step 0 and step 1 to decompose $Q_0 = G_0 \dot{\cup} G_1 \dot{\cup} B_1$, and have added Q_0 and $\{Q_{1,n}\}_n$ into our sparse family, with $E(Q_0) = G_0$, $E(Q_{1,n}) \subseteq G_1$ for every n , $|E(Q_0)| \geq \frac{1}{2}|Q_0|$, $|E(Q_{1,n})| \geq \frac{1}{2}|Q_{1,n}|$ for every n and $|B_1| \leq \frac{1}{2}|B_0|$.

So, in step $i+1$, we assume to have constructed for each $0 \leq j \leq i$ a collection of pairwise disjoint dyadic cubes $\{Q_{j,n}\}_n$ and corresponding $E(Q_{j,n})$ which satisfy

$$E(Q_{j,n}) = \{x \in Q_{j,n} : M_{\mathcal{D}}f(x) \leq C\langle f \rangle_{Q_{j,n}}\} \quad \text{and} \quad |E(Q_{j,n})| \geq \frac{1}{2}|Q_{j,n}|,$$

such that $G_j := \dot{\cup}_n E(Q_{j,n})$ (note then that $E(Q_{j,n}) \subseteq G_j$) and $B_j := \dot{\cup}_n (Q_{j,n} \setminus E(Q_{j,n}))$ satisfy $B_{j-1} = B_j \dot{\cup} G_j$ (in particular, this means that $B_{j-1} = \dot{\cup}_n Q_{j,n}$, so $Q_{j,n} \subseteq B_{j-1}$). Note that by induction, this gives

us the decomposition $Q_0 = G_0 \dot{\cup} \dots \dot{\cup} G_i \dot{\cup} B_i$, out of which only the points in B_i are left to bound by an average in our sparse form.

So, our goal at step $i + 1$ is to find $\{Q_{i+1,n}\}_n$ which satisfy the same properties. We proceed as in step 1, using the fact that by construction we have

$$B_i = \dot{\cup}_m (Q_{i,m} \setminus E(Q_{i,m})) := \dot{\bigcup}_m \{x \in Q_{i,m} : Mf(x) > C\langle f \rangle_{Q_{j,n}}\},$$

which means that we can apply Calderón-Zygmund decomposition to write each of the $Q_{i,m} \setminus E(Q_{i,m})$ as a disjoint union of maximal dyadic cubes. The collection of these cubes over all m is what we define as $\{Q_{i+1,m}\}_m$. For each $Q_{i+1,m}$ we define the E -set as usual:

$$E(Q_{i+1,m}) := \{x \in Q_{i+1,m} : M_{\mathcal{D}}f(x) \leq C\langle f \rangle_{Q_{i+1,m}}\},$$

with C still equal to $2C_{M_{\mathcal{D}}}$. Note that just as in step 1, we defined $Q_{i+1,m}$ as a maximal cube via the Calderón-Zygmund decomposition, which means that for every point $x \in Q_{i+1,m}$ we have $M_{\mathcal{D}}f(x) = M_{\mathcal{D}}(f\chi_{Q_{i+1,m}})(x)$, which allows us to use the weak bound to ensure that

$$|Q_{i+1,m} \setminus E(Q_{i+1,m})| = |\{x \in Q_{i+1,m} : M_{\mathcal{D}}(f\chi_{Q_{i+1,m}})(x) > C\langle f \rangle_{Q_{i+1,m}}\}| \leq C_{M_{\mathcal{D}}} \frac{\|f\chi_{Q_{i+1,m}}\|_1}{2C_{M_{\mathcal{D}}}\langle f \rangle_{Q_{i+1,m}}} = \frac{|Q_{i+1,m}|}{2},$$

ensuring that indeed $|E(Q_{i+1,m})| \geq \frac{1}{2}|Q_{i+1,m}|$ for every m . Defining $G_{i+1} := \dot{\cup}_m E(Q_{i+1,m})$ and $B_{i+1} := \dot{\cup}_m (Q_{i+1,m} \setminus E(Q_{i+1,m}))$ we have that indeed $B_i = B_{i+1} \dot{\cup} G_{i+1}$ holds by construction, which means our iterative step was successful.

Verifications:

In order to show that $\mathcal{S}$ is a $\frac{1}{2}$ -sparse family, it is enough to check that the E -sets are disjoint, since by construction $|E(Q_i)| \geq \frac{2^n-1}{2^n}|Q_i| \geq \frac{1}{2}|Q_i|$ and $|E(Q_{i,n})| \geq \frac{1}{2}|Q_{i,n}|$. This is essentially true by construction, but we take note of the cases:

- Given Q_i and Q_j ($i \neq j$), the fact that $E(Q_i)$ and $E(Q_j)$ are disjoint was ensured in the reduction to a cube portion of the proof;
- Given Q_i and $Q_{j,n}$, we have that $E(Q_i) \subseteq Q_0^c$ and $E(Q_{j,n}) \subseteq Q_{j,n} \subseteq G_j \subseteq Q_0$, which means that $E(Q_i)$ and $E(Q_{j,n})$ are disjoint;
- Given $Q_{i,n}$ and $Q_{i,m}$ with $n \neq m$, we have that by construction the cubes are disjoint, since we wrote $B_i = \dot{\cup}_k Q_{i,k}$, so in particular their E -sets are as well;
- Finally, given $Q_{i,n}$ and $Q_{j,m}$, if (without loss of generality) $i > j$ we have that by construction $E(Q_{j,m}) \subseteq G_j$, whereas $E(Q_{i,n}) \subseteq Q_{i,n} \subseteq B_i$. Since we ensured $Q_0 = G_0 \dot{\cup} \dots \dot{\cup} G_j \dot{\cup} \dots \dot{\cup} G_i \dot{\cup} B_i$, in particular we have that G_j and B_i are disjoint, which implies that $E(Q_{i,n})$ and $E(Q_{j,m})$ are as well.

Now that we have made sure $\mathcal{S}$ is $\frac{1}{2}$ -sparse, all we need to do is show that the sparse form associated to $\mathcal{S}$ indeed bounds the Maximal function almost everywhere. Recall that we proved the result for Q_0^c in

the beginning of the proof for $C = 1$, and for each $E(Q_{i,n})$ we have, by definition, the result for $C = 2C_{\mathcal{D}}$, since if $M_{\mathcal{D}}f(x)$ is bounded by some $C\langle f \rangle_{\tilde{Q}}$ then it is bounded by the sparse form, since every sparse form is positive.

So, to get the result with $C = \max(1, 2C_{\mathcal{D}})$ it is enough to show that almost every $x \in Q_0$ is in some $E(Q_{i,n})$, or equivalently that $x \in G_i$ for some i . Since by construction we have that $Q_0 = G_0 \dot{\cup} G_1 \dot{\cup} \dots \dot{\cup} G_M \dot{\cup} B_M$, for a point to not be in any G_n , it must be in every B_M , so in $\cap_M B_M$. Now, recall that $|E(Q_{M,n})| \geq \frac{1}{2}|Q_{M,n}|$ and $|Q_{M,n} \setminus E(Q_{M,n})| \leq \frac{1}{2}|Q_{M,n}|$, so

$$|B_M| = \sum_m |Q_{M,m} \setminus E(Q_{M,m})| \leq \frac{1}{2} \sum_m |Q_{M,m}| = \frac{1}{2}|B_{M-1}| \implies |B_M| \leq \frac{1}{2^M}|B_0|.$$

As a result, we have $|B_M| \rightarrow 0$, and since $|\cap_M B_M| \leq |B_M|$ this implies $|\cap_M B_M| = 0$, which proves the sparse bound almost everywhere, as intended. $\square$

Remark 2.3.5: In this proof, we decomposed the cube Q_0 into successive “good parts”, G_j , while ensuring that our “bad parts” B_j would become increasingly small. An alternative (but entirely equivalent) point of view would be to think about it as decomposing $T(f\chi_{Q_0})\chi_{Q_0}$. With this notation, the steps above could be written as follows:

$$\begin{aligned} T(f\chi_{Q_0})\chi_{Q_0} &\stackrel{\text{step 0}}{=} T(f\chi_{Q_0})\chi_{G_0} + T(f\chi_{Q_0})\chi_{Q_0 \setminus G_0} \leq C\langle f \rangle_{Q_0}\chi_{Q_0} + T(f\chi_{Q_0})\chi_{Q_0 \setminus G_0} \stackrel{\text{step 1}}{=} \\ &\stackrel{\text{step 1}}{=} C\langle f \rangle_{Q_0}\chi_{Q_0} + \sum_n T(f\chi_{Q_0})\chi_{Q_{1,n}} \stackrel{\star}{=} C\langle f \rangle_{Q_0}\chi_{Q_0} + \sum_n T(f\chi_{Q_{1,n}})\chi_{E(Q_{1,n})} + \sum_n T(f\chi_{Q_{1,n}})\chi_{Q_{1,n} \setminus E(Q_{1,n})} \leq \\ &\leq C\langle f \rangle_{Q_0}\chi_{Q_0} + C \sum_n \langle f \rangle_{Q_{1,n}}\chi_{Q_{1,n}} + \sum_n T(f\chi_{Q_{1,n}})\chi_{Q_{1,n} \setminus E(Q_{1,n})}. \end{aligned}$$

where the step marked with a $\star$ contains the subtle but crucial point $T(f\chi_{Q_0}) = T(f\chi_{Q_{1,n}})$, which we got from the maximality of the cubes $Q_{1,n}$ in the Calderón-Zygmund decomposition. The fact that by the end we have a sum of terms $T(f\chi_{Q_{1,n}})\chi_{Q_{1,n} \setminus E(Q_{1,n})}$, which are entirely analogous to $T(f\chi_{Q_0})\chi_{Q_0 \setminus G_0}$ is what then allows us to perform the iteration step (which can be written under this new notation in a similar way) successfully.

In the maximal function case, there is no big advantage to either point of view, it is essentially a matter of notation. In the general case, it will be useful to think of the problem in terms of decomposing our operator instead of our space, since it will highlight the requirements we must place on the operator.

2.4 The A_2 conjecture

Now that we have proved that sparse forms have the desired $L^p(\omega)$ bounds and have shown sparse domination is viable in the “proof of concept” case of the maximal operator, we are in a very good place to prove the A_2 conjecture, by adapting the proof to the more general class of Calderón-Zygmund operators.

The structure of the proof will be very similar, but it is important to note that in the case of the maximal function we used some properties that were very specific to this operator. As a result, we will

have to introduce some changes in the proof which exploit the properties Calderón-Zygmund operators do satisfy.

Through this section, we will be stating results using the more general sparse forms

$$\mathcal{A}_{S,p} := \sum_{Q \in S} \langle f \rangle_{Q,p} \chi_Q = \sum_{Q \in S} \left(\frac{1}{|Q|} \int_Q f^p \right) \chi_Q$$

whenever appropriate. This change will not be needed for the proof of the A_2 conjecture, but it does not make the proofs any harder and it lays the groundwork for further generalizations of the method of Sparse Domination, which might appear in chapter 3 if I ever bother to write it.

2.4.1 Reduction to a Cube

A flagrant example of the exploitation of maximal function specific properties is the very first step of the proof, the reduction to a cube, which relied heavily on the particular properties of the maximal operator to conclude that it was enough to add the predecessors of Q_0 to the sparse family. As a result, our first order of business is proving that we can restrict our focus to a cube in general, as long as we include a few more points in our averages:

Lemma 2.4.1: Let T be an operator and f be a compactly supported L^1 function such that for every cube Q there is a constant $C > 0$ and a η -sparse family $\mathcal{S}_Q$ contained in Q such that

$$|Tf \chi_{\alpha Q}| \leq C \sum_{\tilde{Q} \in \mathcal{S}_Q} \langle f \rangle_{\alpha \tilde{Q},p} \chi_{\tilde{Q}},$$

with $\alpha \geq 3$ and $\alpha \tilde{Q}$ the dilation of the cube $\tilde{Q}$ by α , fixing the centre. Then, we have that there is a η/α^n -sparse family $\mathcal{S}$ such that

$$|Tf| \leq C \sum_{\tilde{Q} \in \mathcal{S}} \langle f \rangle_{\tilde{Q},p} \chi_{\tilde{Q}}.$$

Proof: The main idea of the proof is to find a nice (disjoint) family of cubes $Q_{i,j}$ which covers the entire space and such that for every one of them $\alpha Q_{i,j}$ contains the support of f . In that case, we have that each point x is contained in some Q_x , which means that $Tf(x) = Tf \chi_{\alpha Q_x}(x)$ can be bounded by the local bound. The fact that the family is disjoint will then allow us to create a sparse family from the sparse families associated to each cube.

So, let us fix $Q = Q_{0,1}$ a cube which contains the support of f . Let $Q_{1,1}, \dots, Q_{1,3^n-1}$ be obtained from Q as the disjoint translations of Q which fit together inside $3Q$ (see picture below). Define $Q_{2,1}, \dots, Q_{2,3^n-1}$ from $3Q$ similarly, and more generally define $Q_{m,1} \dots Q_{m,3^n-1}$ from $3^{m-1}Q$.

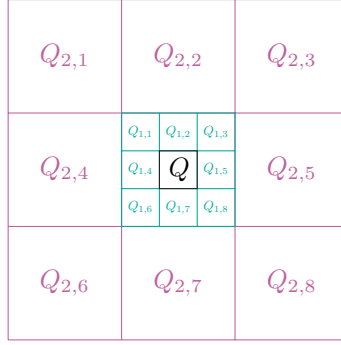


Figure 2.8: In black, the cube Q which contains the support of f . In green, $Q_{1,1}, \dots, Q_{1,8}$ and in pink $Q_{2,1}, \dots, Q_{2,8}$, the first two iterations of the construction of the $Q_{i,j}$.

By construction, it is clear that $Q \subseteq 3Q_{i,j} \subseteq \alpha Q_{i,j}$ (since $\alpha \geq 3$) for every i, j , that all $Q_{i,j}$ are pairwise disjoint and that they cover the space. That means that if we take $\mathcal{S} = \cup_{i,j} \alpha \mathcal{S}_{Q_{i,j}}$, we have that for every $x \in Q$ (with Q_x as defined earlier in the proof)

$$|Tf(x)| = |Tf\chi_{\alpha Q_x}(x)| \leq C \sum_{\tilde{Q} \in \mathcal{S}_{Q_x}} \langle f \rangle_{\alpha \tilde{Q}, p} \chi_{\tilde{Q}} \leq C \sum_{\tilde{Q} \in \mathcal{S}} \langle f \rangle_{\tilde{Q}, p} \chi_{\tilde{Q}},$$

since sparse forms are non-negative. So, we just need to ensure that $\mathcal{S}$ is sparse. This follows by a similar reasoning as the proofs in the first section, by taking $E(\alpha S) := E(S)$. Then, given S_1, S_2 in different families, $E(\alpha S_1)$ and $E(\alpha S_2)$ are disjoint since $E(S_1)$ and $E(S_2)$ are, as a result of the $Q_{i,j}$ being disjoint. Moreover, $|E(\alpha S)| = |E(S)| \geq \eta |S| = \eta \alpha^{-n} |\alpha S|$, making the family $\eta \alpha^{-n}$ -sparse, as intended. $\square$

As a quick note, let us remark that the numerology associated to α might not seem obvious at first but becomes natural once we go through the proof. The α in $\chi_{\alpha Q}$ is crucial to ensure we can encapsulate $\text{supp } f$, since $Q_{i,j}$ could never contain the support by themselves. From there, it is not reasonable to expect to bound it using only sparse forms which are contained in each Q , so we do the next best thing and ask that they all are α -dilations of elements in Q , so we can still keep some semblance of localness and reach all of αQ .

2.4.2 Reduction to the Iteration Step

In terms of the general idea, this proof will follow very similarly, using iteration. The following lemma sums up this idea:

Lemma 2.4.2: Let T be an operator and C_T a constant independent of Q such that for every Q there exists a family $\{Q_j\}_j$ such that

$$|T(f\chi_{3Q})|_{\chi_Q} \leq C_T \langle f \rangle_{3Q, p} \chi_{3Q} + \sum_j |T(f\chi_{3Q_j})|_{\chi_{Q_j}}, \quad (2.5)$$

with Q_j disjoint cubes contained in Q such that $|\cup_j Q_j| \leq \frac{1}{2}|Q|$. Then, for every cube Q_0 there exists a

Sparse family $\mathcal{S} = \mathcal{S}(Q_0)$ such that:

$$|T(f_{3Q_0})| \leq C \sum_{\tilde{Q} \in \mathcal{S}} \langle f \rangle_{\tilde{Q}, p} \chi_{\tilde{Q}}.$$

In particular, by Lemma 2.4.1, T is bounded by a $\frac{1}{2\alpha^n}$ -sparse form.

