SHARP ESTIMATES IN HARMONIC ANALYSIS

By

Guillermo Rey

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Mathematics – Doctor of Philosophy

2015

ABSTRACT

SHARP ESTIMATES IN HARMONIC ANALYSIS

By

Guillermo Rey

We investigate certain sharp estimates related to singular integrals. In particular we give sharp level set estimates for sparse operators, we show how to reduce the problem of estimating Calderón-Zygmund operators by sparse operators, and we study some weighted inequalities for these operators.

ACKNOWLEDGEMENTS

Todo esto, no solo esta tesis, habría sido imposible de no ser por mi familia. Gracias, mamá, por enseñarme a ser feliz, y por aguantarme todos estos años; todavía me acuerdo cuando entedí las matemáticas por primera vez cuando estábamos en Elkridge. Gracias, papá, por todas esas discusiones sobre filosofía, ciencia, informática, y de todo en general; por tener siempe una respuesta a mis preguntas, y por dejarme jugar con todos los aparatos que encontraba. Gracias a los dos por estar siempre ahí. Y gracias a Marta, por ayudarme en todas las grandes decisiones que he tomado; contigo he aprendido muchísimo, y eso que estos dos te llevan mas de veinte años de ventaja. Gracias, Jaime, por hacerme reir tantísimas veces: así es imposible no tener buenas memorias. Y, Pablo, gracias por quererme tanto; eres de las primeras personas con las que quiero estar cada vez que hago una visita. Y por supuesto, gracias al resto de mi familia, que es demasiado grande como para enumerar aquí. Que sepáis que os pongo como ejemplo de la suerte que tengo.

A Micha y Antonio por esos interrailes, y a Luis Ángel. Por ser de los pocos buenos recuerdos que tengo de mi educación antes de ir a la universidad. A Jorge porque contigo siempre se puede hablar de lo que más nos gusta. A Rita, por ser una amiga tan chachi y por enseñarme tantas palabras guays. A Carmen por tener conversaciones interminables conmigo sobre cómo arreglar el mundo. A María porque yo no iba a la cafetería contigo por la calidad de los cafés, sino porque así podía estar más tiempo contigo; porque eras, y sigues siendo, absolutamente imprescindible. A Antón porque después de tanto tiempo yo creo que ya pensamos telepáticamente. Por enseñarme que los problemas sin solución no existen, y porque contigo puedo estar hablando hasta el amanecer. Y a Miranda, porque aún me acuerdo de cuando te conocí, y de las agujetas que tuve al día siguiente de tanto reirme.

A Jose, porque te gustan las matemáticas y la comida tanto como a mí. Por visitarme, viajar conmigo, y ayudarme en el proyecto del que estoy más orgulloso. A Beatriz porque siempre te preocupas por mí y porque abrazas fenomenal. A Adrián por enseñarme tantas cosas, y por todas las que aún no me has enseñado. Y a María, porque puedo contártelo todo y siempre estás ahí cuando te necesito. Porque hablar contigo me alegra el día aunque esté muy lejos.

A los profesores de la UAM que me ayudaron tanto desde el principio. Gracias a Antonio Córdoba, José García-Cuerva, y Fernando Soria. Sobre todo, gracias a Ana Vargas por confiar en mí y enseñarme tanto.

To the graduate students at MSU which made my stay here much more tolerable; thank you Alex for teaching me Bellman functions, thank you Tyler and Ben for all those Beggars nights, and thank you Charlotte and Sami for all the laughs in your new office.

And last, but not least. Thank you to my two wonderful advisors: Ignacio Uriarte-Tuero and Alexander Volberg. Thank you for your constant support and for teaching me so much about mathematics and life. This, obviously, would not have happened without your help.

Guillermo Rey

TABLE OF CONTENTS

LIST OF FIGURES

1

Introduction

In analysis, one often needs to commute a limit with an operator. A famous example of such situation concerns the Fourier transform:

$$\widehat{f}(\xi) := \int_{-\infty}^{\infty} e^{-2\pi i x\xi} f(x)\, dx.$$

We would like to recover the function f from its Fourier transform $\widehat{f}$, and this is certainly possible in some cases, but one should be careful with what exactly do we mean by "recover". In particular, even though the Fourier transform may be well defined, the inverse Fourier transform may not.

A typical way to resolve this issue is to not invert the whole function $\widehat{f}$, but a truncation of it:

$$f_R(x) := \int_{-R}^{R} e^{2\pi i x\xi} \widehat{f}(\xi)\, d\xi.$$

Now the question is: Does f_R tend to f as $R \to \infty$? In what sense?

This is a very old question and there are many ways to answer it. Perhaps, the best known answer is the following: if $f \in L^2(\mathbb{R})$ then $f_R \to f$ in $L^2(\mathbb{R})$, that is

$$\lim_{R\to\infty} \|f_R - f\|_{L^2(\mathbb{R})} = 0. \tag{1.1}$$

The main tool used to prove this result is Plancherel's theorem:

$$\|\widehat{f}\|_{L^2(\mathbb{R})} = \|f\|_{L^2(\mathbb{R})}.$$

If we take this theorem for granted, then we can give a short proof of (1.1):

$$\begin{aligned}\lim_{R\to\infty} \|f_R - f\|_{L^2}^2 &= \lim_{R\to\infty} \|\widehat{f_R} - \widehat{f}\|_{L^2}^2 \\ &= \lim_{R\to\infty} \int_{\mathbb{R}} |f(x)|^2 (1 - \mathbb{1}_{[-R,R]}(x))\, dx, \\ &= 0.\end{aligned}$$

where we have used the Dominated Convergence Theorem in the second to last line.

One could ask what is so special about L^2, apart from Plancherel's theorem. Is this true in, say, L^3?

This is indeed true for $f \in L^p(\mathbb{R})$ for all $1 < p < \infty$, but the proof is considerably more involved. One way to prove it is to study the Fourier multiplier operator

$$\widehat{Hf}(\xi) = \operatorname{sign}(\xi)\widehat{f}(\xi).$$

If one could show that this operator is bounded in L^p, then one would immediately get

$$\|f_R - f\|_{L^p} \to 0$$

since we can reconstruct the multiplier $\mathbb{1}_{[-R,R]}$ by translated, modulated, and dilated versions of the multiplier sign.

So, the problem is reduced to studying the boundedness of H, which is also known as the Hilbert transform (up to a multiplicative constant), and which can also be written as

$$Hf(x) = c\,\text{p.v.} \int_{\mathbb{R}} \frac{f(y)}{x-y}\, dy,$$

for some constant $c \neq 0$.

Hence, we have reduced the problem of studying the convergence properties of the Fourier transform to studying a certain *singular integral* operator. And moreover, quantitative bounds for the singular integral yield quantitative information about the original problem.

Singular integrals can be further generalized in many different directions, and have numerous applications. There is a particularly common class, called Calderón-Zygmund operators, which will be the main theme in this work. Their rigorous definition will be postponed until Chapter 3, but essentially they are operators of the form

$$Tf(x) = \text{p.v.} \int_{\mathbb{R}^d} K(x,y) f(y)\, dy,$$

where the kernel K satisfies certain growth conditions like

$$|K(x,y)| \lesssim \frac{1}{|x-y|^d},$$

and certain regularity conditions like

$$|\nabla K(x,y)| \lesssim \frac{1}{|x-y|^{d+1}}.$$

These operators generalize the Hilbert transform to higher dimensions, where more applications appear. For example, three-dimensional singular integrals are of special interest in fluid dynamics.

In Chapter 2 we study the weak-type boundedness of a certain dyadic model of Calderón-Zygmund operators called dyadic shifts. In this Chapter we obtain the best possible upper-level set estimates for such operators using the Bellman function technique.