Note that this is very similar to our iteration step in the maximal function, but now accounting for the dilations introduced by Lemma 2.4.1.

Proof: By Lemma 2.4.1, if we show that for any cube Q there is a $\frac{1}{2}$ -sparse family $\mathcal{S}(Q)$ such that

$$|T(f\chi_{3Q})|\chi_Q \leq C \sum_{\tilde{Q} \in \mathcal{S}_Q} \langle f \rangle_{3\tilde{Q}, p} \chi_{\tilde{Q}}$$

almost everywhere, as then the lemma gives us a global $\frac{1}{2 \cdot 3^n}$ -sparse family $\mathcal{S}$ whose associated sparse form dominates our operator.

This proof is essentially analogous to the maximal function proof, with the slight change of view in remark 2.3.5. So, let us fix an arbitrary cube Q_0 and define $\mathcal{S}(Q_0)$ iteratively as follows:

0. Add Q_0 to $\mathcal{S}(Q_0)$ and use the hypothesis of the lemma to find disjoint cubes $\{(Q_0)_j\}_j$ contained in Q_0 such that the (2.5) holds and $|\cup_j (Q_0)_j| \leq \frac{1}{2}|Q_0|$. Define $E(Q_0) = Q_0 \setminus \cup_j (Q_0)_j$, which then satisfies $|E(Q_0)| \geq \frac{1}{2}|Q_0|$ and will be disjoint from any other $E(Q_{i,j})$ as long as these are chosen to be contained in $\cup_j (Q_0)_j = Q_0 \setminus E(Q_0)$. Define the collection $\{Q_{1,j} := (Q_0)_j\}_j$, $G_0 = E(Q_0)$ and $B_0 = \cup_j (Q_0)_j$.
1. For each $Q_{1,n}$, use the hypothesis of the lemma again, to find disjoint cubes $\{(Q_{1,n})_j\}_j$ which still satisfy (2.5) and $|\cup_j (Q_{1,n})_j| \leq \frac{1}{2}|Q_{1,n}|$, which means that if we define $E(Q_{1,n}) = Q_{1,n} \setminus \cup_j (Q_{1,n})_j$ we once again get the desired $|E(Q_{1,n})| \geq \frac{1}{2}|Q_{1,n}|$, so we add $\mathcal{S}_1 = Q_{1,n}$ to $\mathcal{S}(Q_0)$ and define the collection $\{Q_{2,m}\}_m$ as the collection of $\{(Q_{1,n})_j\}_j$ for all n , $G_1 = \cup_n E(Q_{1,n})$ and $B_1 = \cup_{n,j} (Q_{1,n})_j$.
- i. Given $\{Q_{i,n}\}_n$, proceed as in 1, using the hypothesis of the lemma to define $E(Q_{i,n}) = Q_{i,n} \setminus \cup_j (Q_{i,n})_j$, add $\mathcal{S}_i := \{Q_{i,n}\}_n$ to the sparse family and define $\{Q_{i+1,n}\}_n$ from the $(Q_{i,n})_j$ for all n, j and $G_i = \cup_n E(Q_{i,n})$ and $B_i = \cup_{n,j} (Q_{i,n})_j = \cup_m Q_{i+1,m}$. $\square$

It is easy to check that we have defined a sparse family, since we ensured the size condition on the E -sets at every step, and defined new E -sets as subsets of pairwise disjoint cubes in the complement of previous E -sets. Moreover, note that by construction we have

$$|B_i| = \sum_n \sum_j |(Q_{i,n})_j| \leq \sum_n \frac{1}{2}|Q_{i,n}| = \frac{1}{2}|B_{i-1}|,$$

so $|B_i| \leq \frac{1}{2^i}|Q_0|$, which means that $|B_i| \rightarrow 0$ and so $|\cap_i B_i| \leq |B_i|$ implies that $|\cap_i B_i| = 0$ and therefore almost every $x \in Q_0$ is outside some B_i .

So, let us fix x and assume without loss of generality that $x \in B_n^c$. Using induction on (2.5), we get that for every m

$$|T(f\chi_{3Q_0})|\chi_{Q_0} \leq C_T \sum_{0 \leq i \leq m, \tilde{Q} \in \mathcal{S}_i} \langle f \rangle_{3\tilde{Q}, p} \chi_{\tilde{Q}} + \sum_{\tilde{Q} \in \mathcal{S}_{m+1}} |T(f\chi_{3\tilde{Q}})|\chi_{\tilde{Q}}.$$

Choosing $m = n$, we have that since $x \in B_n^c$, by definition of $B_n = \cup_m Q_{n+1, m}$ and $\mathcal{S}_{n+1} := \{Q_{n+1, m}\}_m$ we have that $\chi_{\tilde{Q}}(x) = 0$ for every $\tilde{Q} \in \mathcal{S}_{n+1}$, so the inequality above becomes

$$|T(f\chi_{3Q_0})|\chi_{Q_0}(x) \leq C_T \sum_{0 \leq i \leq m, \tilde{Q} \in \mathcal{S}_i} \langle f \rangle_{3\tilde{Q}, p} \chi_{\tilde{Q}}(x) \leq C_T \sum_{\tilde{Q} \in \mathcal{S}(Q_0)} \langle f \rangle_{3\tilde{Q}, p} \chi_{\tilde{Q}}(x),$$

since sparse forms are positive. Since x is an arbitrary point in the complement of a zero measure set, we get the inequality almost everywhere, as intended. As a result, lemma 2.4.1 then gives us a $\frac{1}{2 \cdot 3^n}$ -sparse family $\mathcal{S}$ such that

$$|T(f\chi_{3Q_0})| \leq C \sum_{\tilde{Q} \in \mathcal{S}} \langle f \rangle_{\tilde{Q}, p} \chi_{\tilde{Q}}.$$

2.4.3 Proving the Iteration Step

At this point, we have reduced the A_2 conjecture to proving the iteration step (2.5) for Calderón-Zygmund operators. The proof that follows is a slight adaptation of [Ler16], Lerner's original simplified proof of pointwise sparse dominations, using some ideas from his later paper [LO20]. Note that for this part of the proof we will fix $p = 1$ and return to the notation $\langle f \rangle_Q = \langle f \rangle_{Q, 1}$.

Proposition 2.4.3: Let T be a Calderón-Zygmund operator and f be a L^1 function. Then, there is a constant C_T independent of Q such that for every Q we have a collection of Q_j pairwise disjoint cubes contained in Q such that $|\cup_j Q_j| \leq \frac{1}{2}|Q|$ and

$$|T(f\chi_{3Q})|\chi_Q \leq C_T \langle f \rangle_{3Q} \chi_{3Q} + \sum_j |T(f\chi_{3Q_j})|\chi_{Q_j}.$$

Proof: Inspired by the proof for the maximal function, we would be tempted to define the set

$$A = \{x \in Q : |T(f\chi_{3Q})| > C_A \langle f \rangle_{3Q}\}$$

(for some suitable constant C_A) as the bad set. This will still be our starting point, but we will soon see it will require some tweaking, essentially through two new ideas.

The first change:

In order to introduce the first of the changes, let us assume for a second that we have decomposed the bad set (which we will denote E for evil, since B is required further in the proof) into our intended disjoint cubes $\{Q_j\}$. Then, we could get the following bound:

$$|T(f\chi_{3Q})|\chi_Q = |T(f\chi_{3Q})|\chi_G + |T(f\chi_{3Q})|\chi_E = |T(f\chi_{3Q})|\chi_G + \sum_j |T(f\chi_{3Q})|\chi_{Q_j}.$$

The first term would be bounded by $C_A \langle f \rangle_{3Q} \chi_Q \leq C_A \langle f \rangle_{3Q} \chi_Q$, and the second looks close to what we want, except that T should be applied to $f \chi_{3Q_j}$ instead of $f \chi_{3Q}$. If we use the fact that T is linear, we can rewrite the last term as

$$\begin{aligned} \sum_j |T(f \chi_{3Q})| \chi_{Q_j} &= \sum_j (|T(f \chi_{3Q})| - |T(f \chi_{3Q_j})|) \chi_{Q_j} + \sum_j |T(f \chi_{3Q_j})| \chi_{Q_j} \leq \\ &\leq \sum_j |T(f \chi_{3Q \setminus 3Q_j})| \chi_{Q_j} + \sum_j |T(f \chi_{3Q_j})| \chi_{Q_j}, \end{aligned}$$

where we used the triangle inequality to get

$$|T(f \chi_{3Q})| - |T(f \chi_{3Q_j})| \leq |T(f \chi_{3Q}) - T(f \chi_{3Q_j})| = |T(f \chi_{3Q} - f \chi_{3Q_j})| = |T(f \chi_{3Q \setminus 3Q_j})|,$$

since $Q_j \subseteq Q$. So, we want to control the term $|T(f \chi_{3Q \setminus 3Q_j})|$ in points in Q_j , so it makes sense to consider the maximal operator

$$\mathcal{M}_{Q_0, T} f(x) := \sup_{\tilde{Q} \subseteq Q_0, \tilde{Q} \ni x} \operatorname{esssup}_{\xi \in \tilde{Q}} |T(f \chi_{3Q_0 \setminus 3\tilde{Q}})(\xi)|.$$

While taking the supremum over all cubes containing x is the usual in maximal operators, taking the essential supremum over all points in the cube might seem less obvious, but it is very useful, as it makes it so that given $\tilde{Q} \ni x$,

$$\mathcal{M}_{Q_0, T} f(x) \geq \operatorname{esssup}_{\xi \in \tilde{Q}} |T(f \chi_{3Q_0 \setminus 3\tilde{Q}})(\xi)| \stackrel{a.e.}{\geq} |T(f \chi_{3Q_0 \setminus 3\tilde{Q}})(\xi)|,$$

which allows us to turn upper bounds on $\mathcal{M}_{Q_0, T}$ in one point of a cube into upper bounds of $|T(f \chi_{3Q_0 \setminus 3\tilde{Q}})|$ for almost every point ξ in that cube. In particular, as long as Q_j contains points in a “good set” of $\mathcal{M}_{Q, T}$, then $|T(f \chi_{3Q \setminus 3Q_j})|$ will be bounded by our sparse form in almost every point in Q_j , which deals with the compensating term.

So, our first change will be to add to our “bad set” a portion associated to $\mathcal{M}_{Q, T} f$, namely

$$B = \{x \in Q : \mathcal{M}_{Q, T} f > C_B \langle f \rangle_{3Q}\}.$$

By the reasoning above, as long as we are careful in our future choice of Q_j to ensure that each cube catches a positive measure subset of the points in B^c , then the bound on $\mathcal{M}_{Q, T}$ on this subset extends to a bound on $|T(f \chi_{3Q \setminus 3Q_j})| \chi_{Q_j}$ in the whole Q_j , so it takes care of the introduced compensating term.

In order to control the size of B , we will need some kind of weak bound for $\mathcal{M}_{Q, T} f$. In fact, since it has been helpful for these weak bounds to have the operator applied to $f \chi_{3Q_0}$ instead of f , the useful weak bound will be on the operator

$$\mathcal{M}_T f := \sup_{Q \ni x} \operatorname{esssup}_{\xi \in Q} |T(f \chi_{\mathbb{R}^n \setminus 3Q})(\xi)|,$$

since we have $\mathcal{M}_T(f\chi_{3Q_0}) = \sup_{Q \ni x} \text{esssup}_{\xi \in Q} |T(f\chi_{3Q_0 \setminus 3Q})(\xi)|$, which dominates $\mathcal{M}_{Q_0, T}f$ since the sup is taken over more cubes. The weak bound for $\mathcal{M}_T$ is the content of lemma 2.4.4, which follows this proof.

The second change:

The second change has to do with our choice of Q_j . In the proof for the maximal function, it turned out that our bad set was exactly the bad set in the Calderón-Zygmund decomposition of f , which is no longer the case. So, in order to define our Q_j , we will instead be applying Calderón-Zygmund decomposition on $\chi_{A \cup B}$, with a λ carefully chosen such that the union of the Q_j is not too much bigger than $A \cup B$, so we can still use the control on the size of $A \cup B$ to control the size of Q_j , while at the same time containing a significant portion of B^c , as to deal with the compensating term.

The stucture of the proof:

Picture 2.9 below sums up the structure of the proof described so far:

First, define A and B as above, as the bad sets of $T(f\chi_{3Q})$ and $\mathcal{M}_{Q, T}f$, with constants C_A and C_B such that the measure of $A \cup B$ is small. Then, use Calderón-Zygmund decomposition on $\chi_{A \cup B}$ to get the Q_j , represented in the picture by Q_1, Q_2, Q_3 and Q_4 .

Then, define the good set as represented in green, the complement in Q of the union of the Q_j . Here, $T(f\chi_{3Q})$ is bounded since we are in A^c . In order to introduce the terms we intend to iterate on, we need to deal with the compensating term $\sum_j |T(f\chi_{3Q \setminus 3Q_j})|\chi_{Q_j}$, which we can do by the reasoning above since every Q_j contains a positive measure piece of B^c , represented for Q_4 in dashed blue.

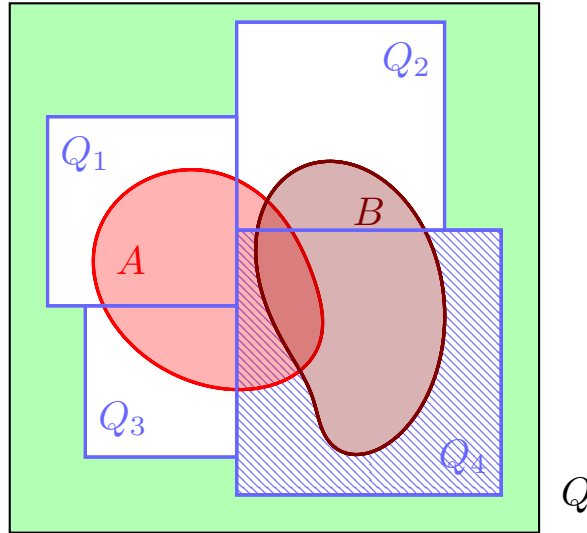


Figure 2.9: Visual representation of the proof structure of Proposition 2.4.3. Note that technically the Q_j should be elements of $\mathcal{D}(Q)$ and their area should be smaller, this picture is for illustrative effects only.

All that is left to do is work out the constants C_A, C_B, λ . In order to apply Calderón-Zygmund decomposition to $\chi_{A \cup B}$ as a subset of Q with height λ , we need to ensure that $\lambda \geq \frac{1}{|Q|} \int_Q \chi_{A \cup B} = \frac{|A \cup B|}{|Q|}$, as otherwise we would just get the trivial decomposition where Q is the bad set. Then, we get a collection of disjoint cubes $\{Q_j\}_j \subseteq \mathcal{D}(Q)$ such that:

- In $Q \setminus (\cup_j Q_j)$, $\chi_{A \cup B} \leq \lambda$. If $\lambda < 1$, this ensures that $A \cup B \subseteq \cup_j Q_j$, since in those points $\chi_{A \cup B} = 1$.

- For each Q_j , we have $\lambda < \frac{1}{|Q_j|} \int_{Q_j} \chi_{A \cup B} = \frac{|(A \cup B) \cap Q_j|}{|Q_j|} \leq 2^n \lambda$. Recall that we wanted Q_j to be somewhat small, which can be ensured using the lower inequality

$$\sum_j |Q_j| \leq \frac{1}{\lambda} \sum_j |(A \cup B) \cap Q_j| \leq \frac{|A \cup B|}{\lambda},$$

and choosing $|A \cup B|$ (via C_A, C_B) and λ adequately. On the other hand, we wanted Q_j to include some of B^c , which can be ensured using the higher inequality:

$$|(A \cup B) \cap Q_j| \leq |Q_j| 2^n \lambda,$$

so it is enough to choose $\lambda < 2^{-n}$ in order for $|Q_j|$ to contain a positive measure set outside $A \cup B$ and as a consequence outside B .

Working out the constants:

To sum up, we need to choose $\lambda, |A \cup B|$ such that $\lambda \geq \frac{|A \cup B|}{|Q|}$, $\lambda < 2^{-n}$ and $\frac{|A \cup B|}{\lambda} \leq \frac{1}{2}|Q| \iff \frac{|A \cup B|}{|Q|} \leq \frac{\lambda}{2}$. It is clear that choosing $\lambda = 2^{-n-1}$ and ensuring that $|A \cup B| \leq 2^{-n-2}|Q|$ satisfies all three.

In order to ensure that $|A \cup B| \leq 2^{-n-2}|Q|$, we will choose C_A and C_B such that $|A|, |B| \leq 2^{-n-3}|Q|$, so that $|A \cup B| \leq |A| + |B| \leq 2^{-n-2}|Q|$:

- Using the weak (1, 1) inequality for T , we get

$$|A| = |\{x \in Q : |T(f\chi_{3Q})| > C_A \langle f \rangle_{3Q}\}| \leq \|T\| \frac{\int_{\mathbb{R}^n} f\chi_{3Q}}{C_A \langle f \rangle_{3Q}} = \|T\| \frac{|3Q|}{C_A} = \frac{\|T\| 3^n}{C_A} |Q|, \quad (2.6)$$

so we choose $C_A = 2^{n+3}\|T\|3^n$, which only depends on T and n .

- Using the weak (1, 1) inequality for $\mathcal{M}_T$ from lemma 2.4.4, whose constant only depends on T , we get

$$\begin{aligned} |B| &= |\{x \in Q : \mathcal{M}_{Q,T} f > C_B \langle f \rangle_{3Q}\}| \leq \\ &\leq |\{x \in Q : \mathcal{M}_T(f\chi_{3Q}) > C_B \langle f \rangle_{3Q}\}| \leq \|\mathcal{M}_T\| \frac{\int_{\mathbb{R}^n} f\chi_{3Q}}{C_B \langle f \rangle_{3Q}} = \|\mathcal{M}_T\| \frac{|3Q|}{C_B} = \frac{\|\mathcal{M}_T\| 3^n}{C_B} |Q|, \end{aligned}$$

so we choose $C_B = 2^{n+3}\|\mathcal{M}_T\|3^n$, which again only depends on T, n .

Proof of the bound:

So, let us define $C_T = C_A + C_B$, with C_A, C_B as defined and let A, B be defined as above. Perform Calderón-Zygmund decomposition at height 2^{-n-2} on $\chi_{A \cup B}$, to get a family of disjoint cubes $\{Q_j\}_j$ with the properties described above. In particular, we have that if we define $E = \cup_j Q_j$ and $G = Q \setminus E$, then $G \subseteq Q \setminus A$, which means that we have the bound

$$|T(f\chi_{3Q})|_{\chi_Q} = |T(f\chi_{3Q})|_{\chi_G} + |T(f\chi_{3Q})|_{\chi_E} \leq C_A \langle f \rangle_{3Q} \chi_G + |T(f\chi_{3Q})|_{\chi_E} \leq C_A \langle f \rangle_{3Q} \chi_Q + |T(f\chi_{3Q})|_{\chi_E},$$

so it is enough to deal with the term $|T(f\chi_{3Q})|_{\chi_E} = \sum_j |T(f\chi_{3Q})|_{\chi_{Q_j}}$. Indeed, we have by the compu-

tations above that

$$\sum_j |T(f\chi_{3Q})|_{\chi_{Q_j}} \leq \sum_j |T(f\chi_{3Q \setminus 3Q_j})|_{\chi_{Q_j}} + \sum_j |T(f\chi_{3Q_j})|_{\chi_{Q_j}}.$$

Given that $|Q_j \cup B^c| \neq 0$ by construction, we have that for almost every $x \in Q^5$

$$|T(f\chi_{3Q \setminus 3Q_j})|(x) \leq \sup_{\tilde{Q} \subseteq Q, \tilde{Q} \ni x} \operatorname{esssup}_{\xi \in \tilde{Q}} |T(f\chi_{3Q \setminus 3\tilde{Q}})(\xi)| = \mathcal{M}_{Q_0, T} f(x)$$

and so since in B^c we have the bound $\mathcal{M}_{Q_0, T} f(x) \leq C_B \langle f \rangle_{3Q}$, we get

$$\begin{aligned} \sum_j |T(f\chi_{3Q \setminus 3Q_j})|_{\chi_{Q_j}} + \sum_j |T(f\chi_{3Q_j})|_{\chi_{Q_j}} &\leq C_B \langle f \rangle_{3Q} \sum_j \chi_{Q_j} + \sum_j |T(f\chi_{3Q_j})|_{\chi_{Q_j}} \leq \\ &\leq C_B \langle f \rangle_{3Q} \chi_Q + \sum_j |T(f\chi_{3Q_j})|_{\chi_{Q_j}}, \end{aligned}$$

since the Q_j are disjoint and contained in Q . As a result, combining both bounds gives

$$\begin{aligned} |T(f\chi_{3Q})|_{\chi_Q} &\leq C_A \langle f \rangle_{3Q} + |T(f\chi_{3Q})|_{\chi_E} \leq C_A \langle f \rangle_{3Q} + C_B \langle f \rangle_{3Q} \chi_Q + \sum_j |T(f\chi_{3Q_j})|_{\chi_{Q_j}} = \\ &= C_T \langle f \rangle_{3Q} \chi_Q + \sum_j |T(f\chi_{3Q_j})|_{\chi_{Q_j}}, \end{aligned}$$

which is the intended bound. □

We now prove the technical lemma, the weak bound on $\mathcal{M}_T$ for Calderón-Zygmund operators:

Lemma 2.4.4: Let T be a Calderón-Zygmund operator and let the operator $\mathcal{M}_T$ be defined as

$$\mathcal{M}_T f(x) = \sup_{Q \ni x} \operatorname{esssup}_{\xi \in Q} |T(f\chi_{\mathbb{R}^n \setminus 3Q})(\xi)|.$$

Then, $\mathcal{M}_T$ is a bounded operator from L^1 to $L^{1, \infty}$.

Proof: In order to show the weak bound, we will bound $\mathcal{M}_T$ almost everywhere by a sum of operators we know satisfy a weak bound. In order to bound $\mathcal{M}_T f(x)$, let us fix an arbitrary cube $Q \ni x$ and an arbitrary $\xi \in Q$, and bound $|T(f\chi_{\mathbb{R}^n \setminus 3Q})(\xi)|$.

First, recall that $T^* f(x) := \sup_{\epsilon} \int_{B_{\epsilon}(x)^c} K(x, y) f(y) dy$ is by classical theory a weak $(1, 1)$ operator, which bounds any $|T(f\chi_{\mathbb{R}^n \setminus B_x})(x)| = \int_{B_x^c} K(x, y) f(y) dy$, since $x \notin \operatorname{supp} f\chi_{\mathbb{R}^n \setminus B_x}$, for any ball B_x around x . This is not exactly the term we have right now, so we need to introduce new terms, since we want to

⁵Note that the essential step in this inequality is that $|T(f\chi_{3Q})|_{\chi_{Q_j}}(x) \leq \operatorname{esssup}_{\xi \in Q} |T(f\chi_{3Q})(\xi)|$ holds for almost every $x \in Q$. This follows from the definition of esssup as the infimum over null measure sets of the supremum of the function in their complement, by defining E_n as a sequence of null-measure sets such that the inequality $\sup_{\xi \in Q \setminus E_n} |T(f\chi_{3Q})(\xi)| - \frac{1}{n} \leq \operatorname{esssup}_{\xi \in Q} |T(f\chi_{3Q})(\xi)|$ holds (this is just the usual real analysis taking a sequence towards the infimum argument). It is clear that any point for which our inequality does not hold must be in some E_n , and so the fact that $E := \cup_n E_n$ has null measure ensures the inequality holds for all but a null-measure set E , as intended.

change both ξ to x and $3Q$ to B_x :

$$|T(f\chi_{\mathbb{R}^n \setminus 3Q})(\xi)| \leq |T(f\chi_{\mathbb{R}^n \setminus 3Q})(\xi) - T(f\chi_{\mathbb{R}^n \setminus B_x})(\xi)| + |T(f\chi_{\mathbb{R}^n \setminus B_x})(\xi) - T(f\chi_{\mathbb{R}^n \setminus B_x})(x)| + |T(f\chi_{\mathbb{R}^n \setminus B_x})(x)|.$$

By linearity of T , the first term is, assuming that $3Q \subseteq B_x$, $|T(f\chi_{\mathbb{R}^n \setminus 3Q})(\xi) - T(f\chi_{\mathbb{R}^n \setminus B_x})(\xi)| = |T(f\chi_{\mathbb{R}^n \setminus 3Q} - f\chi_{\mathbb{R}^n \setminus B_x})(\xi)| = |T(f\chi_{B_x \setminus 3Q})(\xi)|$.⁶ Since ξ is not in the support, we can use the standard kernel form of T :

$$|T(f\chi_{B_x \setminus 3Q})(\xi)| \leq \left| \int_{B_x \setminus 3Q} K(\xi, y) f(y) dy \right| \leq \int_{B_x \setminus 3Q} \frac{1}{|\xi - y|^n} |f(y)| dy.$$

Since $\xi \in Q$ and $y \in (3Q)^c$, we have that $|y - \xi| \geq \ell(Q) \implies |y - \xi|^n \geq |Q|$, which means that as long as we define B_x with radius depending on $\ell(Q)$ then we will be able to bound the $|y - \xi|^{-n}$ term by $c_n |B_x|^{-1}$ with c_n some dimensional constant, which will introduce an average of $|f|$ over a ball containing x which we can then bound by the (weak operator) maximal function of f at x . Indeed, it is easy to check that if B_x has radius $2\text{diam}(Q)$, then it contains $3Q$ (see picture 2.10), and since $\text{diam}(Q) = \sqrt{n}\ell(Q)$ we do have $|Q|^{-1} \geq c_n |B_x|^{-1}$, which concludes the bound for this term:

$$|T(f\chi_{B_x \setminus 3Q})(\xi)| \leq \int_{B_x \setminus 3Q} \frac{1}{|\xi - y|^n} |f(y)| dy \leq \frac{c_n}{|B_x|} \int_{B_x \setminus 3Q} |f(y)| dy \leq \frac{c_n}{|B_x|} \int_{B_x} |f(y)| dy \leq c_n Mf(x),$$

with Mf the maximal function over balls, which is known to be weak $(1, 1)$.

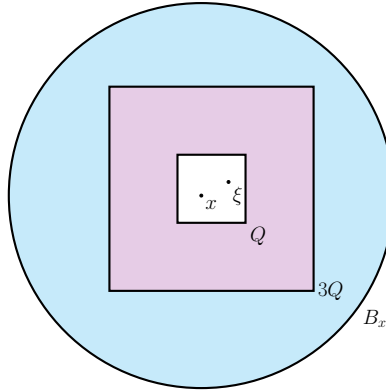


Figure 2.10: In the center, the cube Q which contains x and ξ , contained in the cube $3Q$. In the outside, the cube B_x with centre x and radius $2\text{diam}(Q)$, which contains $3Q$ and therefore Q , x and ξ as well.

The third term was constructed to satisfy $|T(f\chi_{\mathbb{R}^n \setminus B_x})(x)| \leq T^*f(x)$, so all that is left to do is to bound the second term. The key observation is that again x, ξ are not in the support of the function, so we can use the standard kernel formula for T , as well as classic Calderón-Zygmund arguments (which we have seen before, in the proof of Proposition 1.6.3):

$$|T(f\chi_{\mathbb{R}^n \setminus B_x})(\xi) - T(f\chi_{\mathbb{R}^n \setminus B_x})(x)| = \left| \int_{\mathbb{R}^n \setminus B_x} (K(\xi, y) - K(x, y)) f(y) dy \right| \leq \int_{\mathbb{R}^n \setminus B_x} |K(\xi, y) - K(x, y)| |f(y)| dy.$$

⁶It is natural to choose B_x either fully contained in $3Q$ or fully containing $3Q$ so that the difference simplifies to either $T(f\chi_{B_x \setminus 3Q})(\xi)$ or $T(f\chi_{3Q \setminus B_x})(\xi)$; the remaining components will make it clear that a large B_x is the convenient choice.

Now note that for $y \in B_x^c$, we have $|y - x| \geq r(B_x) = 2\text{diam}(Q) \geq 2|x - \xi|$, which means we can apply the Calderón-Zygmund bound $|K(\xi, y) - K(x, y)| \leq C \frac{|x - \xi|^\delta}{|x - y|^{n+\delta}}$, and get, setting $r := r(B_x)$:

$$\int_{\mathbb{R}^n \setminus B_x} |K(\xi, y) - K(x, y)| |f(y)| dy \leq \int_{\mathbb{R}^n \setminus B_x} C \frac{|x - \xi|^\delta}{|x - y|^{n+\delta}} |f(y)| dy \leq C \sum_{k=1}^{\infty} \int_{2^k r \leq |x - y| < 2^{k+1} r} \frac{|x - \xi|^\delta}{|x - y|^{n+\delta}} |f(y)| dy,$$

where we used the fact that the integrand is positive and $y \in B_x^c \implies |x - y| \geq r$. Now, we can simplify the k th term, as $|x - \xi|^\delta \leq r^\delta$ and $|x - y|^{n+\delta} \geq 2^{kn+k\delta} r^{n+\delta}$, so we get:

$$\begin{aligned} C \sum_{k=1}^{\infty} \int_{2^k r \leq |x - y| < 2^{k+1} r} \frac{|x - \xi|^\delta}{|x - y|^{n+\delta}} |f(y)| dy &\leq C \sum_{k=1}^{\infty} 2^{-nk-k\delta} \int_{2^k r \leq |x - y| < 2^{k+1} r} \frac{1}{r^n} |f(y)| dy \leq \\ &\leq C \sum_{k=1}^{\infty} 2^{-nk-k\delta} 2^{nk+n} \frac{1}{(2^{k+1} r)^n} \int_{|x - y| < 2^{k+1} r} |f(y)| dy \leq C c_n 2^n \sum_{k=1}^{\infty} 2^{-k\delta} Mf(x) \lesssim Mf(x), \end{aligned}$$

since $\sum_{k=1}^{\infty} 2^{-k\delta}$ converges and c_n comes from turning the area of the cube into the area of the ball.

Looking back at all the terms, we get by taking the supremum over Q, ξ that

$$\mathcal{M}_T f(x) \lesssim T^* f(x) + Mf(x),$$

since the constants are independent of Q, ξ . We now get that $\mathcal{M}_T f(x)$ is weak $(1, 1)$ by usual arguments:

$$\begin{aligned} |\{\mathcal{M}_T f(x) > \lambda\}| &\leq |\{T^* f(x) + Mf(x) > C^{-1}\lambda\}| \leq \left| \left\{ T^* f(x) > \frac{C^{-1}\lambda}{2} \right\} \right| + \left| \left\{ Mf(x) > \frac{C^{-1}\lambda}{2} \right\} \right| \leq \\ &\leq C_{T^*} \frac{\|f\|_1}{(2C)^{-1}\lambda} + C_M \frac{\|f\|_1}{(2C)^{-1}\lambda} \lesssim \frac{\|f\|_1}{\lambda}, \end{aligned}$$

as intended. □

2.4.4 Proof of the A_2 Conjecture

In this subsection, we put the pieces of the proof together for bookkeeping. We will in fact state and prove the more general version for $1 < p < \infty$, since it follows from the work above.

Theorem 2.4.5 (A_p Conjecture): Let T be a Calderón-Zygmund operator and $\omega \in A_p$. Then there is a constant independent of ω such that for every $f \in L^p(\omega)$ we have

$$\|Tf\|_{L^p(\omega)} \lesssim [\omega]_p^{\max(1, \frac{1}{p-1})} \|f\|_{L^p(\omega)}.$$

Proof: First, since the $L^p(\omega)$ boundedness of Calderón-Zygmund operators is a classical result, we can assume without loss of generality that f is L^1 and compactly supported, since then the $[\omega]_p$ dependence extends to any function in $L^p(\omega)$ by density. So, let us fix such an f , so it satisfies the hypothesis of the previous lemmas and propositions.

By proposition 2.4.3, we have that there is a constant C_T independent of Q such that for every Q we

have a collection of Q_j pairwise disjoint cubes contained in Q such that $|\cup_j Q_j| \leq \frac{1}{2}|Q|$ and

$$|T(f\chi_{3Q})|\chi_Q \leq C_T \langle f \rangle_{3Q} \chi_{3Q} + \sum_j |T(f\chi_{3Q_j})|\chi_{Q_j}.$$

By lemma 2.4.2, we have that this implies for every Q the existence of a $\frac{1}{2}$ -sparse family $\mathcal{S}(Q)$ such that

$$|T(f\chi_{3Q})|\chi_Q \leq C_T \sum_{\tilde{Q} \in \mathcal{S}} \langle f \rangle_{\tilde{Q}} \chi_{\tilde{Q}},$$

which by lemma 2.4.1 implies the existence of a $\frac{1}{2 \cdot 3^n}$ -sparse family $\mathcal{S}$ such that

$$|Tf| \leq C \sum_{\tilde{Q} \in \mathcal{S}} \langle f \rangle_{\tilde{Q}} \chi_{\tilde{Q}} = \mathcal{A}_{\mathcal{S}} f.$$

By proposition 2.2.4, we know that $\mathcal{A}_{\mathcal{S}} f$ has the bound above, which trivially extends to T :

$$\|Tf\|_{L^p(\omega)} \leq \|\mathcal{A}_{\mathcal{S}} f\|_{L^p(\omega)} \lesssim [\omega]_p^{\max(1, \frac{1}{p-1})} \|f\|_{L^p(\omega)},$$

proving the A_2 conjecture and more generally the bound for general $1 < p < \infty$. □

Chapter 3

The Sparse Revolution

In chapter 2, we proved the A_2 conjecture as an adaptation of Lerner's proof published in 2015 in [Ler16], as the culmination of many sequentially simpler proofs of the same result. This proof was done via a particularly strong sparse bound, namely an almost everywhere pointwise bound of the operator by a sparse form, and use it to extract information about the original operator's $1 < p < \infty$ strong weighted bounds.