In Chapter 3 we show how to reduce the problem of obtaining estimates for Calderón-Zygmund operators (as well as other similar operators) to studying the simple dyadic shifts introduced in the previous Chapter. In Chapter 4 we study the embedding of A_1 weights into A_∞; these are classes of absolutely continuous measures which play a key role in the study of weighted inequalities for singular integrals. In Chapter 5 we give short proofs of two weak-type bounds (for singular integrals and for square functions) using the machinery introduced in the previous chapters.

2

Sharp weak-type bounds for positive dyadic shifts

Guillermo Rey and Alexander Reznikov

Advances in Mathematics, Vol 254, March 2014.

2.1 Introduction

The purpose of this article is to study the weak-type $(1,1)$ boundedness of the operator

$$\mathcal{A}f = \sum_{Q\in\mathcal{D}(I)} \alpha_Q \langle f\rangle_Q \mathbb{1}_Q.$$

Here I denotes any finite interval in $\mathbb{R}$, $\langle f\rangle_I = \frac{1}{|I|}\int_I f$, $\mathcal{D}(I)$ denotes the dyadic grid consisting of dyadic subintervals of I and $\{\alpha_J\}_{J\in\mathcal{D}(I)}$ is a Carleson sequence adapted to I, i.e.: $\alpha_J \geq 0$ for all $J \in \mathcal{D}(I)$ and

$$\sup_{J\in\mathcal{D}(I)} \frac{1}{|J|} \sum_{K\in\mathcal{D}(J)} \alpha_K |K| = C < \infty.$$

These operators have recently appeared in the works of A. K. Lerner [20] and [21], where

α_K was a binary sequence, although the ideas go back to [9]. Hence, we will call them *Lerner operators* in the sequel. Here we find the exact Bellman function describing the local boundedness of $\mathcal{A}$ from L^1 to $L^{1,\infty}$.

It is easy to see that the operator $\mathcal{A}$ is bounded in L^2. This, together with a decomposition of Calderón-Zygmund type, can be used to prove an estimate of the form

$$\Big|\Big\{x \in I : \Big|\sum_{J\in\mathcal{D}(I)} \alpha_J \langle f\rangle_J \mathbb{1}_J(x)\Big| > \lambda\Big\}\Big| \le \frac{C}{\lambda}\int_I |f|.$$

However, here we precisely describe how the best constant in the above inequality changes with respect to the parameters of the problem.

The main result of the article is the following theorem:

Theorem 2.1. *Let A, λ and t be positive numbers and I an interval in $\mathbb{R}$, then*

$$\sup \frac{1}{|I|}\Big|\Big\{x \in I : \sum_{J\in\mathcal{D}(I)} \alpha_J \langle f\rangle_J \mathbb{1}_J(x) > \lambda\Big\}\Big| = \begin{cases} \frac{2At}{A\lambda+t} & \text{if } 0 \le t \le A\lambda \le \lambda, \\ \sqrt{\frac{At}{\lambda}} & \text{if } 0 \le A \le \min\Big(\frac{t}{\lambda}, \frac{\lambda}{t}\Big), \\ 1 & \text{otherwise.} \end{cases}$$

Where the supremum is taken over all nonnegative functions f with $\langle f\rangle_I = t$ and all nonnegative sequences $\{\alpha_J\}_{J\in\mathcal{D}(I)}$ with Carleson constant at most 1 which satisfy

$$\frac{1}{|I|}\sum_{J\in\mathcal{D}(I)} \alpha_J |J| = A.$$

We also provide a sequence of examples which, in the limit, attain the supremum of the previous result. See the last section for details on the structure of such examples.

As an immediate corollary we have the following local weak-type (1,1) estimate:

Corollary 2.2. *For any nonnegative $f \in L^1([0,1))$ and for any Carleson sequence $\{\alpha_J\}_{J\in\mathcal{D}([0,1))}$*

with constant at most 1 *we have the sharp bound*

$$\left|\left\{x \in [0,1) : \mathcal{A}f(x) > \lambda\right\}\right| \leq \begin{cases} \frac{2\|f\|_{L^1}}{\lambda+\|f\|_{L^1}} & \text{if } \|f\|_{L^1} \leq \lambda \\ 1 & \text{if } \|f\|_{L^1} \geq \lambda, \end{cases}$$

which in particular implies that

$$\|\mathcal{A}f\|_{L^{1,\infty}([0,1))} \leq 2\|f\|_{L^1([0,1))},$$

and that the constant 2 *is sharp.*

Operators similar to these were recently studied in [29], [33], [31] and [32], however their results are slightly different from ours. They consider the supremum taken over all functions f satisfying

$$\int_I f = s \quad \text{and} \quad \int_I G(f) = t,$$

where G is a strictly convex function satisfying $G(x)/x \to \infty$ as $x \to \infty$. This does not include the question of boundedness from L^1 to $L^{1,\infty}$. Our method of proof is different than the one used in the articles cited above, where they use the deep combinatorial properties of these operators. See also the monograph [40] by A. Osękowski for related results. We instead follow the ideas in [45] and [46] to solve the Bellman PDE and prove its sharpness. This problem is also closely related to studying Haar shifts, the main difference being that Haar shifts are not positive operators. It has been shown however, see [5], that Lerner-type operators can be used to bound Haar shifts. The reader can find results similar to ours in [44], [34] and [38].

The article is organized as follows. In Section 2 we explain how the Bellman function technique is used to compute the supremum in Theorem 2.1. In Section 3 we give a supersolution to the Bellman variational problem which serves as an upper bound for the exact Bellman function. Finally, in Section 4 we show that the function we found in the previous section is the exact Bellman function, we also give a sequence of examples which, in the limit, extremize the inequality of Theorem 2.1.

2.2 The Bellman function technique

Consider the function defined in $\Omega = \{(t, A, \lambda) : 0 \leq t, 0 \leq A \leq 1, \lambda \in \mathbb{R}\}$

$$\mathbb{B}(t, A, \lambda) = \sup\Big\{\frac{1}{|I|}\Big|\Big\{x \in I : \sum_{J\in\mathcal{D}(I)} \alpha_J \langle f\rangle_J \mathbb{1}_J > \lambda\Big\}\Big|\Big\},$$

where the supremum is taken over all all nonnegative functions f on I with $\langle f\rangle_I = t$ and all Carleson sequences $\{\alpha_J\}_{j\in\mathcal{D}(I)}$ with constant at most 1 and

$$A = \frac{1}{|I|}\sum_{J\subseteq I} \alpha_J |J|.$$

Note that I is *not* a parameter in $\mathbb{B}$, this is because the supremum is invariant under dilations and translations in I, and hence independent of I.

The Bellman function technique, which first appeared in the 1995 preprint version of [36], is based on showing that $\mathbb{B}$ solves a certain minimization problem. One first shows that $\mathbb{B}$ satisfies a kind of concavity property and explicitly computes $\mathbb{B}$ in a subdomain natural to the problem (this is usually easy). Then one shows that any continuous positive function satisfying these conditions majorizes $\mathbb{B}$, which reduces the problem to finding the smallest function which satisfies these properties. Finally one has to actually find such a function, this is usually the hardest part. The reader can find insightful introductions in [37] and [39], see also [36], [45], and [46] for more examples of this technique.

Let us begin by describing more precisely the concavity property which $\mathbb{B}$ satisfies:

Lemma 2.3 (Main inequality)**.**

$$\mathbb{B}(t,A,\lambda) \geq \frac{1}{2}\Big(\mathbb{B}(t_1,A_1,\lambda') + \mathbb{B}(t_2,A_2,\lambda')\Big) \tag{2.1}$$

whenever

$$t = \frac{t_1+t_2}{2}, \quad A = \frac{A_1+A_2}{2} + \alpha \quad \textit{and} \quad \lambda = \lambda' + \alpha t$$

and $\alpha \geq 0$.

Proof. Consider any dyadic interval I, any function $f \geq 0$ satisfying

$$\langle f\rangle_{I_-} = t_1 \quad \text{and} \quad \langle f\rangle_{I_+} = t_2$$

and any Carleson sequence $\{\alpha_J\}_{J\in\mathcal{D}(I)}$ with constant at most 1 on I satisfying

$$\frac{1}{|I_-|}\sum_{J\in\mathcal{D}(I_-)} \alpha_J|J| = A_1, \quad \frac{1}{|I_-|}\sum_{J\in\mathcal{D}(I_+)} \alpha_J|J| = A_2 \quad \text{and} \quad \alpha_I = \alpha.$$

Suppose also that $\lambda = \lambda' + \alpha t$.

Since $\langle f\rangle_I = t$ then we must have

$$\mathbb{B}(t,A,\lambda) \geq \frac{1}{|I|}\Big|\Big\{x\in I : \sum_{J\in\mathcal{D}(I)} \alpha_J\langle f\rangle_J \mathbb{1}_J(x) > \lambda\Big\}\Big|$$

since the supremum defining $\mathbb{B}$ is taken over a larger space.

Observe now that

$$
\begin{aligned}
\frac{1}{|I|}\Big|\Big\{x \in I : \sum_{J \in \mathcal{D}(I)} \alpha_J \langle f \rangle_J \mathbb{1}_J(x) > \lambda\Big\}\Big| = \\
\frac{1}{2|I_-|}\Big|\Big\{x \in I_- : \sum_{J \in \mathcal{D}(I)} \alpha_J \langle f \rangle_J \mathbb{1}_J(x) > \lambda\Big\}\Big| + \\
\frac{1}{2|I_+|}\Big|\Big\{x \in I_+ : \sum_{J \in \mathcal{D}(I)} \alpha_J \langle f \rangle_J \mathbb{1}_J(x) > \lambda\Big\}\Big| \\
= \frac{1}{2|I_-|}\Big|\Big\{x \in I_- : \sum_{J \in \mathcal{D}(I_-)} \alpha_J \langle f \rangle_J \mathbb{1}_J(x) > \lambda - \alpha_I t\Big\}\Big| + \\
\frac{1}{2|I_+|}\Big|\Big\{x \in I_+ : \sum_{J \in \mathcal{D}(I_+)} \alpha_J \langle f \rangle_J \mathbb{1}_J(x) > \lambda - \alpha_I t\Big\}\Big| \\
= \frac{1}{2|I_-|}\Big|\Big\{x \in I_- : \sum_{J \in \mathcal{D}(I_-)} \alpha_J \langle f \rangle_J \mathbb{1}_J(x) > \lambda'\Big\}\Big| + \\
\frac{1}{2|I_+|}\Big|\Big\{x \in I_+ : \sum_{J \in \mathcal{D}(I_+)} \alpha_J \langle f \rangle_J \mathbb{1}_J(x) > \lambda'\Big\}\Big|
\end{aligned}
$$

and thus the claim follows. □

Also, we trivially see that $\mathbb{B}$ must satisfy the following "obstacle" condition:

$$\mathbb{B}(t, A, \lambda) = 1 \quad \text{whenever } \lambda < 0. \tag{2.2}$$

As we described in the beginning of the section, the function $\mathbb{B}$ is a minimizer in the space of positive functions which satisfy these properties. The following proposition makes this precise:

Proposition 2.4. *Suppose a continuous function F satisfies inequality* (2.1) *together with the obstacle condition* (2.2)*, then we must have*

$$\mathbb{B}(t, A, \lambda) \leq F(t, A, \lambda).$$

Proof. Let $f \geq 0$ be an integrable function on an interval I and let $\{\alpha_J\}_{J \in \mathcal{D}(I)}$ be a Carleson sequence with constant at most 1, then for all fixed λ we have (by (2.1))

$$
\begin{aligned}
F(\langle f\rangle_I, A, \lambda) &= F\left(\frac{\langle f\rangle_{I_-}+\langle f\rangle_{I_+}}{2}, \frac{A_-+A_+}{2}+\alpha_I, \lambda\right) \\
&\geq \frac{1}{2}\Big(F(\langle f\rangle_{I_-}, A_-, \lambda-\alpha_I\langle f\rangle_I)+F(\langle f\rangle_{I_+}, A_+, \lambda-\alpha_I\langle f\rangle_I)\Big),
\end{aligned}
$$

where $A = \frac{1}{|I|}\sum_{J\subseteq I}\alpha_J|J|$ and $A_\pm$ is defined analogously for I_- and I_+.

If we iterate this inequality we obtain

$$
F(\langle f\rangle_I, A, \lambda) \geq \frac{1}{2^N}\sum_{J\subset I,\, |J|=2^{-N}|I|} F(\langle f\rangle_J, A_J, \lambda - \sum_{k=1}^{N}\alpha_{J^{(k)}}\langle f\rangle_{J^{(k)}}\mathbb{1}_{J^{(k)}}(c_J)),
$$

where $A_J = \frac{1}{|J|}\sum_{P\subseteq J}\alpha_P|P|$.

If we assume a priori that the Carleson sequence α is finite then we can let $N\to\infty$ and obtain

$$
\begin{aligned}
F(\langle f\rangle_I, A, \lambda) &\geq \frac{1}{|I|}\int_I F(f(x), A(x), \lambda-\mathcal{A}f(x))\,dx \\
&\geq \frac{1}{|I|}\int_{\{x\in I:\lambda-\mathcal{A}f(x)<0\}} 1\,dx \qquad\qquad \text{by (2.2)} \\
&= \frac{1}{|I|}|\{x\in I : \mathcal{A}f(x)>\lambda\}|.
\end{aligned}
$$

Here $A(x)$ is almost everywhere-defined as the limit of $A(J)$ as $J\to x$, this is easily seen to exist almost everywhere by the Lebesgue differentiation theorem.

Letting the number of non-zero elements of $\{\alpha_J\}_{J\in\mathcal{D}(I)}$ tend to infinity and then taking the supremum in the definition of $\mathbb{B}$ we obtain

$$
F(\langle f\rangle_I, A, \lambda) \geq \mathbb{B}(\langle f\rangle_I, A, \lambda).
$$

□

Remark 2.5. Note that we don't know yet if the function $\mathbb{B}$ is continuous, thus finding a minimizer in the space of continuous functions might not give us the true Bellman function. It turns out, however, that assuming continuity (actually C^1 smoothness) we are able to find a positive function satisfying (2.1) and (2.2) which moreover is best possible without the a priori assumption of smoothness. We show this in the last section.

We have therefore seen that finding any positive continuous function F satisfying (2.1) and (2.2) will give us an upper bound for $\mathbb{B}$. In the next section we find such a function.

2.3 Finding the Bellman function candidate

Our goal now is to find the smallest continuous function F satisfying (2.1) and (2.2). As we remarked after Proposition 2.4, we will assume a priori that F is C^1. Moreover, we will restrict the minimization space even more by requiring F to have the same kind of homogeneity that the true $\mathbb{B}$ must have, i.e.:

$$\mathbb{B}(\eta t, A, \eta\lambda) = \mathbb{B}(t, A, \lambda) \quad \forall \eta > 0, \lambda > 0.$$

This in principle might make our candidate for Bellman function larger than the one we could find without requiring such homogeneity. However, the optimal Bellman function satisfies this identity, so requiring F to also satisfy it will not prevent us from finding it. Assuming smoothness we can write the Main Inequality (2.1) as a concavity condition, together with a monotonicity property along certain characteristics. More precisely, if F is a smooth positive function, then (2.1) together with (2.2) and the above homogeneity is equivalent to the following conditions:

1. F is nonnegative, and concave in the first two variables.

2. $F(t, A, \lambda)$ is increasing in the direction $(0, 1, t)$.

3. $F(st, A, s\lambda) = F(t, A, \lambda)$ for all $s > 0$.