It turns out, the key ideas in the Sparse Domination technique can be extended, on the one hand by considering weaker bounds by sparse forms - we will focus on bilinear Sparse Bounds as an example of such weakening - and on the other hand by exploring how to extract weighted weak and endpoint bounds for operators which are in some way dominated by sparse forms.

As a result, in following years, and up to the present, the "Sparse Revolution"¹ ensued, where this simple and somewhat elementary idea of using sparse forms to prove weighted bounds was developed in several other reincarnations of the same principle, such as bilinear Sparse Bounds, for many other diverse classes of operators relevant in Harmonic Analysis.

In this last chapter we will mention some of these different reincarnations, including a negative result for the multiparameter case, and reference some of the literature on the topic, though obviously not exhaustively. Further work not addressed in this thesis includes most generalizations to multilinear and vector-valued settings [CDO18; LN22; CS24; Li18], the Sparse T_1 Theorem [LM17], the Helicoidal Method [BM18; BM21] and Multilinear Matrix Weights [KN24].

The series of lectures [BC20b] by D. Beltran and J. M. Conde Alonso was a very valuable resource for a good part of this chapter, especially for subsection 3.2, as well as Rodrigo Duarte's informal lecture notes.

3.1 Generalized Pointwise Domination

In chapter 2, most of our work was done in greater generality than needed for the Calderòn-Zygmund case, by considering sparse forms with L^p averages instead. This was in view to the generalized The-

¹This expression was coined by M.C Pereyra in the survey [Per19], which includes references to other applications of sparse domination, some of which we do not address here, particularly in section 8.

orem in Pointwise Sparse Domination proved by Lerner and Ombrosi in their 2020 paper [LO20]. This generalized version is two-fold: firstly, the weak bound on T gets replaced by the weaker property of satisfying the so-called W_q condition, and secondly the weak bound for $\mathcal{M}_T$ gets replaced by a weak bound for the operator $\mathcal{M}_{T,\alpha}^\#$ (to be defined shortly), which plays a very similar role in the proof, but is more flexible and easier to prove weak bound results for.

Definition 3.1.1 (W_q Condition): Given an operator T , we say it satisfies the W_q property if there exists a non-increasing function $\psi_T, q(\lambda)$ for $0 < \lambda < 1$ such that the inequality

$$|\{x \in Q : |T(f\chi_Q)(x)| > \psi_{T,q}(\lambda)\langle f \rangle_{q,Q}\}| \leq \lambda|Q|$$

holds for any cube Q , $f \in L^q(Q)$ and $0 < \lambda < 1$.

Remark 3.1.2: This is indeed a generalization of the notion of weak (q, q) operator: given such an operator T we have (with $\|T\|$ denoting the weak (q, q) norm of T)

$$|\{x \in Q : |T(f\chi_Q)(x)| > \alpha\langle f \rangle_{q,Q}\}| \leq \|T\|^q \frac{\int_Q |f|^q}{\alpha^q \langle f \rangle_{q,Q}^q} = \frac{\|T\|^q}{\alpha^q} |Q|,$$

so taking $\alpha = \frac{\|T\|}{\lambda^{1/q}}$ gives $\lambda|Q|$ in the right-hand side and shows that T satisfies the W_q condition with the non-increasing function $\psi_{T,q}(\lambda) := \frac{\|T\|}{\lambda^{1/q}}$.

Definition 3.1.3 ($\mathcal{M}_{T,\alpha}^\#$ Operator): Given an operator T and some $\alpha \geq 0$, we define the operator $\mathcal{M}_{T,\alpha}^\#$ as

$$\mathcal{M}_{T,\alpha}^\# f(x) = \sup_{Q \ni x} \operatorname{esssup}_{x', x'' \in Q} |T(f\chi_{\mathbb{R}^n \setminus \alpha Q})(x') - T(f\chi_{\mathbb{R}^n \setminus \alpha Q})(x'')|.$$

Remark 3.1.4: It is clear that a weak bound for $\mathcal{M}_T$ implies the same weak bound for $\mathcal{M}_{T,3}^\#$, since $\operatorname{esssup}_{x', x'' \in Q} |T(f\chi_{\mathbb{R}^n \setminus \alpha Q})(x') - T(f\chi_{\mathbb{R}^n \setminus \alpha Q})(x'')| \leq 2 \operatorname{esssup}_{\xi \in Q} |T(f\chi_{\mathbb{R}^n \setminus \alpha Q})(\xi)|$. However, the advantage of using $\mathcal{M}_{T,\alpha}^\#$ is to prove the weak bound in an easier way, without passing through $\mathcal{M}_T$. In the case of Calderón-Zygmund operators, the simpler bound shown in [LO20] uses $\alpha = 5\sqrt{n}$.

We now present this more general theorem, and sketch the necessary changes in the proof, which follows unchanged, except for proposition 2.4.3.

Theorem 3.1.5: Let T be a sublinear operator satisfying the W_q condition, such that the associated operator $\mathcal{M}_{T,\alpha}^\#$ satisfies a weak (r, r) bound for some $\alpha \geq 3$, with $1 \leq q, r < \infty$. Then, there exists a $\frac{1}{2\alpha^n}$ -sparse family $\mathcal{S}(f)$ such that (taking $s = \max(q, r)$)

$$|Tf(x)| \lesssim \sum_{Q \in \mathcal{S}} \langle f \rangle_{s,Q} \chi_Q(x),$$

holds almost everywhere, for every compactly supported $f \in L^s$.

Sketch of Proof: Up to the iteration step, the proof in chapter 2 holds. In proving the iteration step, assuming we have a decomposition of Q into a good set G and a union of cubes Q_j , we reach the

following decomposition by the method in the proof of proposition 2.4.3:

$$|T(f\chi_{\alpha Q})\chi_Q(x)| \leq |T(f\chi_{\alpha Q})\chi_G(x)| + \sum_j |T(f\chi_{\alpha Q \setminus \alpha Q_j})\chi_{Q_j}(x)| + \sum_j |T(f\chi_{\alpha Q_j})\chi_{Q_j}(x)|.$$

In the right hand side, the last term is the one we want to proceed with the iteration, the first term is bounded by sparse forms, and the middle term was previously handled using the operator $\mathcal{M}_T$. Since we no longer have that tool in our disposal, we can use the inequality

$$|T(f\chi_{\alpha Q \setminus \alpha Q_j})\chi_{Q_j}(x)| \leq |T(f\chi_{\alpha Q \setminus \alpha Q_j})\chi_{Q_j}(x) - T(f\chi_{\alpha Q \setminus \alpha Q_j})\chi_{Q_j}(x')| + |T(f\chi_{\alpha Q \setminus \alpha Q_j})\chi_{Q_j}(x')|,$$

to bound that term by

$$\sum_j |T(f\chi_{\alpha Q \setminus \alpha Q_j})| \leq \sum_j |T(f\chi_{\alpha Q \setminus \alpha Q_j})\chi_{Q_j}(x) - T(f\chi_{\alpha Q \setminus \alpha Q_j})\chi_{Q_j}(x')| + \sum_j |T(f\chi_{\alpha Q \setminus \alpha Q_j})\chi_{Q_j}(x')|,$$

where we have the freedom to choose a convenient x' so that the last term is controlled, and the first term can now be handled using $\mathcal{M}_{T,\alpha}^\#$ in a very similar way to how we used $\mathcal{M}_T$ in chapter 2. There is an added subtlety in this case, since $|T(f\chi_{\alpha Q \setminus \alpha Q_j})\chi_{Q_j}(x')|$ is bounded in the good points by $\langle f \rangle_{s,Q_j}$ instead of $\langle f \rangle_{s,Q}$, but this can be fixed by bounding $\langle f \rangle_{s,Q_j}$ by M_s and adding the points where $M_s \gtrsim \langle f \rangle_{s,Q}$ (with an appropriate constant) to the bad set², and then dealing with them the usual way, by leaving some good points in each cube so we can choose a good x' .

Note also that we are using a weak (r, r) or W_q condition, while working with a sparse form $\langle f \rangle_{s,Q}$ with $s \geq r, q$. This makes no meaningful difference in the proof, as it just introduces another instance of the Hölder inequality $(\frac{s}{r}, (\frac{s}{r})')$ when estimating the size of the bad sets. For example, for a weak (r, r) operator S :

$$|\{x : S(f\chi_Q) > c\langle f \rangle_{s,Q}\}| \lesssim \frac{\int_Q |f|^r}{|Q|^{-\frac{r}{s}}(\int_Q |f|^s)^{\frac{r}{s}}} \leq \frac{|Q|^{1-\frac{r}{s}} \left(\int_Q |f|^{r\frac{s}{r}}\right)^{\frac{r}{s}}}{|Q|^{-\frac{r}{s}}(\int_Q |f|^s)^{\frac{r}{s}}} = |Q|.$$

Regarding the W_q condition, it has very little impact in the proof, as we can use this inequality instead of the usual weak inequality in step (2.6), with the constant $\psi_{T,q}(\delta)$ where δ is such that we want this set to have size at most $\delta|Q|$.

The full proof can be found in [LO20] in the proof of Theorem 1.1. □

Remark 3.1.7: This result can be further generalized by assuming the W_r condition instead of the weak bound (r, r) for $\mathcal{M}_{T,\alpha}^\#$ as well, with the change in the proof being analogous to requiring the W_q condition for T . An interesting consequence is that this allows us to conclude sparse bounds for Tf as long as T and $\mathcal{M}_{T,\alpha}^\#$ satisfy the W_q and W_r conditions for f itself.

It is also possible to replace sparse forms by Orlicz averages, which is the content of Theorem 3.3 of [LO20], and this can be extended to the multilinear case, as seen in Theorem 3.4 of the same paper.

This generalized version provides a unified approach to proving sparse bounds for several op-

²Note that M_s satisfies a weak (s, s) bound; see discussion about M_s right before proposition 3.1.8

erators, many examples of which can be found in section 5 of [LO20]. In particular, the estimate $\mathcal{M}_{T,\alpha}^\# f(x) \lesssim M(|f(x)|^r)^{1/r}$ holds for many of the examples, which is a significant advantage of this generalized theorem, avoiding a proof of weak bounds for the operator $\mathcal{M}_T$.

Regarding the consequences of a sparse bound with averages $\langle f \rangle_{p,Q}$, it is possible to adapt the proofs in subsection 2.2 for this situation, by using the bound $\langle f \rangle_{p,Q} \leq M_p(f)$, for $M_p(f) = \sup_Q \langle f \rangle_{p,Q} \chi_Q$ in place of $\langle f \rangle_Q \leq M(f)$. Now, we have

$$M_p f = \sup_Q \left(\frac{1}{|Q|} \int_Q f^p \right)^{1/p} = \left(\sup_Q \frac{1}{|Q|} \int_Q f^p \right)^{1/p} = M(f^p)^{\frac{1}{p}},$$

which means that as long as $\frac{q}{p} > 1$ we can use the $L^{\frac{q}{p}}$ bound of the maximal operator to get:

$$\int (M_p f)^q = \int M(f^p)^{\frac{q}{p}} \leq \int (f^p)^{\frac{q}{p}} = \int f^q \implies \|M_p f\|_{L^q} \leq \|f\|_{L^q},$$

and we can similarly see that it also inherits a weak (p, p) bound from the weak $(1, 1)$ bound for M . So, we can adapt the proof of proposition 2.2.4 to get the following proposition:

Proposition 3.1.8: Let T be a sublinear operator such that for every $f \in L^q(\omega)$ there is a sparse family of cubes $\mathcal{S}$ such that

$$|Tf| \lesssim \sum_{Q \in \mathcal{S}} \langle f \rangle_{p,Q} \chi_Q.$$

Then, for every $q > p$ and $\omega \in A_{\frac{q}{p}}$ we have

$$\|Tf\|_{L^q(\omega)} \lesssim [\omega]_{\frac{q}{p}}^{\max\{1, \frac{1}{q-1}\}} \|f\|_{L^q(\omega)}.$$

A full proof of this proposition can be found in [Ler16], in the appendix. Note that the fact that we have $[\omega]_{\frac{q}{p}}$ instead of $[\omega]_q$ might appear surprising at first, but it is intuitive enough in light of the discussion above: L^q bounds on M_p are extracted from $L^{\frac{q}{p}}$ bounds on M , so it is natural that a $\frac{q}{p}$ appears in this context.

3.2 Bilinear Sparse Domination

There are two observations which motivate thinking of bilinear sparse domination. The first is that if an operator is pointwise dominated by a sparse form, by proposition 2.2.4 we automatically get strong weighted bounds for every $1 < p < \infty$, and even if we bound it by a sparse form with $p > 1$ we still get by proposition 3.1.8 the full range $p < q < \infty$. That means that if we have any hope of using Sparse Domination techniques for operators with smaller intervals of boundedness, like $r < p < q$, we need to find a weaker notion of Sparse Domination.

The second observation arises when following the proof of the strong bound in proposition 2.2.4: the first step is to bound the L^p norm by the supremum of the bilinear form $\langle \mathcal{A}_S f, g \rangle$ with g of norm 1 in the dual space. This means that if we are to prove an L^p bound for Tf , we just need a bound on the bilinear

form $\langle Tf, g \rangle \lesssim \sum_{Q \in \mathcal{S}} |Q| \langle f \rangle_{1,Q} \langle g \rangle_{1,Q}$.

In the end of the previous subsection, we saw that if we instead allow a bound of the form $\langle Tf, g \rangle \lesssim \sum_{Q \in \mathcal{S}} |Q| \langle f \rangle_{p,Q} \langle g \rangle_{1,Q}$ we decrease the range from $1 < q < \infty$ to $p < q < \infty$. It is therefore plausible that in order to get a smaller range, it is worth investigating what happens when $\langle g \rangle_{1,Q}$ gets replaced by $\langle g \rangle_{r,Q}$ by some r . Indeed, if we focus on the unweighted case, we see that:

$$\sum_{Q \in \mathcal{S}} |Q| \langle f \rangle_{p,Q} \langle g \rangle_{r,Q} \chi_Q \leq \frac{1}{\eta} \sum_{Q \in \mathcal{S}} |E_Q| \langle f \rangle_{p,Q} \langle g \rangle_{r,Q} \leq \int M_p f \cdot M_r g.$$

In order to apply Hölder in this step and be able to exploit the boundedness of M_p and M_r , we need to choose q such that $q > p$, so M_p is bounded in L^q and $q' > r \iff q < r'$, so M_r is bounded in $L^{q'}$. So, we obtain the intended bound for any q in the range $p < q < r'$ if $p \leq r'$ or $r < q < p'$ otherwise (this is obtained swapping q and q'). This brings us to the following definition:

Definition 3.2.1 ((p, r) bilinear sparse bound): Given an operator T we say it satisfies a (p, r) bilinear sparse bound if for every f there is a sparse family of cubes $\mathcal{S}$ such that

$$\langle Tf, g \rangle \lesssim \sum_{Q \in \mathcal{S}} |Q| \langle f \rangle_{p,Q} \langle g \rangle_{r,Q}.$$

Remark 3.2.2: The pointwise case consists implies a $(1, 1)$ bound, which is why we get the full range $1 < p < \infty$. This is the best case scenario. On the other hand, if we had $q = r'$ then we'd have $q < p < q$ and we would therefore get nothing out of the proposition. In a sense, the best (r, s) bilinear bound we can prove is a measure of the size of an operator, in a similar way L^p bounds are.

It has since been proven that we are able to extract weighted bounds from (p, r) bilinear bounds. For example, in [BFP16] several operators are bounded by sparse forms with averages over 5-dilated cubes, and the following proposition is proven in section 6:

Proposition 3.2.3: Let $p < r$ and q be such that $p < q < r$. Let T be a bounded operator in L^p such that for every $f \in L^p$ and $g \in L^{p'}$ there is a sparse collection of cubes $\mathcal{S}$ such that

$$\langle Tf, g \rangle \lesssim \sum_{Q \in \mathcal{S}} |Q| \langle f \rangle_{p,5Q} \langle g \rangle_{r,5Q}.$$

Then for every $\omega \in A_{\frac{q}{p}} \cap RH_{(\frac{q}{p})}'$ and $f \in L^p(\omega)$ we have

$$\|Tf\|_{L^p(\omega)} \lesssim \left([\omega]_{A_{\frac{q}{p}}} [\omega]_{RH_{(\frac{q}{p})}'}^3 \right)^\alpha,$$

with $\alpha = \max \left\{ \frac{1}{q-p}, \frac{r-1}{r-q} \right\}$.