4. $F(t, A, \lambda) = 1$ whenever $\lambda < 0$

Indeed, if we let $\alpha = 0$ in (2.1) we see that $\mathbb{B}$ is concave in the variables (t, A). If we set $A_1 = A_2 = A$ and $t_1 = t_2 = t$ then we see, by varying α, that $\mathbb{B}(t, A, \lambda)$ is increasing in the direction $(0, 1, t)$. This shows that any smooth F satisfying (2.1) and (2.2), and which is also homogeneous in the above sense, must also satisfy properties (1) through (4). Moreover, if F is any smooth function satisfying properties (1) through (4), then it also must satisfy the main inequality (2.1) and the obstacle condition (2.2). To see this observe that using property (1) we obtain (2.1) but with $\alpha = 0$, now property (2) allows us to insert an α as in the hypotheses for the main inequality since it describes the path along which F is increasing. The homogeneity and obstacle conditions are exactly (3) and (4) respectively, so this proves the equivalence.

Using the homogeneity property, we can reduce to finding $M : (0, \infty) \times [0, 1] \to [0, \infty)$ such that if

$$F(x, y, z) = \begin{cases} M(x/z, y) & \text{if } z > 0 \\ 1 & \text{if } z \leq 0, \end{cases}$$

then F satisfies (1) through (4). These properties, when translated to the function M, become:

1. M is concave.

2. $M_y - x^2 M_x \geq 0$.

3. $M(x, y) \to 1$ when $x \to \infty$.

The second of these properties tells us that M is increasing along the characteristics

$$\begin{cases} \dot{x}(t) & = -x^2 \\ \dot{y}(t) & = 1. \end{cases}$$

Observe that these characteristics foliate $[0, \infty) \times [0, 1]$. Also, if we move backwards in time along a characteristic which starts at $(x_0, 1)$ with $x_0 \geq 1$, then this characteristic is above the curve $y = \frac{1}{x}$ and furthermore the characteristic tends to (∞, y_f) for some $0 < y_f < 1$. Using the fact that $M(x, y) \to 1$ as $x \to \infty$ and that we should decrease if we move backwards along these characteristics, we must have

$$M(x, y) \geq 1 \quad \text{whenever } y \geq \frac{1}{x}.$$

However, we may assume (if our goal is to find the true Bellman function) that $M \leq 1$ since the true Bellman function $\mathbb{B}$ obviously cannot be larger than 1, so we will actually impose

$$M(x, y) = 1 \quad \text{whenever } y \geq \frac{1}{x}.$$

Observe that $\mathbb{B}(0, 0, 1)$ is 0 and consider the straight line joining the point $(0, 0)$ with (x_1, y_1), where $x_1 y_1 = 1$. Observe also that the pointwise minimum of any two positive continuous functions satisfying (2.1) and (2.2) will give us a smaller function which also satisfies these properties.

We know that the function M should be 1 at (x_1, y_1) and that, along this line, M should be concave. The smallest concave curve joining these two points is obviously a straight line, so if defining M in this way produces a smooth concave function satisfying the monotonicity property (2) then the optimal M should be such a function. Joining the point $(0, 0)$ with the points $(x_1, y_1) \in [0, \infty) \times [0, 1]$ satisfying $x_1 y_1 = 1$ covers everything in the subdomain $0 \leq y \leq \min(x, x^{-1})$, so let us define M here by

$$M(x, y) = \sqrt{xy}.$$

This function is linear along straight lines joining $(0, 0)$ with the boundary curve $xy = 1$ and is 1 at this boundary. It is furthermore concave and satisfies the monotonicity property (2), so if we knew that $\mathbb{B}$ is continuous then $\mathbb{B}(x, y, 1)$ must be defined as above in this subdomain.

We are therefore left with defining M in the upper triangle $\Omega_T = \{0 \leq x \leq y \leq 1\}$. Inspired by the linear behavior of M in the first domain, we make the ansatz that M is actually 1-homogeneous in the whole domain.

Let $f(x) = M(x, 1)$ for $0 \leq x \leq 1$, then if M is 1-homogeneous we should have

$$M(x, y) = yf(x/y).$$

If we want condition (2) to hold then we should have

$$f(x/y) - (x/y)f'(x/y) - x^2 f'(x/y) \geq 0.$$

We expect this to be an equality on the boundary, which is when $y = 1$, so we will assume that

$$f(x) - f'(x)(x + x^2) = 0.$$

This ordinary differential equation has the solutions

$$f(x) = C\frac{x}{1+x},$$

and we should furthermore have $f(1) = M(1, 1) = 1$. So $C = 2$ and therefore

$$f(x) = \frac{2x}{x+1} \implies M(x, y) = \frac{2xy}{x+y}$$

whenever $1 \geq y \geq x \geq 0$. One easily verifies that M satisfies all the requirements in this subdomain, so we just have to show that the whole function M is concave, but this immediately follows from the fact that M is concave in each subdomain and that M is C^1 (as can be easily seen).

This gives us that

$$M(x,y) = \begin{cases} \frac{2xy}{x+y} & \text{if } 0 \le x \le y \le 1 \\ \sqrt{xy} & \text{if } 0 \le y \le \min(x, x^{-1}), \end{cases}$$

which using the homogeneity gives us the full function of Theorem 2.1.

2.4 Optimality

In this section we show that the function found in the previous section is actually the exact Bellman function. We first we need a simple technical lemma which will allow us to deduce that $\mathbb{B}(\cdot,\cdot,1)$ must be superlinear along lines joining $(0,0,1)$ to $(x,1,1)$.

Lemma 2.6. *Let* $f : [0,1] \to [0,\infty)$ *be a function which satisfies*

$$f\Big(\frac{x+y}{2}\Big) \ge \frac{1}{2}f(x) + \frac{1}{2}f(y) \tag{2.3}$$

for all $0 \le x \le y \le 1$*. Then we must have*

$$f(x) \ge f(1)x$$

for all $x \in [0,1]$.

Proof. We can assume without loss of generality that $x \in (0,1)$ and that $f(1) = 1$. Using (2.3) we have

$$f(x_0 + \lambda(1 - x_0)) \ge \lambda \tag{2.4}$$

for all dyadic rationals $\lambda \in [0,1]$, i.e.: numbers of the form $\lambda = k2^{-N}$ for $0 \le k \le 2^N$. For every $N \in \mathbb{N}$ let k_N be the unique integer in $0 \le k \le 2^N x$ which satisfies

$$\Big|x - \frac{k}{2^N}\Big| < \frac{1}{2^N}$$

(this exists because the sequence $k \mapsto k2^{-N}$ is an arithmetic sequence of step 2^{-N}).

Observe that then, if we define

$$x_N := \frac{2^N x - k_N}{2^N - k_N} = \frac{x - \frac{k_N}{2^N}}{1 - \frac{k_N}{2^N}},$$

we must have $0 \leq x_N \leq \frac{1}{2^N(1-x)}$, so in particular $x_N \to 0$ as $N \to \infty$.

But then

$$\lambda := \frac{k_N}{2^N} = \frac{x - x_N}{1 - x_N}$$

is a dyadic rational and plugging it into (2.4), with x_N playing the role of x_0, yields

$$f(x) \geq \frac{x - x_N}{1 - x_N},$$

so letting $N \to \infty$ completes the proof. $\square$

Using this lemma, together with the Main Inequality (2.1) we immediately have the following corollary:

Corollary 2.7. *We have the following identity:*

$$\mathbb{B}(x, y, 1) = M(x, y)$$

for all x, y in the subdomain $0 \leq y \leq \min(x, x^{-1})$.

Proof. We showed in the previous section that $\mathbb{B}(x, y, 1) \leq M(x, y)$ for all $(x, y) \in \Omega'$. To show the reverse inequality notice that the Main Inequality (2.1) together with Lemma 2.6 imply

$$\mathbb{B}(x, y, 1) \geq \lambda \mathbb{B}\Big(\frac{x}{\lambda}, \frac{y}{\lambda}, 1\Big). \tag{2.5}$$

We would be done if we can show that $\mathbb{B}(x, y, 1) = 1$ whenever $xy = 1$. Indeed, then we can just use equation (2.5) with $\lambda = \sqrt{xy}$.

Fix $(x, y) \in \Omega'$ with $xy = 1$ and consider the function

$$f_n = \frac{2^n x}{2^n - 1} \mathbb{1}_{[0,1-2^{-n})}.$$

If I is the interval $[0, 1)$ then obviously $\langle f_n \rangle_I = x$. Consider also the Carleson sequence $\{\alpha_J\}_{J \in \mathcal{D}(I)}$ defined by

$$\alpha_J = \begin{cases} \frac{y}{1-2^{-n}} & \text{if } J = [2^{-n}(k-1), 2^{-n}k) \text{ and } k \in \{1, \ldots, 2^n - 1\} \\ 0 & \text{otherwise.} \end{cases}$$

Then we have

$$\frac{1}{|I|} \sum_{J \in \mathcal{D}(I)} \alpha_J |J| = y \sum_{k=1}^{2^n - 1} \frac{2^{-n}}{1 - 2^{-n}} = y.$$

Also,

$$\mathcal{A} f_n(t) = \begin{cases} \left(\frac{2^n}{2^n - 1} \right)^2 & \text{if } 0 \leq t < 1 - 2^{-n} \\ 0 & \text{otherwise,} \end{cases}$$

hence

$$\mathbb{B}(x, y, 1) \geq 1 - 2^{-n}$$

for all $n \geq 1$. Letting $n \to \infty$ yields the claim. $\square$

Remark 2.8. Observe that using the constant function $f(t) = x\mathbb{1}_I(t)$ and the one-term Carleson sequence which is y on I and 0 everywhere else, one obtains that $\mathcal{A}f = xy\mathbb{1}_I$, hence $\mathbb{B}(x, y, 1) = 1$ for all $xy > 1$.

Using Lemma 2.6 in the same way, we just have to show that $\mathbb{B}(x, 1, 1) = \frac{2x}{x+1}$ to prove that $\mathbb{B}(x, y, 1) = M(x, y)$ in the rest of the domain, however this turns out to be harder.

Theorem 2.9. *Fix $x \in (0, 1)$ and let $\epsilon > 0$. For any interval I there exists a nonnegative function f on I with $\langle f \rangle_I = x$ and a Carleson sequence $\{\alpha_J\}_{J \in \mathcal{D}(I)}$ with Carleson constant* *at most one and verifying*

$$\frac{1}{|I|}\sum_{J\in\mathcal{D}(I)}\alpha_J|J| = 1$$

such that

$$\frac{1}{|I|}\Big|I\cap\Big\{\sum_{J\in\mathbb{D}(I)}\alpha_J\langle f\rangle_J\mathbb{1}_J > 1\Big\}\Big| = \frac{2x}{x+1}+O(\epsilon).$$

To prove this we will use the Main Inequality (2.1) iteratively to give a decomposition of f consisting of constant functions on certain dyadic intervals, this also gives us the construction of the sequence $\{\alpha_J\}_{J\in\mathcal{D}(I)}$. The basic idea is to, starting with a point $(x,1)$ in Ω', use (2.1) to split this point into another point $(x_+,1)$ on the boundary and some point (x_-,A_-). The point (x_-,A_-) is then absorbed back into the initial point and we apply the same procedure to the point $(x_+,1)$ until we get to a point past the obstacle $xy\geq 1$ (where extremizers consist of constant functions together with one-term Carleson sequences as in the Remark after Corollary 2.7).

In order to illustrate the idea we will first prove the lower bound for $\mathbb{B}$ without explicitly constructing the example. The way in which we prove the lower bound will make the construction more intuitive.

Theorem 2.10. *The Bellman function $\mathbb{B}$ satisfies*

$$\mathbb{B}(x,1,1) = \frac{2x}{x+1}$$

for all $x\in[0,1]$.

Proof. Let $E(x,y)=\mathbb{B}(x,y,1)$, then using the Main Inequality (2.1) we see that we have the following behavior:

$$E(t,1)\geq\frac{1}{2}E\Big(\frac{t_1}{1-\alpha t},A_1\Big)+\frac{1}{2}E\Big(\frac{t_2}{1-\alpha t},A_2\Big)$$

whenever $t = \frac{t_1+t_2}{2}$ and $1 = \frac{A_1+A_2}{2} + \alpha$. Letting $\epsilon > 0$, $x = t$ and $A_2 = 1$ we get

$$E(x,1) \geq \frac{1}{2}\Big(E\Big(x - 2\epsilon, 1 - \frac{2\epsilon}{x}\Big) + E\Big(x_+, 1\Big)\Big),$$

where

$$x_+ = x\frac{1+\epsilon}{1-\epsilon} + 2\epsilon.$$

Since $\mathbb{B}$ is superlinear in the first two variables and $\mathbb{B}(0,0,1) = 0$, we must have

$$E\Big(x - 2\epsilon, 1 - \frac{2\epsilon}{x}\Big) \geq \Big(1 - \frac{2\epsilon}{x}\Big)E(x,1)$$

so putting everything together we obtain

$$E(x,1) \geq \frac{x}{x+2\epsilon}E(x_+,1). \tag{2.6}$$

If we define inductively $x_{n+1} = x_n\frac{1+\epsilon}{1-\epsilon} + 2\epsilon$ and $x_0 = x$, then we easily see that

$$x_n = \delta^n\Big(\frac{1}{1-\epsilon} + x\Big) - \frac{1}{1-\epsilon},$$

where $\delta = \frac{1+\epsilon}{1-\epsilon}$.

We want to stop the iteration once $x_n \geq 1$, and this happens when

$$\delta^n \geq \frac{2-\epsilon}{1+x(1-\epsilon)},$$

let $N = N(\epsilon, x)$ be the smallest integer for which the above inequality does not hold. Then iterating (2.6) N times we get (since $E(1,1) = 1$)

$$E(x,1) \geq \prod_{j=0}^{N} \frac{x_j}{x_j + 2\epsilon},$$

it just suffices to give a lower bound for the right hand side.

To this end observe that

$$\begin{aligned}\prod_{j=0}^{N}\frac{x_j}{x_j+2\epsilon} &\geq \exp\Big(-\sum_{j=0}^{N}\log\Big(1+\frac{2\epsilon}{x_j}\Big)\Big)\\ &\geq \exp\Big(-\sum_{j=0}^{N}\frac{2\epsilon}{x_j}\Big)\\ &= \exp\Big(-2\epsilon\sum_{j=0}^{N}\frac{1}{x_j}\Big).\end{aligned}$$

Let us estimate $-2\epsilon\sum_{j=0}^{N}\frac{1}{x_j}$. Using the explicit formula for x_n we have

$$\begin{aligned}-2\epsilon\sum_{j=0}^{N}\frac{1}{x_j} &= -2\epsilon\sum_{j=0}^{N}\frac{1}{\delta^j\big(\frac{1}{1-\epsilon}+x\big)-\frac{1}{1-\epsilon}}\\ &= -2\epsilon\sum_{j=0}^{N}\Big(\frac{1}{\delta^j\big(\frac{1}{1-\epsilon}+x\big)-\frac{1}{1-\epsilon}}-\frac{1}{\delta^j\big(1+x\big)-1}\Big)+\sum_{j=0}^{N}\frac{2\epsilon}{1-\delta^j(1+x)}.\end{aligned}$$

The first term tends to 0 as $\epsilon\to 0$ and the second is a Riemann sum, indeed (recalling the definition of $N = N(x,\epsilon)$:

$$\begin{aligned}\sum_{j=0}^{N}\frac{2\epsilon}{1-\delta^j(1+x)} &= (1-\epsilon)\sum_{j=0}^{N}\frac{\delta^j\frac{2\epsilon}{1-\epsilon}}{\delta^j(1-\delta^j(1+x))}\\ &= (1-\epsilon)\sum_{j=0}^{N}f(\delta^j)(\delta^{j+1}-\delta^j)\\ &= \int_1^{\frac{2}{1+x}} f(y)\,dy+O(\epsilon),\end{aligned}$$

as $\epsilon\to 0$ and where

$$f(y)=\frac{1}{y(1-y(x+1))}.$$

It is easy to see that

$$\int_1^{\frac{2}{1+x}}\frac{1}{y(1-y(x+1))}\,dy=\log\Big(\frac{2x}{x+1}\Big),$$

which completes the proof of the lower bound.