Remark 3.2.4: There are other weighted bounds that can be obtained from bilinear bounds. For example, [Li17] generalizes proposition 3.2.3 and considers two weight inequalities.

³The constant $[\omega]_{RH_q}$ is related to the Reverse Hölder Inequality (lemma 1.5.2), which is a very important property of A_p weights. It represents the best constant when $1 + \epsilon = q$, so it is formally defined as $\sup_Q \left(\frac{1}{|B|} \int_B w(x)^q dx \right)^{1/q} \left(\frac{1}{|B|} \int_B w(x) dx \right)^{-1}$.

There is a particularly simple situation where we can obtain bilinear bounds for operators in an elementary way. This is the case of operators which are localized at a single scale, and we only need to require two properties:

Definition 3.2.5 (L^p improving estimates): We say T satisfies a L^p improving estimate if it is bounded from L^p to L^r , with $r > p$.

Definition 3.2.6 (Localized at scale s): We say that T is localized at scale s if there is an $\alpha > 1$ such that for any cube Q with $\ell(Q) \leq s$ we have $\langle T(f\chi_Q), g \rangle = \langle T(f\chi_Q), g\chi_{\alpha Q} \rangle$, for every f, g .

Given these two properties, we are able to obtain a (p, r) bilinear bound, where r is such that an L^p improving estimate from L^p to L^r holds. This is the content of the next proposition:

Proposition 3.2.7: Let T be a linear operator which satisfies an L^p improving estimate $L^p \rightarrow L^r$ and is localized at scale s . Then, T satisfies a (p, r') bilinear bound.

Proof: Let $\mathcal{Q}$ be a collection of disjoint cubes of size s which cover the space. Then, since T is linear we have

$$|\langle Tf, g \rangle| \leq \sum_{Q \in \mathcal{Q}} |\langle T(f\chi_Q), g \rangle| = \sum_{Q \in \mathcal{Q}} |\langle T(f\chi_Q), g\chi_{\alpha Q} \rangle|,$$

where we used the fact that T is localized at scale s to introduce the term $g\chi_{\alpha Q}$, which will be crucial to introduce an average. We now want to apply Hölder's inequality, and this is where L^p improving estimates are key: if we were to apply (p, p') Hölder, then we would get a (p, p') bilinear bound, which is not helpful. So, we apply (r, r') Hölder, since T can take the L^r norm and turn it into a L^p norm:

$$\sum_{Q \in \mathcal{Q}} |\langle T(f\chi_Q), g\chi_{\alpha Q} \rangle| \leq \sum_{Q \in \mathcal{Q}} \|T(f\chi_Q)\|_{L^r} \|g\chi_{\alpha Q}\|_{L^{r'}} \leq \sum_{Q \in \mathcal{Q}} \|f\chi_Q\|_{L^p} \|g\chi_{\alpha Q}\|_{L^{r'}}.$$

First, note that since $\mathcal{Q}$ consists of disjoint cubes, $\{\alpha Q : Q \in \mathcal{Q}\}$ is a sparse family, with $E(\alpha Q) = Q$. So, we can use $\|f\chi_Q\|_{L^p} \leq \|f\chi_{\alpha Q}\|_{L^p}$ so that the integrals of f and g are taken over the same cube. All that is left to do is introducing the averages and making sure we get a $|\alpha Q|$ term:

$$\sum_{Q \in \mathcal{Q}} \|f\chi_{\alpha Q}\|_{L^p} \|g\chi_{\alpha Q}\|_{L^{r'}} = \sum_{Q \in \mathcal{Q}} |\alpha Q|^{\frac{1}{p}} \langle f \rangle_{p, \alpha Q} |\alpha Q|^{\frac{1}{r'}} \langle g \rangle_{r', \alpha Q}.$$

Now, recall we chose $\ell(Q) = s$, so $|\alpha Q| = (\alpha s)^n$, so $|\alpha Q|^{\frac{1}{p}} |\alpha Q|^{\frac{1}{r'}} = |\alpha Q| |\alpha Q|^{-1 + \frac{1}{p} + \frac{1}{r'}} = |\alpha Q| (s\alpha)^{\frac{n}{p} - \frac{n}{r'}}$, which is a constant, so we get:

$$\sum_{Q \in \mathcal{Q}} |\alpha Q|^{\frac{1}{p}} \langle f \rangle_{p, \alpha Q} |\alpha Q|^{\frac{1}{r'}} \langle g \rangle_{r', \alpha Q} = (s\alpha)^{\frac{n}{p} - \frac{n}{r'}} \sum_{Q \in \mathcal{Q}} |\alpha Q| \langle f \rangle_{p, \alpha Q} \langle g \rangle_{r', \alpha Q},$$

proving the (p, r') bilinear bound. □

Remark 3.2.8: One can see easily in the proof that it is enough to require that the bound in definition 3.2.5 holds for any $f\chi_Q$, with $\ell(Q) \leq s$.

We now show a prototypical example for this proposition and prove it satisfies L^p improving and is localized at scale s , from which we automatically get a bilinear bound.

Example 3.2.9: Let us consider the operator $A_s f(x) := \frac{1}{|B(x,s)|} \int_{B(x,s)} f(y) dy = \frac{\chi_{B(0,s)}}{|B(0,s)|} \star f$. This operator is clearly linear, so we show it satisfies the two necessary properties. Let us fix any p and then any $p < r < \infty$.

- By Young's inequality for convolutions,

$$\|A_s f(x)\|_{L^p} = \left\| \frac{\chi_{B(0,s)}}{|B(0,s)|} \star f \right\|_{L^p} \leq \left\| \frac{\chi_{B(0,s)}}{|B(0,s)|} \right\|_{L^1} \|f\|_{L^p} = \|f\|_{L^p}.$$

On the other hand, we have

$$\left| \frac{1}{|B(x,s)|} \int_{B(x,s)} f(y) dy \right| \leq |B(x,s)|^{-1} \|f\|_{L^p} \|\chi_{B(x,s)}\|_{L^{\frac{1}{p'}}} = |B(x,s)|^{-1+\frac{1}{p'}} \|f\|_{L^p} \lesssim s^{\frac{n}{p}} \|f\|_{L^p},$$

with an arbitrary x , so we get a bound for $\|f\|_{L^\infty}$. This means we can use interpolation, with

$$\frac{1}{r} = \frac{\theta}{p} + \frac{1-\theta}{\infty} \implies \theta = \frac{p}{r}, \text{ to get}$$

$$\|A_s f(x)\|_{L^r} \leq \|A_s f(x)\|_{L^p}^{\frac{p}{r}} \|A_s f(x)\|_{L^\infty}^{1-\frac{p}{r}} \leq \|f\|_{L^p}^{\frac{p}{r}} s^{-\frac{n}{p}(1-\frac{p}{r})} \|f\|_{L^p}^{1-\frac{p}{r}} = s^{\frac{n}{r}-\frac{n}{p}} \|f\|_{L^p},$$

which not only proves the L^p improving estimate but it does so with the constant $s^{\frac{n}{r}-\frac{n}{p}}$, which deletes the dependence on s from proposition 3.2.7.

- Since $A_s f(x)$ takes an average of f on $B(x,s)$, if $f\chi_Q$ is supported in Q with $\ell(Q) \leq s$, then any point x outside of sQ will be at least s away from the support of $f\chi_Q$, meaning $B(x,s) \cap \text{supp}(f\chi_Q)$ has zero measure and therefore $A_s(f\chi_Q)(x) = 0$ in that case. This means that the support of $A_s(f\chi_Q)$ is contained in sQ , and therefore $\langle A_s(f\chi_Q), g \rangle = \langle A_s(f\chi_Q), g\chi_{sQ} \rangle$, so A_s is localized at scale s .

As a result, proposition 3.2.7 implies that A_s satisfies a (p, r') bound, which implies bounds for A_s for q in $p < q < r$.

In this case, we had a lot of freedom to choose p, r , so while this is an example where bilinear sparse domination is easy to prove, it is not an example of an operator which benefits from the versatility of these bounds regarding restricted strips of boundedness.

While the properties required in 3.2.7, it is often possible to apply this idea to operators that are not localized at a single scale, by breaking them up into single-scale pieces (typically restricting the domain using powers of two), as long as the bounds for each scale include some decay that then allows summation over the scales. This is a strategy used to prove single-scale bounds for operators, sometimes combined with other ideas, for example the techniques used to prove the pointwise bound. For example, such a mix of techniques is used in [BFP16] to find a bilinear bound for non-integral singular operators, and apply it to several operators, like Riesz Transforms and Fourier Multipliers.

Several other operators have also had sparse bounds proven via single-scale inspired techniques, including Spherical Maximal Functions [Lac19], the Bilinear Spherical Maximal Function [Bor+23], Pseudodifferential Operators [BC20a], Bochner-Riesz Multipliers [LMR19], certain Radon Transforms [Obe19] and Oscillatory and Random Singular Integrals [LS17].

In addition to these papers where specific operators are treated, there are a few general/abstract theorems on bilinear Sparse Domination:

- In [Con+17], an abstract theorem of domination by bilinear sparse forms is proven for bilinear forms of the form $\Lambda(f_1, f_2) = \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K(x, y) f_1(y) dy f_2(x) dx$ under certain conditions and then used to study rough singular integrals.
- In [Bar17], the theorem in [Con+17] is generalized to a multilinear setting and used to study specific operators, such as Rough Bilinear Singular Integrals or multilinear Calderón-Zygmund operators.
- In [Con+22], general metric spaces of homogeneous type are considered instead of just $\mathbb{R}^n$ using the ideas of single-scale L^p improving introduced above. This leads to a general theorem, which is then followed by applications, such as Calderón Zygmund theorem and Radon Transforms.
- In [BRS24], proposition 3.2.7 is proved and used as a base case for a more general theorem on Multi-scale operators. This general theorem is then applied to several operators, including Maximal Operators, Square Functions and Fourier Multipliers.

3.3 Weak and Endpoint Bounds

So far, we have focused on the cases where $1 < p < \infty$, where pointwise Sparse Bounds prove the respective (oftentimes sharp) quantitative weighted $L^p(\omega)$ bounds and bilinear Sparse Bounds prove similar bounds for strips $r < p < q$. In this chapter, we address some of the work done for endpoints and other related results.

In chapter 2, we saw that a big advantage of Sparse Forms is that they have the “correct” weighted bounds for $1 < p < \infty$, in the sense that they agree with the bounds we expect for Calderón-Zygmund operators. In general, Calderón-Zygmund operators are not bounded from $L^1(\omega)$ to $L^1(\omega)$, but they do satisfy a weak $(1, 1)$ bound. It turns out, Sparse Bounds are able to recover this bound as well. In the following proposition, we prove this for a $(1, r)$ bilinear bound in the unweighted case, following the proof in appendix B of [Con+17].

Proposition 3.3.1: Let T be a sublinear operator such that there exist $1 \leq r < \infty$ and $C > 0$ such that for every f, g bounded of compact support there exists a sparse family of cubes $\mathcal{S}$ such that

$$|\langle Tf, g \rangle| \leq C \sum_{Q \in \mathcal{S}} |Q| \langle f \rangle_{1, Q} \langle g \rangle_{r, Q}.$$

Then T satisfies a weak $(1, 1)$ bound.

Proof: We prove this bound via duality for the weak L^1 bound. Namely, it is enough to prove

$$\sup_{\|f\|_{L^1}=1} \sup_{E \subseteq \mathbb{R}^d} \inf_{\substack{E' \subseteq E \\ |E| \leq 2|E'|}} \sup_{|g| \leq 1_{E'}} |\langle Tf, g \rangle|, \quad (3.1)$$

where g is measurable and E has finite measure. Indeed, we want to show that for any f with $\|f\|_{L^1} = 1$ we have

$$\lambda |\langle x : Tf(x) > \lambda \rangle| \lesssim C$$

Let λ be fixed, and let $E_1 = \{x : Tf(x) > \lambda\}$, $E_2 = \{x : Tf(x) < -\lambda\}$, so that $\{x : |Tf(x)| > \lambda\} = E_1 \cup E_2$. Then, it is enough to show $\lambda |E_i| \lesssim C$ for $i = 1, 2$.

Assuming $|E_i| < \infty$ (we treat $|E_i| = \infty$ below) we know there is some $E'_i \subseteq E_i$ with $|E_i| \leq 2|E'_i|$ such that

$$\sup_{g \leq 1_{E'_i}} |\langle Tf, g \rangle| \leq C + \varepsilon \Rightarrow |\langle Tf, \chi_{E'_i} \rangle| \leq C + \varepsilon.$$

Since $E'_i \subseteq E_i$, $|Tf| \geq \lambda$ on E'_i , so $|\langle Tf, \chi_{E'_i} \rangle| \geq \lambda |E'_i|$. This implies $\lambda |E_i| \leq 2\lambda |E'_i| \leq 2|\langle Tf, \chi_{E'_i} \rangle| \leq 2C + 2\varepsilon$, so taking $\varepsilon = C$ gives $\lambda |E_i| \leq 4C$, proving the weak bound.

If $|E_i| = \infty$, the result follows by applying the above reasoning to $|E_i \cap B_M|$, which has finite measure, whenever it has positive measure, to get $|E_i \cap B_M| \leq 4C$ for every M . Since $E_i \cap B_M \nearrow_M E_i$ this implies $|E_i| \leq 4C$ as well by monotone convergence.

So, now we wish to show (3.1) holds. Let us fix $f \in L^1$ with $\|f\|_{L^1} = 1$ and E a set of finite positive measure. We can use the sparse bound to prove that the equivalent of (3.1) holds with $|\langle C \sum_{Q \in \mathcal{S}} |Q| \langle f \rangle_{1,Q} \langle g \rangle_{r,Q} \rangle|$ instead. We know from previous Sparse Form bounds such as the ones in subsection 2.2 that we can make an integral of the maximal operators appear, to which we can apply Hölder with a convenient choice of p . In this case, we will get $M_r(g) = \sup_Q \langle g \rangle_{r,Q}$ from $\langle g \rangle_{r,Q}$ instead. So we wish to find bounds for Mf and $M_r g$, using the fact that g is quite integrable and we can control its support.

Let us first note that while g is integrable for every choice of p , $M_r g$ introduces a power of r which makes it so that L^p bounds for M_r are only guaranteed to hold for $p > r$ (recall the discussion about M_r right before proposition 3.1.8). Now, since $g \leq \chi_{E'}$,

$$\|M_r g\|_{L^p} \lesssim \|g\|_{L^p} \leq \|\chi_{E'}\|_{L^p} = |E'|^{\frac{1}{p}} \leq |E|^{\frac{1}{p}},$$

for any $p > r$ and any $E' \subseteq E$ we choose.

Regarding Mf , we know it will be weak L^1 by the weak $(1, 1)$ bound for the maximal function. This will be very helpful to get an $L^{p'}$ bound: not only does it provide a lower endpoint for L^p interpolation, but it also guarantees that if we restrict Mf to a subset of the form $\{x : Mf(x) \leq C\}$ as to use L^∞ norm as an upper endpoint, the remaining subset has relatively small measure, which hopefully allows us to choose the support of g in such a way that we can safely ignore these points. Let $A = \{x : Mf(x) \leq C\}$,

with some fixed C we will choose later. Then, we get

$$\|Mf\|_{L^{p'}(A)} \leq \|Mf\|_{L^\infty(A)}^{1-\frac{1}{p'}} \|Mf\|_{L^{1,\infty}}^{\frac{1}{p'}} \lesssim C^{1-\frac{1}{p'}} \|f\|_{L^1}^{\frac{1}{p'}} = C^{1-\frac{1}{p'}} = C^{\frac{1}{p}}.$$

If we want to bound the Sparse Form by a constant, we want to choose $C = c|E|^{-1}$, with c to be chosen appropriately, to cancel out the $|E|^{\frac{1}{p}}$ from the contribution of g . Note that with this choice we have $|A^c| \lesssim |E|/c$, so with a nice choice of c it is small in comparison to $|E|$, which is promising since we need $|E| \leq 2|E'|$.