□

Let us now use these ideas to construct the example. There are two basic steps in the iteration: first we split the point $(x,1)$ into (x_-,A_-) and $(x_+,1)$, then we absorb (x_-,A_-) into $(x,1)$ and obtain a lower bound for $E(x,1)$ in terms of $E(x_+,1)$, we then iterate this until $x_+>1$, where we stop because we know that $E(x_+,1)$ must be 1 there. These two steps are imposing a certain self-similarity on f and the Carleson sequence α in terms of (f_+,α_+). The following Lemma, which is based on the ideas from [46], makes this precise.

Lemma 2.11. *Fix an interval I and let g_+ be a nonnegative function on I_+. Suppose also that α^+ is a Carleson sequence adapted to I_+ with constant at most 1 and such that*

$$\frac{1}{|I_+|}\sum_{J\in\mathcal{D}(I_+)}\alpha_J^+|J|=1.$$

If $\langle g_+\rangle_{I_+}=x\frac{1+\epsilon}{1-\epsilon}+2\epsilon$ for some $x\in(0,1)$ and a sufficiently small $\epsilon>0$, then we can construct a function f on I and a Carleson sequence α adapted to I with constant at most 1 such that $\langle f\rangle_I=x$,

$$\frac{1}{|I|}\sum_{J\in\mathcal{D}(I)}\alpha_J|J|=1 \tag{2.7}$$

and

$$\frac{1}{|I|}\Big|I\cap\Big\{\sum_{J\in\mathcal{D}(I)}\alpha_J\langle f\rangle_J\mathbb{1}_J>1\Big\}\Big|\geq\Big(\frac{x}{x+2\epsilon}\Big)\frac{1}{|I_+|}\Big|I_+\cap\Big\{\sum_{J\in\mathcal{D}(I_+)}\alpha_J^+\langle g_+\rangle_J\mathbb{1}_J>1\Big\}\Big|. \tag{2.8}$$

Proof. We will assume without loss of generality that $I=[0,1)$, also denote $\alpha=\frac{\epsilon}{x}$. Define α_J to be α if $J=I$ and α_J^+ for $J\in\mathcal{D}(I_+)$.

Define f to be $(1-\epsilon)g_+$ on I_+ and denote $g_- = (1-\epsilon)^{-1} f \mathbb{1}_{I_-}$, then

$$\begin{aligned}
\frac{1}{|I|}\Big|y \in I : \sum_{J\in\mathcal{D}(I)} \alpha_J \langle f\rangle_J \mathbb{1}_J(y) > 1\Big| &= \\
&\frac{1}{2|I_-|}\Big|y \in I_- : \sum_{J\in\mathcal{D}(I_-)} \alpha_J \langle f\rangle_J \mathbb{1}_J(y) > 1-\epsilon\Big| \\
&+ \frac{1}{2|I_+|}\Big|y \in I_+ : \sum_{J\in\mathcal{D}(I_+)} \alpha_J \langle f\rangle_J \mathbb{1}_J(y) > 1-\epsilon\Big| \\
&= \frac{1}{2|I_-|}\Big|y \in I_- : \sum_{J\in\mathcal{D}(I_-)} \alpha_J \langle g_-\rangle_J \mathbb{1}_J(y) > 1\Big| \\
&\qquad + \frac{1}{2|I_+|}\Big|y \in I_+ : \sum_{J\in\mathcal{D}(I_+)} \alpha_J \langle g_+\rangle_J \mathbb{1}_J(y) > 1\Big|.
\end{aligned}$$

Let $I_j = [e_j, e_{j+1})$, where $e_j = \frac{1}{2} - 2^{-j}$, and suppose that $\alpha_{\widehat{I_j}} = 0$ for $j \geq 1$ and $\alpha_{I_-} = 0$, then

$$\frac{1}{2|I_-|}\Big|y \in I_- : \sum_{J\in\mathcal{D}(I_-)} \alpha_J \langle g_-\rangle_J \mathbb{1}_J(y) > 1\Big| = \frac{1}{2}\sum_{j=1}^{\infty} 2^{-j} \frac{1}{|I_j|}\Big|y \in I_j : \mathcal{A}(g_- \mathbb{1}_j)(y) > 1\Big|. \quad (2.9)$$

Let $\theta = 1 - 2\alpha$ and write

$$\theta = \sum_{j=1}^{\infty} 2^{-j} b_j$$

for some binary sequence $\{b_j\}_{j\in\mathbb{N}}$ (i.e.: write θ in binary).

For a given interval J let $S_J f$ be the scaled version of f adapted to J, i.e.: if $J = [a, b)$ then

$$S_J f(x) = f\Big(\frac{x-a}{b-a}\Big).$$

Abusing notation, let us also denote by $S_J\alpha$ the scaled version of the Carleson sequence α to the dyadic subinterval J of I, then we have

$$\frac{1}{|J|}\Big|\Big\{y \in J : \sum_{K\in\mathcal{D}(J)} (S_J\alpha)_K \langle S_J f\rangle_K \mathbb{1}_K(y) > 1\Big\}\Big| = \frac{1}{|I|}\Big|\Big\{y \in I : \sum_{K\in\mathcal{D}(I)} \alpha_K \langle f\rangle_K \mathbb{1}_K(y) > 1\Big\}\Big|.$$

Suppose that $(1-\epsilon)f$, when restricted to I_j, agrees with $S_{I_j} f$ for all $j \geq 1$ such that $b_j = 1$

and is 0 otherwise. Suppose furthermore that the Carleson sequence α also satisfies the same similarity, i.e.: if we scale to I the restriction of α to I_j we obtain α again. If we denote by Ξ the left-hand side in (2.8) then we could use (2.9) to obtain

$$\Xi = \frac{1}{2}\sum_{j=1}^{\infty} 2^{-j} b_j \Xi + \frac{1}{2|I_+|}\left|y \in I_+ : \sum_{J\in\mathcal{D}(I_+)} \alpha_J \langle g_+\rangle_J \mathbb{1}_J(y) > 1\right|,$$

hence

$$\begin{aligned}\Xi &= \left(\frac{1}{1+2\alpha}\right)\frac{1}{|I_+|}\left|y \in I_+ : \sum_{J\in\mathcal{D}(I_+)} \alpha_J \langle g_+\rangle_J \mathbb{1}_J(y) > 1\right| \\ &= \left(\frac{x}{x+2\epsilon}\right)\frac{1}{|I_+|}\left|y \in I_+ : \sum_{J\in\mathcal{D}(I_+)} \alpha_J \langle g_+\rangle_J \mathbb{1}_J(y) > 1\right|,\end{aligned}$$

which is what we wanted. Note also that we could use the same method to compute the average of f and it yields precisely the right amount: x.