In order to use the bounds above, we need to be able to restrict from Q not only to E_Q but to $E_Q \cap A$, and we already have $|Q| \leq \frac{1}{\eta}|E_Q|$, so we need a similar bound for $|E_Q \cap A|$, namely $|E_Q \cap A| \geq \tilde{c}|E_Q|$, for which it is enough to prove a bound of the form $|E_Q \cap A^c| \leq (1 - \tilde{c})|E_Q|$, with $0 < \tilde{c} < 1$. We have $|E_Q \cap A^c| \leq |Q \cap A^c|$, so if we ensure $|Q \cap A^c| \leq (1 - \tilde{c})\eta|Q|$ for every Q in the sparse form then we get that

$$|E_Q \cap A^c| \leq |Q \cap A^c| \leq (1 - \tilde{c})\eta|Q| \leq (1 - \tilde{c})|E_Q| \implies |E_Q \cap A| \geq \tilde{c}|E_Q| \geq \tilde{c}\eta|Q|. \quad (3.2)$$

So, we want to be able to restrict to cubes for which $|Q \cap A^c| \leq (1 - \tilde{c})\eta|Q|$, but if we take the union of all cubes for which this inequality fails, in order to try to find a convenient E' , there is no reason why we should be able to make its measure small relative to E , since there is a lot of overlap. However, we have just the tool to deal with unwanted overlap: considering dyadic cubes instead. We will see this is enough, by taking a cube Q for which $|Q \cap A^c| > (1 - \tilde{c})\eta|Q|$ and considering the (at most) 2^n cubes Q_i of size $\frac{1}{2}\ell(Q) < \ell(Q_i) \leq \ell(Q)$ which cover Q . Since $\sum_{i=1}^{2^n} |Q_i \cap Q \cap A^c| = |Q \cap A^c| > (1 - \tilde{c})\eta|Q|$, there must be a term in the sum such that $|Q_i \cap Q \cap A^c| > 2^{-n}(1 - \tilde{c})\eta|Q|$, so we have

$$|Q_i \cap A^c| > 2^{-n}(1 - \tilde{c})\eta|Q| \geq 2^{-n}(1 - \tilde{c})\eta|Q_i|,$$

which means that Q_i is a dyadic cube which satisfies a similar bound, with a slightly different constant. It is easy to check that $3Q_i$ contains Q , so we can consider

$$\mathcal{Q} := \{Q \text{ maximal dyadic} : |Q \cap A^c| > 2^{-n}(1 - \tilde{c})\eta|Q|\} \text{ and } B := \cup_{Q \in \mathcal{Q}} 3Q$$

and take the support of g to be outside of B . If we do, then by the reasoning above any Q with $|Q \cap A^c| > (1 - \tilde{c})\eta|Q|$ would be inside one of these dyadic cubes in $\mathcal{Q}$, so it would be contained in the triple of the maximal dyadic containing it, which would imply $Q \subseteq B$. So, any g with support in $E \setminus B$ would be identically zero in Q , and so the associated term in the sparse form would be 0, so it can be omitted.

Before we put all the pieces together, all that's left to do is work out the constants. Let us fix $\tilde{c} = \frac{1}{2}$ so that $(1 - \tilde{c})\eta = \frac{\eta}{2}$. Then,

$$|B| \leq \sum_{Q \in \mathcal{Q}} |3Q| \leq 3^n \sum_{Q \in \mathcal{Q}} |Q| \leq 3^n 2^{-n-1} \eta \sum_{Q \in \mathcal{Q}} |Q \cap A^c| \leq 3^n 2^{-n-1} \eta |A^c|,$$

since we chose $\mathcal{Q}$ in such a way that the cubes in the sum are disjoint. Now, by the note above, $|A^c| \lesssim |E|/c$, so by choosing $c = 3^n$ we get that $|B| \leq 2^{-n-1}\eta|E|$, and so $E' := E \setminus B$ has measure at

least $\frac{1}{2}|E|$, as intended.

So, we can use (3.2) to restrict our sparse forms to A , using the same methods as in subsection 2.2:

$$\begin{aligned} |\langle Tf, g \rangle| &\leq C \sum_{Q \in \mathcal{S}} |Q| \langle f \rangle_{1,Q} \langle g \rangle_{r,Q} \leq \frac{2}{\eta} |E_Q \cap A| \langle f \rangle_{1,Q} \langle g \rangle_{r,Q} \lesssim \int_A Mf \cdot M_r g \lesssim \\ &\lesssim \|Mf\|_{L^p} \|M_r g\|_{L^{p'}(A)} \lesssim |E|^{-\frac{1}{p}} |E|^{\frac{1}{p}} \lesssim 1, \end{aligned}$$

using the bounds for Mg and $M_r f$ obtained above and finishing the proof. $\square$

It turns out, in the $p = 1$ case we are also able to get weighted bounds, which was first proven for Calderón-Zygmund operators in [HP15], and later improved in different ways in [DLR16], [FN19] and [NS24]. In particular, the following theorem is proven in [FN19], with a simplified proof appearing in [NS24]:

Theorem 3.3.2: Let T satisfies Sparse Domination, in the sense that for any bounded compactly supported f and any $g \in L^1$ there is a sparse collection $\mathcal{S}$ such that

$$\int_{\mathbf{R}^d} |Tf| |g| dx \lesssim \sum_{Q \in \mathcal{S}} \langle f \rangle_{1,Q} \langle g \rangle_{1,Q} |Q|,$$

then if $w \in A_1$ then

$$\|Tf\|_{L_w^{1,\infty}(\mathbf{R}^d)} \lesssim [w]_1 (1 + \log[w]_{FW}^4) \|f\|_{L_w^1(\mathbf{R}^d)}$$

for every $f \in L^1$.

Remark 3.3.3: These papers contain more endpoint results beyond just weak $(1, 1)$ bounds, such as weak bounds in the endpoints of (r, s) bilinear sparse domination introduced in the previous subsection. As an example, [NS24] contains these weak bounds, as well as references to related work.

As it often happens, beyond these general results endpoint bounds end up being operator-dependent and studied case-by-case. There are a few papers where endpoints are treated for specific operators, using Sparse Domination to varying degrees, such as Bochner-Riesz multipliers on the plane [KL18], Walsh-Fourier multipliers of Marcinkiewicz type [Che+18], nonhomogeneous trees [Con+24], the composition of Calderon-Zygmund operators [Hu20], Commutators of Rough Singular Integral Operators [LTH20] and Classes of Multiplier Transformations [BRS23].

3.4 Multiparameter Operators

It is a well-known fact that multiparameter operators tend to be more complicated than their single parameter counterparts. The strong maximal function (where averages are taken over rectangles with sides parallel to the axes instead of cubes) is a good example of this added complexity, since it fails to satisfy the weak $(1, 1)$ bound we've come to expect of maximal functions.

⁴The term $[w]_{FW}$ denotes the Fujii-Wilson constant $\sup_Q \frac{1}{w(Q)} \int_Q M(w\chi_Q) dx$ proposed by Fujii and later Wilson as the definition of $[w]_\infty$, and is sometimes denoted by $[w]_\infty$ instead.

Even then, it is a natural question to consider whether sparse domination methods can be applied in the multiparameter setting. In this section, we will see that even bounding the strong maximal function - a relatively simple multiparameter operator - by $(1, 1)$ bilinear (bi-parameter) sparse forms with dyadic rectangles turns out to be impossible. This was proven in [Bar+19], in which the following exposition was based.

Given the bi-parameter character of the strong maximal function M_S , it is intuitive that we could never get a bound while only considering sparse families of cubes, since they could never capture the behavior of thin, tall rectangles. Indeed, by proposition 3.3.1 such a bound would imply that the strong maximal function satisfies a weak bound. So, we consider sparse forms associated to rectangles with sides parallel to the axis instead. It is possible to adapt the proof of proposition 3.3.1 to this case, to get that such a bound would imply the strong maximal function satisfies the bound

$$|\{x : M_S f(x) > \lambda\}| \leq \int \frac{|f(x)|}{\lambda} \left(1 + \log^+ \left(\frac{|f(x)|}{\lambda}\right)\right). \quad (3.3)$$

In this sense, it would be plausible that a bound of this form exists, since it is well known that the strong maximal function does satisfy the bound above (see [JMZ35; CF75]).

A possible way to prove the failure of a weak $(1, 1)$ bound for M_S is via a dirac delta, made formal via a sequence of functions converging to it (see [SM93], chapter 10, section 2.3). It is therefore natural to consider a dirac delta if we are investigating a possible sparse bound for M_S .

In this case, there is a simple characterization for the strong maximal function, $M_S(\delta)(x, y) = \sup_{(0,0),(x,y) \in R} \frac{1}{|R|} = \frac{1}{|x||y|}$. We also have that $\langle \delta \rangle_R = \frac{1}{|R|}$ whenever $(0, 0) \in R$, so if we want to bound $M_S(\delta)(x, y)$ by a sparse form the easiest thing to do is to find $|R|$ such that a multiple of $\frac{1}{|R|}$ is greater than $\frac{1}{|x||y|}$.

Let us assume to simplify notation $x, y \geq 0$, as the other cases are analogous. There are maximal powers of two such that $2^a < x \leq 2^{a+1}$ and $2^b < y \leq 2^{b+1}$, so $\frac{1}{xy} \leq \frac{2^a}{2^b}$. Even though we do not have $(x, y) \subseteq [0, 2^a] \times [0, 2^b]$ (so we can't use this rectangle as R above, we do have $(x, y) \subseteq [0, 2^{a+1}] \times [0, 2^{b+1}]$ and $\frac{1}{xy} \leq \frac{2^a}{2^b} \implies \frac{1}{xy} \leq 4 \frac{2^{a+1}}{2^{b+1}}$. This gives us

$$M_S(\delta)(x, y) \leq \sum_{a,b \in \mathbb{Z}} \langle \delta \rangle_{[0,2^a] \times [0,2^b]} \chi_{[0,2^{a+1}] \times [0,2^{b+1}]}(x, y),$$

which is a sparse bound, as one can see that taking $E([0, 2^a] \times [0, 2^b]) = [2^{a-1}, 2^a] \times [2^{b-1}, 2^b]$ makes this family $\frac{1}{4}$ -sparse.

So, it again seems plausible such a bound could exist in general. However, it turns out that we can disprove even the bilinear formulation of sparse domination with dyadic rectangles for M_S , via a clever choice of collections of dirac deltas at different points. In what follows, we will not worry about the limiting arguments and will work directly with these measures. The theorem we will prove, assuming a technical lemma which is outside of the scope and we present a reference for, is the following:

Theorem 3.4.1: Let C be any constant and $\mathcal{S}$ any sparse family of dyadic rectangles. Then, there exist

measures μ, ν such that

$$|\langle \mathcal{M}_S \mu, \nu \rangle| \geq C \sum_{R \in \mathcal{S}} \langle |\mu| \rangle_R \langle |\nu| \rangle_R |R|.$$

Remark 3.4.2: The dyadic rectangles in this theorem are with respect to the standard dyadic grid in $\mathbb{R}^2$ and in particular have sides parallel to the axes.

Remark 3.4.3: The analogous theorem holds for compactly supported integrable functions instead of measures via a limiting argument which we omit, similar to the one in [SM93], chapter 10, section 2.3.

Let us consider μ, ν which are both sums of point masses at points⁵ $p_1, \dots, p_n$ and $z_1, \dots, z_m$. Then,

$$\langle M_S(\mu), \nu \rangle = \sum_{i=1}^m M_S(\mu)(z_i) = \sum_i \sup_{R \ni z_i} \sum_{j=1}^n \frac{1}{|R|} \delta_{p_j}(R),$$

so this quantity is larger when more of the p_j are contained in small rectangles containing each z_i . If we want to find examples where $\langle M_S(\mu), \nu \rangle$ is too large to compensate for, then we want that for each z_i there is at least one p_j near it. It is however not very helpful to have more p_j s in the same rectangle, as $\langle M_S(\mu), \nu \rangle$ would only increase by a factor of the number of p_j , whereas no new rectangles would be needed from $\sum_{R \in \mathcal{S}} \langle \mu \rangle_R \langle \nu \rangle_R$ to cover that term. The idea is then to have many q_j , each near a p_i , but such that they are far away from each other, so that the right-hand-side has a hard time covering everything.

Morally, the idea is that $\langle M_S(\mu), \nu \rangle$ sees how close the p_j s and q_i s are to one another, but the distances within each set does not make a lot of difference, whereas $\langle \mu \rangle_R \langle \nu \rangle_R$ covers the p_j s and q_i s separately without exploiting this closeness of the sets, so having the p_j s and q_i s far from each other within each set will require either a collection of larger rectangles or a collection with a large amount of smaller rectangles densely packed (since the q_i are all close to common points, making the rectangles necessarily close to each other in some sense) to cover them all. These collections will respectively not be able to compete with the reciprocal of the small rectangles $\langle M_S(\mu), \nu \rangle$ sees and not be able to be sparse since they are so densely packed. It is easy to see why this problem is very particular to the bi-parameter case, as it allows points to be close to a common third point, but far from each other.

So, we want to find two sets of points $\mathcal{P}$ and $\mathcal{Z}$ with the properties above. Before we define them, we note that given two points, it is important to know the dimensions of the smallest dyadic cube containing both, or at least its area. So, we will define a notion of “distance”⁶ which keeps track of that information:

Definition 3.4.4: Given two points p, q , we define $\text{dist}(p, q) := \min\{|R|^{\frac{1}{2}} : p, q \in R, R \text{ dyadic cube}\}$. In particular, whenever $p \neq q$ the infimum is a minimum and $\text{dist}(p, q)^2$ is the area of the smallest dyadic rectangle containing both p and q , which is an integer power of two.

We now define the first type of set, $\mathcal{P}_m$, which is the simplest. All that we require is that we have a high enough number of points in the unit square, while keeping them far enough from each other. The existence of such a set is content of the following lemma:

⁵In reality, we will want the average, so that the measures have “the same size” regardless of the number of points. However, for the sake of simplifying the math in the motivating section we consider sums.

⁶This “distance” is not an actual distance, as it fails the triangle inequality; imagine two intersecting thin dyadic rectangles, one horizontal and one vertical, and three points in the endpoints.

Lemma 3.4.5: For every natural m , there is a set $\mathcal{P}_m \subseteq [0, 1]^2$, which satisfies the following two properties:

(P.1) $\text{dist}(p, q) \geq 2^{-m}$ for every $p \neq q$ in $\mathcal{P}_m$.

(P.2) $\#(\mathcal{P}_m) = 2^{2m+1}$

We will sketch how to construct $\mathcal{P}_m$, starting on $m = 0$ and inducting up. When $m = 0$, we want two points at distance 1, which is easy to achieve by choosing any points in the two blue squares in the following picture:

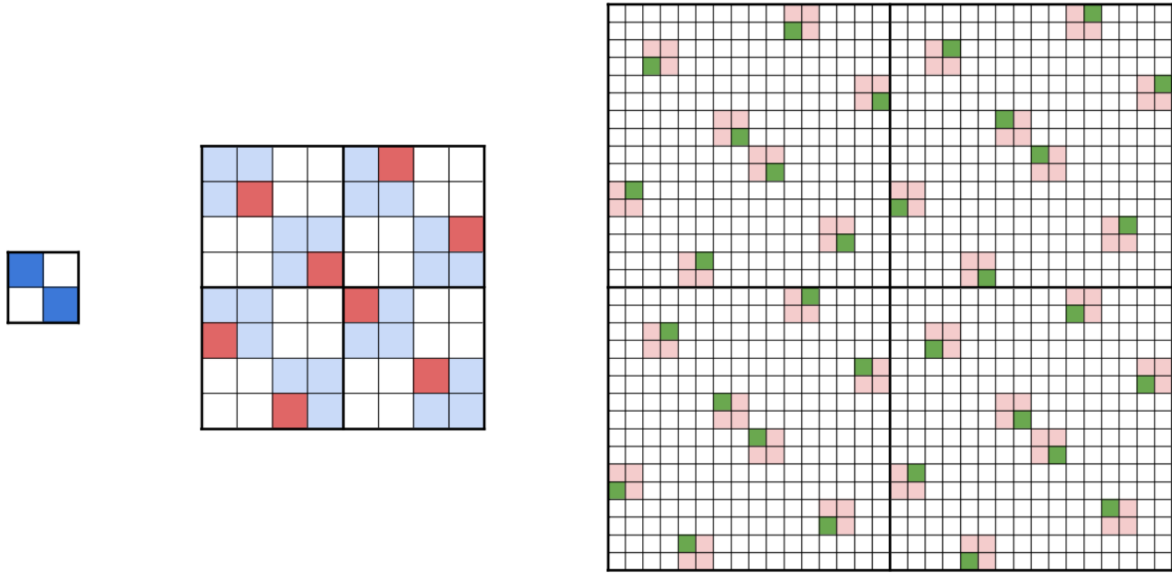


Figure 3.1: Examples of possible $\mathcal{P}_0, \mathcal{P}_1, \mathcal{P}_2$: to obtain these sets, choose any point inside each of the blue, red and green squares, respectively.

To obtain $\mathcal{P}_m$ from $\mathcal{P}_{m-1}$, divide $[0, 1]^2$ into its four dyadic children and repeat the pattern of squares of $\mathcal{P}_{m-1}$ in each of them. This is illustrated in light blue in $\mathcal{P}_1$ and light red in $\mathcal{P}_2$. One can prove that any pair of points in different squares are at 2^{-m} away, possibly except for the 4 different copies of each square, which depending on the choice of points can have a pair of points in a common $2^{-m-1} \times 1$ or a $1 \times 2^{-m-1}$ rectangle. To ensure that does not happen, think of the four children of each light-colored square, and choose a child for each in such a way that they are not in the same row or column. This is illustrated in darker red in $\mathcal{P}_1$ and in green in $\mathcal{P}_2$. A formal proof of the construction can be found in [Bar+19], section 3.1.

The second type of sets $\mathcal{Z}_m$ is defined so that it has a good number of points (this will be property (Z.1) below), each of which is close to exactly one point in $\mathcal{P}_m$ (property (Z.2) below), while still being roughly at a specified distance from it (property (Z.3)). We also need that there are not too many points of $\mathcal{Z}_m$ inside any given rectangle, so we ensure a bound on $\mathcal{Z}_m \cap R$ depending on the size of R (properties (Z.4) and (Z.5)). The following lemma ensures the existence of such a set:

Lemma 3.4.6: For every natural m , every $1 \ll k \ll m$ and $\mathcal{P}_m$ as in lemma 3.4.5, there is a set $\mathcal{Z}_m$ satisfying all the following properties, with constants independent of k, m :

$$(Z.1) \quad \#(\mathcal{Z}_m) \gtrsim m2^{2m}$$

$$(Z.2) \quad \text{For every } z \in \mathcal{Z}_m \text{ there is a unique } p \in \mathcal{P}_m, \text{ denoted } p(z), \text{ such that } \text{dist}(p(z), z)^2 < 2^{-2m-1}.$$

$$(Z.3) \quad \text{For every } z \in \mathcal{Z}_m, \text{dist}(p(z), z)^2 \sim 2^{-2m-k}.$$

$$(Z.4) \quad \text{For every dyadic rectangle } R \text{ with } |R| \leq 2^{-2m-2} \text{ which intersects } \mathcal{P}_m, \text{ we have } \#(R \cap \mathcal{Z}_m) \lesssim k.$$

$$(Z.5) \quad \text{For every dyadic rectangle } R \text{ with } |R| \geq 2^{-2m-1}, \text{ we have } \#(R \cap \mathcal{Z}_m) \lesssim 2^{2m}mk|R|.$$

We omit the proof of lemma 3.4.6, which can be found in [Bar+19], sections 3.2 and 3.3.

Using this sets, we can define our measures:

Definition 3.4.7: Given $\mathcal{P}_m$ and $\mathcal{Z}_m$ as in lemmas 3.4.5 and 3.4.6, we define

$$\mu_m = \frac{1}{\#\mathcal{P}_m} \sum_{p \in \mathcal{P}_m} \delta_p \quad \text{and} \quad \nu_m = \frac{1}{\#\mathcal{Z}_m} \sum_{z \in \mathcal{Z}_m} \delta_z.$$

We are now ready to prove the theorem, via two propositions, establishing respectively a lower bound for $\langle M_S(\mu_m), \nu_m \rangle$ and an upper bound for $\sum_{R \in \mathcal{S}} \langle \mu_m \rangle_R \langle \nu_m \rangle_R$ for any sparse family of dyadic rectangles.

Proposition 3.4.8: For any natural m and μ_m and ν_m as in definition 3.4.7, we have

$$\langle M_S(\mu_m), \nu_m \rangle \gtrsim 2^k.$$

Proof: Note that by definition of ν_m , $\langle M_S(\mu_m), \nu_m \rangle$ is the average of $M_S(\mu_m)(z)$ over all $z \in \mathcal{Z}_m$. So, it is enough to prove $M_S(\mu_m)(z) \gtrsim 2^k$ for every $z \in \mathcal{Z}_m$. Since M_S is the supremum of $\langle \mu_m \rangle_R = \frac{1}{\#\mathcal{P}_m} \frac{\#(\mathcal{P}_m \cap R)}{|R|}$ over $R \ni z$, so we can just use $\langle \mu_m \rangle_{R'}$, for an appropriate R as a lower bound.

By definition of $\mathcal{P}_m, \mathcal{Z}_m$, there is by (Z.2), (Z.3) a point $p(z)$ with $\text{dist}(p(z), z)^2 \sim 2^{-2m-k}$, so taking R' to be the rectangle corresponding to this distance, we have

$$M_S(\mu_m)(z) \geq \langle \mu_m \rangle_{R'} = \frac{1}{\#\mathcal{P}_m} \frac{\#(\mathcal{P}_m \cap R')}{|R'|} \gtrsim \frac{1}{2^{2m+1}} \frac{1}{2^{-2m-k}} = 2^{k-1} \gtrsim 2^k,$$

where we used (P.2) to bound $\#(\mathcal{P}_m)$, the existence of $p(z)$ to bound $\#(\mathcal{P}_m \cap R')$ below by 1 and (Z.3) to bound $|R'| \lesssim 2^{-2m-k}$. $\square$

Inspecting the proof, we can see that we only used the existence of a point $\mathcal{P}_m$ close enough to any point in $\mathcal{Z}_m$. For the next proposition, we will need to use all the other properties of the sets.

Proposition 3.4.9: Let $\mathcal{S}$ be an η -sparse family of dyadic rectangles, m any integer and ν_m, μ_m be as in definition 3.4.7. Then,

$$\sum_{R \in \mathcal{S}} \langle \mu_m \rangle_R \langle \nu_m \rangle_R |R| \lesssim \frac{1}{\eta} k \left(1 + \frac{2^{2k}}{m} \right)$$

Proof: Recall that by definition of $\mathcal{P}_m$ and $\mathcal{Z}_m$, the measures μ_m and ν_m are supported in $[0, 1]^2$. So, given a rectangle R , we define $\bar{R} = R \cap [0, 1]^2$, since all our information is concentrated in $[0, 1]^2$. Note $\bar{R}$ is also a dyadic rectangle, since it is the product of the intersection of the coordinates, and intersections of dyadic intervals are dyadic. Then, we have:

$$\langle \mu_m \rangle_R = \frac{\mu_m(R)}{|R|} = \frac{\mu_m(\bar{R})}{|\bar{R}|} \frac{|\bar{R}|}{|R|} = \langle \mu_m \rangle_{\bar{R}} \frac{|\bar{R}|}{|R|} \quad \text{and similarly} \quad \langle \nu_m \rangle_R = \langle \nu_m \rangle_{\bar{R}} \frac{|\bar{R}|}{|R|}.$$

We will now show that it is enough to find a bound for $\langle \mu_m \rangle_{\bar{R}} \langle \nu_m \rangle_{\bar{R}} \leq B$, and it extends to a bound for $\sum_{R \in \mathcal{S}} \langle \mu_m \rangle_R \langle \nu_m \rangle_R |R|$. Since both $R, \bar{R}$ are dyadic, $\frac{|\bar{R}|}{|R|}$ is a non-positive power of 2. So, we can split $\mathcal{S}$ into collections $\mathcal{S}_j := \{R \in \mathcal{S} : \frac{|\bar{R}|}{|R|} = 2^{-j}\}$. Let us define $R_j = \cup_{R \in \mathcal{S}_j} R$. Then, we have

$$\begin{aligned} \sum_{R \in \mathcal{S}} \langle \mu_m \rangle_R \langle \nu_m \rangle_R |R| &= \sum_{j \geq 0} \sum_{R \in \mathcal{S}_j} \langle \mu_m \rangle_R \langle \nu_m \rangle_R |R| \leq \frac{1}{\eta} \sum_{j \geq 0} \sum_{R \in \mathcal{S}_j} \langle \mu_m \rangle_{\bar{R}} \langle \nu_m \rangle_{\bar{R}} \left(\frac{|\bar{R}|}{|R|} \right)^2 |E(R)| \leq \\ &\frac{1}{\eta} \sum_{j \geq 0} \sum_{R \in \mathcal{S}_j} B 2^{-2j} |E(R)| \leq \frac{1}{\eta} B \sum_{j \geq 0} 2^{-2j} |R_j|, \end{aligned}$$

which means that all we need to do is bound $|R_j|$, in order to show the sum is finite and the bound $\leq B$ extends to the bilinear sparse form. We have that we can write $|R_j|$ as

$$\left| \left\{ x \in R \in \mathcal{S} : \frac{|\bar{R}|}{|R|} = 2^{-j} \right\} \right| \leq \left| \left\{ x \in \mathbb{R}^2 : \frac{|R \cap [0, 1]^2|}{|R|} = 2^{-j} \right\} \right| \leq \left| \left\{ x \in \mathbb{R}^2 : \frac{|R \cap [0, 1]^2|}{|R|} > 2^{-j-1} \right\} \right|.$$

The key observation is that $\frac{|R \cap [0, 1]^2|}{|R|}$ can be bounded above by $M_S(\chi_{[0, 1]^2})$. Even though M_S does not satisfy a weak $(1, 1)$ inequality, we can still use (3.3) to bound the measure:

$$\left| \left\{ x \in \mathbb{R}^2 : \frac{|R \cap [0, 1]^2|}{|R|} > 2^{-j-1} \right\} \right| \leq \left| \left\{ x \in \mathbb{R}^2 : M_s(\chi_{[0, 1]^2}) > 2^{-j-1} \right\} \right| \leq \int_{[0, 1]^2} \frac{1}{2^{-j-1}} \left(1 + \log^+ \left(\frac{1}{2^{-j-1}} \right) \right),$$

which is up to a log base change constant $(j+2)2^{j+1} \lesssim (j+1)2^{j+1}$. Using this bound, we get:

$$\frac{1}{\eta} B \sum_{j \geq 0} 2^{-2j} |R_j| \leq \frac{1}{\eta} B \sum_{j \geq 0} 2^{-2j} (j+1) 2^{j+1} \lesssim B \sum_{j \geq 1} 2^{-j} j \lesssim \frac{1}{\eta} B,$$

since the series $\sum_{j \geq 1} 2^{-j} j$ converges.

So, all we need to do now is show that for every $\bar{R}$ we have $\langle \mu_m \rangle_{\bar{R}} \langle \nu_m \rangle_{\bar{R}} \lesssim k \left(1 + \frac{2^{2k}}{m} \right)$. We will do it in two cases, $|\bar{R}| > 2^{-2m-1}$ and $|\bar{R}| \leq 2^{-2m-1}$ due in no small part to the differences that (Z.4), (Z.5) introduce:

($|\bar{R}| \geq 2^{-2m-1}$) Since

$$\langle \mu_m \rangle_{\bar{R}} = \frac{\mu_m(\bar{R})}{|\bar{R}|} = \frac{1}{\#(\mathcal{P}_m)} \frac{\#(\mathcal{P}_m \cap \bar{R})}{|\bar{R}|} \quad \text{and} \quad \langle \nu_m \rangle_{\bar{R}} = \frac{1}{\#(\mathcal{Z}_m)} \frac{\#(\mathcal{Z}_m \cap \bar{R})}{|\bar{R}|},$$

and we have lower bounds for $\#(\mathcal{P}_m)$, $\#(\mathcal{Z}_m)$ and $|\bar{R}|$, we need an upper bound for $\#(\mathcal{P}_m \cap \bar{R})$ and $\#(\mathcal{Z}_m \cap \bar{R})$. For the latter, we can use (Z.5), while for the former we need to do a little more

work, using (P.1) and (P.2):

We can split $[0, 1]^2$ into $|\bar{R}|^{-1}$ rectangles of size $|\bar{R}|$, each of which can have at most $|\bar{R}|/2^{-2m-1}$ points, since by (P.1) there can only be one inside each dyadic rectangle of size 2^{-2m-1} and we can consider the tiling into dyadic rectangles of size 2^{-2m-1} within the tiling of rectangles of size $|\bar{R}| \geq 2^{-2m-1}$ (otherwise the distance between two points would be at most the square root of 2^{-2m-1} , which is less than 2^{-m}). This means there are at most $|\bar{R}|^{-1}|\bar{R}|2^{2m+1} = 2^{2m+1}$ points in $[0, 1]^2$, and since by (P.2) we must have exactly 2^{2m+1} points, this means each rectangle, including $\bar{R}$, must have exactly $|\bar{R}|/2^{-2m-1} = |\bar{R}|2^{2m+1}$. Putting all of these bounds together, we get:

$$\begin{aligned}\langle \mu_m \rangle_{\bar{R}} &= \frac{1}{\#(\mathcal{P}_m)} \frac{\#(\mathcal{P}_m \cap \bar{R})}{|\bar{R}|} \leq \frac{1}{2^{2m+1}} \frac{|\bar{R}|2^{2m+1}}{|\bar{R}|} = 1, \\ \langle \nu_m \rangle_{\bar{R}} &= \frac{1}{\#(\mathcal{Z}_m)} \frac{\#(\mathcal{Z}_m \cap \bar{R})}{|\bar{R}|} \leq \frac{1}{m2^{2m}} \frac{2^{2m}mk|\bar{R}|}{|\bar{R}|} = k,\end{aligned}$$

yielding $\langle \mu_m \rangle_{\bar{R}} \langle \nu_m \rangle_{\bar{R}} \leq k \leq k \left(1 + \frac{2^{2k}}{m}\right)$.

($|\bar{R}| < 2^{-2m-1}$) Again, we have

$$\langle \mu_m \rangle_{\bar{R}} = \frac{\mu_m(\bar{R})}{|\bar{R}|} = \frac{1}{\#(\mathcal{P}_m)} \frac{\#(\mathcal{P}_m \cap \bar{R})}{|\bar{R}|} \quad \text{and} \quad \langle \nu_m \rangle_{\bar{R}} = \frac{1}{\#(\mathcal{Z}_m)} \frac{\#(\mathcal{Z}_m \cap \bar{R})}{|\bar{R}|}.$$

The bounds for $\#(\mathcal{P}_m)$ and $\#(\mathcal{Z}_m)$ still hold, but we need to find bounds for the remaining terms.

Since $|\bar{R}| < 2^{-2m-1}$, by (P.1) there can be at most one point of $\mathcal{P}_m$ in $\bar{R}$. If $\bar{R}$ does not contain at least one point in $\mathcal{Z}_m$ and one in $\mathcal{P}_m$, then $\langle \mu_m \rangle_{\bar{R}} \langle \nu_m \rangle_{\bar{R}} = 0$ and the bound is trivial. So, there is at least one $z \in \mathcal{Z}_m \cap \bar{R}$ and $p \in \mathcal{P}_m \cap \bar{R}$. Since $|\bar{R}| < 2^{-2m-1}$ and contains both z and p , by (Z.2) we must have $p = p(z)$ and so by (Z.3) we must have $|\bar{R}| \gtrsim 2^{-2m-k}$, which provides a lower bound for $|\bar{R}|$ we can use to get an upper bound. All we need now is an upper bound for $\#(\mathcal{Z}_m \cap \bar{R})$, for which we can just use (Z.4). Combining all these bounds, we get:

$$\begin{aligned}\langle \mu_m \rangle_{\bar{R}} &= \frac{1}{\#(\mathcal{P}_m)} \frac{\#(\mathcal{P}_m \cap \bar{R})}{|\bar{R}|} \leq \frac{1}{2^{2m+1}} \frac{1}{2^{-2m-k}} = 2^k, \\ \langle \nu_m \rangle_{\bar{R}} &= \frac{1}{\#(\mathcal{Z}_m)} \frac{\#(\mathcal{Z}_m \cap \bar{R})}{|\bar{R}|} \leq \frac{1}{m2^{2m}} \frac{k}{2^{-2m-k}} = \frac{k}{m} 2^k,\end{aligned}$$

which yields $\langle \mu_m \rangle_{\bar{R}} \langle \nu_m \rangle_{\bar{R}} \leq k \frac{k}{m} 2^{2k} \leq k \left(1 + \frac{2^{2k}}{m}\right)$, concluding the proof. $\square$

With propositions 3.4.8 and 3.4.9, we are now ready to prove theorem 3.4.1:

Proof (of theorem 3.4.1): Let $m = k2^{2k}$. Let there be a C for which a sparse bound holds. Then, we'd have:

$$2^k \lesssim \langle M_S(\mu_m), \nu_m \rangle \leq C \sum_{R \in S} \langle \mu_m \rangle_R \langle \nu_m \rangle_R |R| \lesssim \frac{C}{\eta} \frac{1}{k} \left(1 + \frac{2^{2k}}{k2^{2k}}\right) \leq \frac{2C}{\eta}.$$

This is a contradiction, since $2^k \rightarrow \infty$ as $k \rightarrow \infty$, but $2C\eta^{-1}$ is constant. $\square$

It is interesting to note that the ideas above can be extended to other settings. In [Bar+19], this is

done in three directions: in section 4.1 the equivalent theorem for the bi-parameter martingale transform is proven, in section 4.2 it is proven that an analogous bound can't hold for the n -dimensional Strong Maximal function, by showing that a bound for dimension n would yield a bound for dimension $n - 1$, and in section 4.3 other sparse forms are considered, namely those with Orlicz averages. In [FLW23], this proof is generalized in order to show that Zygmund type and Flag type maximal functions similarly allow no multiparameter sparse bound.