Therefore we just have to show that we can find a function f and a Carleson sequence α satisfying these self-similarity conditions. Let us start with f: define the operator T by

$$Tf = (1-\epsilon)\sum_{j=1}^{\infty} b_j \mathbb{1}_{I_j} S_{I_j} f + (1-\epsilon)\mathbb{1}_{I_+} g_+.$$

We need to show that T has a fixed point in $L^1(I)$; we will do this following the steps of the proof of the Banach fixed point theorem. Let $f_0 = (1-\epsilon) g_+ \mathbb{1}_{I_+}$ and define inductively

$$f_{n+1} = Tf_n.$$

We should show that f_n is a Cauchy sequence in $L^1(I)$, but observe that

$$\begin{aligned}
\|f_{n+1} - f_n\|_{L^1(I)} &= (1-\epsilon) \int_{I_-} \Big| \sum_{j=1}^{\infty} b_j \mathbb{1}_{I_j} S_{I_j}(f_n) - \sum_{j=1}^{\infty} b_j \mathbb{1}_{I_j} S_{I_j}(f_{n-1}) \Big| \\
&= (1-\epsilon) \sum_{j=1}^{\infty} b_j \int_{I_j} |S_{I_j}(f_n) - S_{I_j}(f_{n-1})| \\
&= (1-\epsilon) \sum_{j=1}^{\infty} b_j |I_j| \int_I |f_n - f_{n-1}| \\
&= (1-\epsilon) \int_I |f_n - f_{n-1}| \sum_{j=1}^{\infty} b_j 2^{-j-1} \\
&= \frac{(1-\epsilon)(1-2\alpha)}{2} \int_I |f_n - f_{n-1}|.
\end{aligned}$$

The constant $\xi := \frac{(1-\epsilon)(1-2\alpha)}{2}$ is strictly less than 1 and by induction we have

$$\|f_{n+1} - f_n\|_{L^1(I)} \lesssim \xi^n,$$

hence the sequence is Cauchy. This finishes the proof of existence for f since we can just define f to be the limit in L^1 of the sequence f_n defined above.

To show the existence of the Carleson sequence we can follow the same steps as above, but now we don't have to deal with convergence issues. Indeed, start with a sequence as in the beginning of the proof and define inductively the $(n+1)$-th sequence α^{n+1} by inserting the entire dyadic tree of α^n at each I_j. At each step we are only changing the value of the sequence at deeper and deeper levels, so we can just define α_K as the the value of α_K^n, where n is the first integer at which the sequence α_K^n stabilizes. □

We are now ready to prove Theorem 2.9, we will use the same ideas and notation as in the proof of Theorem 2.10. Given ϵ, I and $x \in (0,1)$ let N be the smallest integer such that

$$\delta^n \geq \frac{2-\epsilon}{1 + x(1-\epsilon)}.$$

Let f_1 be the constant function x on I_+ and let α^1 be the one-term Carleson sequence which is 1 at I_+. Now define the function f_{n+1} and Carleson sequence α^{n+1} inductively by applying Lemma 2.11 to the function $g_+ := S_{I_+}(f_n)$ and the Carleson sequence $S_{I_+}(\alpha^n)$; let $f = f_N$ and $\alpha = \alpha^N$. Then, as in the proof of Theorem 2.10, we have

$$\frac{1}{|I|}\Big|\Big\{y \in I : \sum_{J \in \mathcal{D}(I)} \alpha_J \langle f \rangle_J \mathbb{1}_J > 1\Big\}\Big| \geq \exp\Big(\sum_{j=0}^{N} \frac{2\epsilon}{1 - \delta^j(1+x)}\Big),$$

which we showed to be

$$\frac{2x}{x+1} + O(\epsilon),$$

and this is what we wanted to prove.

3

Dyadic models for singular integrals

José M. Conde-Alonso and Guillermo Rey

Mathematische Annalen, October 2015.

3.1 Introduction

One particularly useful way to study many important operators in Harmonic Analysis is that of decomposing them into sums of simpler dyadic operators. An example of a recent striking result using this strategy is the proof of the sharp weighted estimate for the Hilbert transform by S. Petermichl [42]. This was a key step towards the full A_2 theorem for general Calderón-Zygmund operators, finally proven by T. Hytönen in [14]. Of course there are many instances of this useful technique, but we will not try to give a thorough historical overview here.

The proof in [14] was a *tour de force* which was the culmination of many previous partial efforts by others, see [14] and the references therein. Hytönen did not only prove the A_2 theorem, but he also showed that general Calderón-Zygmund operators could be represented as averages of certain simpler "Haar shifts" in the spirit of [42]. The sharp weighted bound then followed from the corresponding one for these simpler operators. Later, A. Lerner gave a simplification of the A_2 theorem in [22] which avoided the use of most of the complicated machinery in [14]; it mainly relied on a general pointwise estimate for functions in terms of positive dyadic operators which had already been proven in [20].

The weighted result for the positive dyadic shifts that this contribution reduced the problem to had already been shown before in [18], see also [4] and [5]. More precisely, the proof of Lerner (essentially) gave the following pointwise estimate for general Calderón-Zygmund operators T: for every dyadic cube Q

$$|Tf(x)| \lesssim \sum_{m=0}^{\infty} 2^{-\delta m} \mathcal{A}_{\mathcal{S}}^m |f|(x) \quad \text{for a.e. } x \in Q, \tag{3.1}$$

where $\delta > 0$ depends on the operator T, $\mathcal{S}$ are collections of dyadic cubes (belonging to same dyadic grid for each fixed $\mathcal{S}$) which depend on f, T and m, and $\mathcal{A}_{\mathcal{S}}^m$ are positive dyadic operators defined by

$$\mathcal{A}_{\mathcal{S}}^m f(x) = \sum_{Q \in \mathcal{S}} \langle f \rangle_{Q^{(m)}} \mathbb{1}_Q(x),$$

where $Q^{(m)}$ denotes the m-th dyadic parent of Q. Moreover, the collections $\mathcal{S}$ in (3.1) are *sparse* in the usual sense: given $0 < \eta < 1$, we say that a collection of cubes $\mathcal{S}$ belonging to the same dyadic grid is η-sparse if for all cubes $Q \in \mathcal{S}$ there exist measurable subsets $E(Q) \subset Q$ with $|E(Q)| \geq \eta |Q|$ and $E(Q) \cap E(Q') = \emptyset$ unless $Q = Q'$. A collection is called simply sparse if it is $\frac{1}{2}$-sparse.

From this pointwise estimate Lerner continues the proof by showing that bounding the operator norm of each $\mathcal{A}_{\mathcal{S}}^m$ can be reduced to just estimating the operator norm of $\mathcal{A}_{\mathcal{S}'}^0$ in the same space for all possible sparse collections $\mathcal{S}'$. More precisely, he shows that

$$\|\mathcal{A}_{\mathcal{S}}^m f\|_{\mathbb{X}} \lesssim (m+1) \sup_{\mathscr{D}, \mathcal{S}'} \|\mathcal{A}_{\mathcal{S}'}^0 f\|_{\mathbb{X}}, \tag{3.2}$$

where the supremum is taken over all dyadic grids $\mathscr{D}$ and all sparse collection $\mathcal{S}' \subset \mathscr{D}$, and where $\mathbb{X}$ is any Banach function space, in the sense of [1], Chapter 1.

It is at this point where the duality of $\mathbb{X}$ is needed in the argument; the operators $\mathcal{A}_{\mathcal{S}}^m$ do not lend themselves to Lerner's pointwise formula, while their adjoints do. Consequently, the question of what to do when no duality is present was left open. Our main result answers this question by proving a stronger (though localized) statement: the operators $\mathcal{A}_{\mathcal{S}}^m$ are actually pointwise bounded by positive dyadic 0-shifts:

Theorem 3.1. *Let P be a cube and $\mathcal{S}$ a sparse collection of dyadic subcubes Q such that $Q^{(m)} \subseteq P$, then for all nonnegative integrable functions f on P there exists another sparse collection $\mathcal{S}'$ of dyadic subcubes of P such that*

$$\mathcal{A}_{\mathcal{S}}^m f(x) \lesssim (m+1)\mathcal{A}_{\mathcal{S}'}^0 f(x) \quad \forall x \in P \tag{3.3}$$

In fact, we prove Theorem 3.1 in a slightly more general setting: first, the statement is proven for a certain natural multilinear generalization of the operators $\mathcal{A}_{\mathcal{S}}^m$. Second, the sparse collection $\mathcal{S}$ is replaced by a more general Carleson sequence. The relevant details are given in the next section.

The novelty in our approach is two-fold: we directly attack the pointwise estimate for the operators $\mathcal{A}^m$, instead of bounding their norm in various spaces. Also, in proving the pointwise bound we develop an algorithm that constructively selects those cubes which will form the family $\mathcal{S}'$. This algorithm has "memory" in a certain sense: each iteration takes into account the previous steps, a feature which is crucial in our method to ensure that $\mathcal{S}$ is sparse.