There is some work in [BP17] to try to establish Sparse Domination for the bi-parameter Martingale Transform, which we know by [Bar+19] requires some alternative method, since a bilinear Sparse Bound can't hold. The main new idea in [BP17] is the introduction of dyadic Square Functions, and even though it is possible to find weighted bounds with this approach, they are nowhere near sharp. In particular, they are worse than those one can obtain using Square Functions directly. Recall that the bi-parameter Martingale Transform and the dyadic Square function are given respectively by

$$Tf = \sum_{R \in \mathcal{D}} \epsilon_R \langle f, h_R \rangle h_R, \quad \sup_R |\epsilon_R| \leq C \quad \text{and} \quad Sf(x)^2 = \sum_{R \in \mathcal{D}} |\langle f, h_R \rangle|^2 \frac{1_R(x)}{|R|}.$$

In [BP17], the following theorem is proven:

Theorem 3.4.10: Let T be the bi-parameter martingale transform and S be the dyadic square function. Then, there is a sparse family of rectangles $\mathcal{S}$ such that

$$|\langle Tf, g \rangle| \lesssim \left(\sup_R |\epsilon_R| \right) \sum_{R \in \mathcal{S}} |R| \langle Sf \rangle_R \langle Sg \rangle_R.$$

One can use usual arguments to extract a weighed bound from this form, but we need to account for the square operator, using the bound $\|Sf\|_{L^2(w)} \lesssim [w]_{A_2}^2 \|f\|_{L^2(w)}$, which can similarly to $\|M_S f\|_{L^2(w)} \lesssim [w]_{A_2}^2 \|f\|_{L^2(w)}$ be established by iterating one-parameter results. As one can see, these dependences are already quadratic instead of linear, and they pile up when computing the full bound:

$$\begin{aligned} \sum_{R \in \mathcal{S}} |R| \langle Sf \rangle_R \langle Sg \rangle_R &\leq \frac{1}{\eta} \sum_{R \in \mathcal{S}} |E(R)| \langle Sf \rangle_R \langle Sg \rangle_R \leq \int M_S(S(f)) M_S(S(g)) \omega^{\frac{1}{2}} \omega^{-\frac{1}{2}} \leq \\ &\leq \|M_S(S(f))\|_{L^2(\omega)} \|M_S(S(g))\|_{L^2(\omega^{-1})} \leq [\omega]_2^4 \|S(f)\|_{L^2(\omega)} \|S(g)\|_{L^2(\omega^{-1})} \leq [\omega]_2^8 \|f\|_{L^2(\omega)} \|g\|_{L^2(\omega^{-1})}. \end{aligned}$$

By duality, we get the dependence $[\omega]_2^8$ in the A_2 constant. It is worth pointing out that it is known that $\|Tf\|_{L^2(w)} \lesssim [w]_{A_2}^3 \|f\|_{L^2(w)}$ holds for the Martingale Transform. This means that even without passing through the square function, the $[\omega]_2^4$ introduced by the bounds on M_S would already stop the estimate from being sharp. Similarly, as long as we need to pass through the square function, the $[\omega]_2^4$ term its bound introduces already stops the estimate from being sharp, which means a better bound on M_S would not be enough to get a sharp estimate here. With all of this in mind, it is not yet clear if it is possible to use Sparse Domination to get sharp multiparameter weighed bounds, or how one would go about doing it.

Bibliography

- [AIS01] Kari Astala, Tadeusz Iwaniec, and Eero Saksman. “Beltrami operators in the plane”. In: *Duke Mathematical Journal* 107.1 (2001), pp. 27–56. DOI: 10.1215/S0012-7094-01-10713-8. URL: <https://doi.org/10.1215/S0012-7094-01-10713-8>.
- [Bar+19] Alex Barron et al. “Sparse domination and the strong maximal function”. In: *Advances in Mathematics* 345 (2019), pp. 1–26.
- [Bar17] Alex Barron. “Weighted estimates for rough bilinear singular integrals via sparse domination”. In: *New York J. Math.* 23 (2017), pp. 779–811. ISSN: 1076-9803. URL: http://nyjm.albany.edu:8000/j/2017/23_779.html.
- [BC20a] David Beltran and Laura Cladek. “Sparse bounds for pseudodifferential operators”. In: *Journal d’Analyse Mathématique* 140 (2020), pp. 89–116.
- [BC20b] David Beltran and José M. Conde-Alonso. *Scheme Lectures: The Harmonic Analysis and PDE Research Network in the UK*. (Accessed 2025-06-20). 2020. URL: https://www.youtube.com/playlist?list=PLsDn5JyJXoYJh56xxK87JFxxu_Xrho0dY.
- [BFP16] Frédéric Bernicot, Dorothee Frey, and Stefanie Petermichl. “Sharp weighted norm estimates beyond Calderón–Zygmund theory”. In: *Analysis & PDE* 9.5 (2016), pp. 1079–1113.
- [BM18] Cristina Benea and Camil Muscalu. “The helicoidal method”. In: *Theta Ser. Adv. Math.* 20 (2018), pp. 45–96.
- [BM21] Cristina Benea and Camil Muscalu. “Sparse domination via the helicoidal method”. In: *Revista matemática iberoamericana* 37.6 (2021), pp. 2037–2118.
- [Bor+23] Tainara Borges et al. “Sparse bounds for the bilinear spherical maximal function”. In: *Journal of the London Mathematical Society* 107.4 (2023), pp. 1409–1449.
- [BP17] Alex Barron and Jill Pipher. “Sparse domination for bi-parameter operators using square functions”. In: *arXiv preprint arXiv:1709.05009* (2017).
- [BRS23] David Beltran, Joris Roos, and Andreas Seeger. “Endpoint sparse domination for classes of multiplier transformations”. In: *Mathematische Zeitschrift* 305.1 (2023), p. 11.
- [BRS24] David Beltran, Joris Roos, and Andreas Seeger. *Multi-scale sparse domination*. Vol. 298. 1491. American Mathematical Society, 2024.

- [Buc93] Stephen M. Buckley. “Estimates for operator norms on weighted spaces and reverse Jensen inequalities”. In: *Transactions of the American Mathematical Society* 340.1 (1993), pp. 253–272.
- [CDO18] Amalia Culiuc, Francesco Di Plinio, and Yumeng Ou. “Domination of multilinear singular integrals by positive sparse forms”. In: *Journal of the London Mathematical Society* 98.2 (2018), pp. 369–392.
- [CF74] Ronald Coifman and Charles Fefferman. “Weighted norm inequalities for maximal functions and singular integrals”. In: *Studia Mathematica* 51 (1974), pp. 241–250.
- [CF75] A. Córdoba and Robert Fefferman. “A geometric proof of the strong maximal theorem”. In: *Annals of Mathematics* 102.1 (1975), pp. 95–100.
- [Che+18] Wei Chen et al. “Endpoint sparse bounds for Walsh-Fourier multipliers of Marcinkiewicz type”. In: *arXiv preprint arXiv:1805.06060* (2018).
- [CMP12] David Cruz-Urbe, José María Martell, and Carlos Pérez. “Sharp weighted estimates for classical operators”. In: *Advances in Mathematics* 229.1 (2012), pp. 408–441.
- [Con+17] José M. Conde-Alonso et al. “A sparse domination principle for rough singular integrals”. In: *Analysis & PDE* 10.5 (2017), pp. 1255–1284.
- [Con+22] José M. Conde-Alonso et al. “A metric approach to sparse domination”. In: *Annali di Matematica Pura ed Applicata (1923-)* 201.4 (2022), pp. 1–37.
- [Con+24] José M. Conde-Alonso et al. “Endpoint estimates and sparse domination in nonhomogeneous trees”. In: *arXiv preprint arXiv:2410.23047* (2024).
- [CR16] José M. Conde-Alonso and Guillermo Rey. “A pointwise estimate for positive dyadic shifts and some applications”. In: *Mathematische Annalen* 365 (2016), pp. 1111–1135.
- [CS24] Xi Cen and Zichen Song. “The multilinear fractional sparse operator theory I: pointwise domination and weighted estimate”. In: *arXiv preprint arXiv:2412.21121* (2024).
- [DLR16] Carlos Domingo-Salazar, Michael T. Lacey, and Guillermo Rey. “Borderline weak-type estimates for singular integrals and square functions”. In: *Bulletin of the London Mathematical Society* 48.1 (2016), pp. 63–73.
- [Dua22] Rodrigo Duarte. “Controllo por Operadores Esparsos”. In: *Private Lecture Notes* (2022).
- [Duo01] Javier Duoandikoetxea. *Fourier Analysis*. Vol. 29. Graduate Studies in Mathematics. Translated and revised from the 1995 Spanish original by David Cruz-Urbe. American Mathematical Society, Providence, RI, 2001, pp. xviii+222. ISBN: 0-8218-2172-5. DOI: 10.1090/gsm/029. URL: <https://doi.org/10.1090/gsm/029>.
- [FLW23] Guillermo Flores, Ji Li, and Lesley Ward. “Zygmund type and flag type maximal functions, and sparse operators”. In: *Proceedings of the American Mathematical Society* 151.02 (2023), pp. 555–567.

- [FN19] Dorothee Frey and Zoe Nieraeth. “Weak and Strong Type $A_1 - A_\infty$ Estimates for Sparsely Dominated Operators”. In: *The Journal of Geometric Analysis* 29 (2019), pp. 247–282.
- [Gra09] L. Grafakos. *Modern Fourier Analysis*. Graduate Texts in Mathematics. Springer New York, 2009. ISBN: 9780387094342. URL: <https://books.google.pt/books?id=B-zebZafcLUC>.
- [Gra14] L. Grafakos. *Classical Fourier Analysis*. Graduate Texts in Mathematics. Springer New York, 2014. ISBN: 9781493911943. URL: <https://books.google.pt/books?id=94FxBQAAQBAJ>.
- [Hän18] Timo S. Hänninen. “Equivalence of sparse and Carleson coefficients for general sets”. In: *Arkiv för Matematik* 56.2 (2018), pp. 333–339. DOI: 10.4310/ARKIV.2018.v56.n2.a8. URL: <https://doi.org/10.4310/ARKIV.2018.v56.n2.a8>.
- [HMW73] Richard Hunt, Benjamin Muckenhoupt, and Richard Wheeden. “Weighted norm inequalities for the conjugate function and Hilbert transform”. In: *Transactions of the American Mathematical Society* 176 (1973), pp. 227–251.
- [HP15] Tuomas Hytönen and Carlos Pérez. “The $L(\log L)^\epsilon$ endpoint estimate for maximal singular integral operators”. In: *Journal of Mathematical Analysis and Applications* 428.1 (2015), pp. 605–626.
- [Hu20] Guoen Hu. “Weighted weak type endpoint estimates for the compositions of Calderón–Zygmund operators”. In: *Journal of the Australian Mathematical Society* 109.3 (2020), pp. 320–339.
- [Hyt12] Tuomas P. Hytönen. “The sharp weighted bound for general Calderón–Zygmund operators”. In: *Annals of mathematics* (2012), pp. 1473–1506.
- [Hyt14] Tuomas P. Hytönen. “The A_2 theorem: remarks and complements”. In: *Harmonic analysis and partial differential equations* 612 (2014), pp. 91–106.
- [JMZ35] Børge Jessen, Jan Marcinkiewicz, and Antoni Zygmund. “Note on the differentiability of multiple integrals”. In: *Fundamenta Mathematicae* 25.1 (1935), pp. 217–234.
- [KL18] Robert Kesler and Michael T. Lacey. “Sparse endpoint estimates for Bochner–Riesz multipliers on the plane”. In: *Collectanea Mathematica* 69 (2018), pp. 427–435.
- [KN24] Spyridon Kakaroumpas and Zoe Nieraeth. “Multilinear matrix weights”. In: *arXiv preprint arXiv:2412.15026* (2024).
- [Lac17] Michael T. Lacey. “An elementary proof of the A_2 bound”. In: *Israel Journal of Mathematics* 217 (2017), pp. 181–195.
- [Lac19] Michael T. Lacey. “Sparse bounds for spherical maximal functions”. In: *Journal d’Analyse Mathématique* 139.2 (2019), pp. 613–635.
- [Ler13a] Andrei K. Lerner. “A simple proof of the A_2 conjecture”. In: *International Mathematics Research Notices* 2013.14 (2013), pp. 3159–3170.
- [Ler13b] Andrei K. Lerner. “On an estimate of Calderón–Zygmund operators by dyadic positive operators”. In: *Journal d’Analyse Mathématique* 121.1 (2013), pp. 141–161.

- [Ler16] Andrei K. Lerner. “On pointwise estimates involving sparse operators”. In: *New York J. Math.* 22 (2016), pp. 341–349. ISSN: 1076-9803. URL: http://nyjm.albany.edu:8000/j/2016/22_341.html.
- [Li17] Kangwei Li. “Two weight inequalities for bilinear forms”. In: *Collectanea mathematica* 68 (2017), pp. 129–144.
- [Li18] Kangwei Li. “Sparse domination theorem for multilinear singular integral operators with Lr-Hörmander condition”. In: *Michigan Mathematical Journal* 67.2 (2018), pp. 253–265.
- [LM17] Michael T. Lacey and Darío Mena Arias. “The sparse T1 theorem”. In: *Houston J. Math.* 43.1 (2017), pp. 111–127. ISSN: 0362-1588.
- [LMR19] Michael T. Lacey, Dario Mena, and Maria Carmen Reguera. “Sparse bounds for Bochner–Riesz multipliers”. In: *Journal of Fourier Analysis and Applications* 25 (2019), pp. 523–537.
- [LN19] Andrei K. Lerner and Fedor Nazarov. “Intuitive dyadic calculus: the basics”. In: *Expositiones Mathematicae* 37.3 (2019), pp. 225–265.
- [LN22] Emiel Lorist and Zoe Nieraeth. “Sparse domination implies vector-valued sparse domination”. In: *Mathematische Zeitschrift* 301.1 (2022), pp. 1–35.
- [LO20] Andrei K. Lerner and Sheldy Ombrosi. “Some remarks on the pointwise sparse domination”. In: *The Journal of Geometric Analysis* 30.1 (2020), pp. 1011–1027.
- [LS17] Michael T. Lacey and Scott Spencer. “Sparse bounds for oscillatory and random singular integrals”. In: *New York J. Math.* 23 (2017), pp. 119–131. ISSN: 1076-9803. URL: http://nyjm.albany.edu:8000/j/2017/23_119.html.
- [LTH20] Jiacheng Lan, Xiangxing Tao, and Guoen Hu. “Weak type endpoint estimates for the commutators of rough singular integral operators”. In: *Math. Inequal. Appl.* 23.4 (2020), pp. 1179–1195. ISSN: 1331-4343, 1848-9966. DOI: 10.7153/mia-2020-23-91. URL: <https://doi.org/10.7153/mia-2020-23-91>.
- [Moe12] Kabe Moen. “Sharp weighted bounds without testing or extrapolation”. In: *Archiv der Mathematik* 99.5 (2012), pp. 457–466.
- [Muc72] Benjamin Muckenhoupt. “Weighted norm inequalities for the Hardy maximal function”. In: *Transactions of the American Mathematical Society* 165 (1972), pp. 207–226.
- [NS24] Zoe Nieraeth and Cody B. Stockdale. “Endpoint weak-type bounds beyond Calderón-Zygmund theory”. In: *arXiv preprint arXiv:2409.08921* (2024).
- [Obe19] Richard Oberlin. “Sparse bounds for a prototypical singular Radon transform”. In: *Canadian Mathematical Bulletin* 62.2 (2019), pp. 405–415.
- [Per19] María Cristina Pereyra. “Dyadic harmonic analysis and weighted inequalities: the sparse revolution”. In: *New Trends in Applied Harmonic Analysis, Volume 2: Harmonic Analysis, Geometric Measure Theory, and Applications*. Springer, 2019, pp. 159–239.

- [Pet07] Stefanie Petermichl. “The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical characteristic”. In: *American journal of mathematics* (2007), pp. 1355–1375.
- [Ros62] Marvin Rosenblum. “Summability of Fourier series in $L_p(d\mu)$ ”. In: *Transactions of the American Mathematical Society* 105.1 (1962), pp. 32–42.
- [Rub84] JL Rubio de Francia. “Factorization theory and Ap weights”. In: *Amer. J. Math* 106.3 (1984), pp. 533–547.
- [SM93] Elias M. Stein and Timothy S. Murphy. *Harmonic Analysis (PMS-43): Real-Variable Methods, Orthogonality, and Oscillatory Integrals. (PMS-43)*. Princeton University Press, 1993. ISBN: 9780691032160. (Visited on 06/25/2025).