As a corollary of Theorem 3.1, we find an analogue of (3.1) for Calderón-Zygmund operators with more general moduli of continuity (see the next section for the precise definition). In particular, we obtain the following pointwise estimate for Calderón-Zygmund operators:

Corollary 3.2. *If P is a dyadic cube, f is an integrable function supported on P and T is* *a Calderón-Zygmund operator whose kernel has modulus of continuity ω, then*

$$|Tf(x)| \lesssim \sum_{m=0}^{\infty} \omega(2^{-m})(m+1)\mathcal{A}^0_{\mathcal{S}_m}|f|(x) \quad \text{for a.e. } x \in P, \tag{3.4}$$

where $\mathcal{S}_m$ are sparse collections belonging to at most 3^d different dyadic grids.

Moreover, if we know that ω satisfies the logarithmic Dini condition:

$$\int_0^1 \omega(t)\Big(1+\log\Big(\frac{1}{t}\Big)\Big)\frac{dt}{t} < \infty, \tag{3.5}$$

then we can find sparse collections $\{\mathcal{S}'_1, \ldots, \mathcal{S}'_{3^d}\}$, belonging to possibly different dyadic grids, such that

$$|Tf(x)| \lesssim \sum_{i=1}^{3^d} \mathcal{A}^0_{\mathcal{S}'_i}|f|(x) \quad \text{for a.e. } x \in P. \tag{3.6}$$

The factor m in (3.2) precluded a naive adaptation of the proof in [23] to an A_2 theorem with the usual Dini condition:

$$\int_0^1 \omega(t)\frac{dt}{t} < \infty, \tag{3.7}$$

since the sum

$$\sum_{m=0}^{\infty} \omega(2^{-m})(m+1) \simeq \int_0^1 \omega(t)\Big(1+\log\frac{1}{t}\Big)\frac{dt}{t} \tag{3.8}$$

could diverge for some moduli ω satisfying only (3.7). Moreover, it was shown in [12] that the weak-type $(1,1)$ norm of the adjoints of the operators $\mathcal{A}^m_{\mathcal{S}}$ was at least linear in m, even in the unweighted case, so using duality prevented an extension of this type. However, although our argument does not quite give an A_2 theorem for Calderón-Zygmund operators satisfying the Dini condition (we still need (3.8) to be finite), our proof avoids the use of duality and the study of the adjoint operators $(\mathcal{A}^m_{\mathcal{S}})^*$. It thus removes at least one of the obstructions to possible proofs of the A_2 theorem with the Dini condition which follow this strategy. Hence, removing the linear factor of m in Theorem 3.1 remains as an open problem.

Apart from being interesting in its own right, a bound for Calderón-Zygmund operators by these sums of positive 0-shifts in cases where there is no duality has interesting applications, some of which we describe later. Before, let us state a second corollary to Theorem 3.2:

Corollary 3.3. *Let $\|\cdot\|_{\mathbb{X}}$ be a function quasi-norm (see section 3.2) and T a Calderón-Zygmund operator satisfying the logarithmic Dini condition, then*

$$\|Tf\|_{\mathbb{X}} \lesssim \sup_{\mathscr{D},\mathcal{S}} \|\mathcal{A}^0_{\mathcal{S}}|f|\|_{\mathbb{X}}, \tag{3.9}$$

where the supremum is taken over all dyadic grids $\mathscr{D}$ and all sparse collections $\mathcal{S} \subset \mathscr{D}$.

We now describe two immediate applications of our result. First we can continue the program, initiated in [6] and extended in [28], which aims to extend the sharp weighted estimates for Calderón-Zygmund operators to their multilinear analogues (as in [10]). In particular we obtain

Theorem 3.4. *Let T be a multilinear Calderón-Zygmund operator. Suppose $1 < p_1, \ldots, p_k < \infty$, $\frac{1}{p} = \frac{1}{p_1} + \cdots + \frac{1}{p_k}$ and $\vec{w} \in A_{\vec{P}}$. Then*

$$\|T\vec{f}\|_{L^p(v_{\vec{w}})} \lesssim [\vec{w}]_{A_{\vec{P}}}^{\max(1,\frac{p_1'}{p},\ldots,\frac{p_k'}{p})} \prod_{i=1}^{k} \|f_i\|_{L^p(w_i)}. \tag{3.10}$$

The same theorem was proven in [28] but with the additional hypothesis that p had to be at least 1. The proof of this theorem is an application of the result in [28] which proved the same estimate (without the condition $p \geq 1$) but for a multilinear analogue of the operators $\mathcal{A}^m_{\mathcal{S}}$, together with Theorem 3.1. In fact, we will need a multilinear version of Theorem 3.1 which we state and prove in the next section.

Our second application is a sharp aperture weighted estimate for square functions which extends a result in [24]. In particular:

Theorem 3.5. *Let $\alpha > 0$, then the square function $S_{\alpha,\psi}$ for the cone in $\mathbb{R}^{d+1}_{+}$ of apperture α and the standard kernel ψ satisfies*

$$\|S_{\alpha,\psi}f\|_{L^{p,\infty}(\mathbb{R}^d,w)} \lesssim \alpha^d [w]_{A_p}^{1/p} \|f\|_{L^p(\mathbb{R}^d,w)} \quad \textit{for } 1 < p < 2$$

and

$$\|S_{\alpha,\psi}f\|_{L^{2,\infty}(\mathbb{R}^d,w)} \lesssim \alpha^d [w]_{A_2}^{1/2}(1 + \log[w]_{A_2})\|f\|_{L^2(\mathbb{R}^d,w)}. \tag{3.11}$$

An analogous result was shown in [24] for $2 < p < 3$:

$$\|S_{\alpha,\psi}f\|_{L^{p,\infty}(\mathbb{R}^d,w)} \lesssim \alpha^d [w]_{A_p}^{1/2}(1 + \log[w]_{A_p})\|f\|_{L^p(\mathbb{R}^d,w)}.$$

The proof relies on the use of Lerner's pointwise formula and previous results by Lacey and Scurry [19]. However, in [24] the requirement of $p > 2$ was necessary for the same reason why the proof of the multilinear weighted estimates required $p \geq 1$ (a certain space had no satisfactory duality properties). Theorem 3.1 can be used in almost the same way as with the weighted multilinear estimates to prove Theorem 3.5. Indeed, the proofs in [19] and [24] reduce the problem to estimating certain discrete positive operators which can be seen to be particular instances of the positive multilinear m-shifts used in the proof of Theorem 3.4.

As was noted in [19], estimate (3.11) can be seen as an analogue of the result in [26] stablishing the endpoint weighted weak-type estimate for Calderón-Zygmund operators

$$\|Tf\|_{L^{1,\infty}(w)} \lesssim [w]_{A_1}(1 + \log[w]_{A_1})\|f\|_{L^1(w)}.$$

See also [34] for a similar estimate from below and more information on the sharpness of this estimate, known as the weak A_1 conjecture. In this direction, it seems reasonable that Lacey and Scurry's proof in [19] could be adapted to the multilinear setting, however we will not pursue this problem here.

Finally, as a third application of our results, it is possible to give a more direct proof of the result in [15] for the q-variation of Calderón-Zygmund operators satisfying the logarithmic Dini condition by using the pointwise estimate analogous to (3.1) in [15] and then applying Theorem 3.1. However, we will not pursue this argumentation either.

Shortly before uploading this preprint, Andrei Lerner kindly communicated to the authors that he, jointly with Fedor Nazarov, had independently proven a theorem very similar to Corollary 3.2 [25]. Though the hypothesis are the same, their result differs from the one in this note in that we give a localized pointwise estimate while their pointwise estimate is valid for all of $\mathbb{R}^d$. However, our result seems to be as powerful in the applications.

3.2 Pointwise domination

The goal of this section is the proof of Theorem A and its consequences as stated in the introduction. We will prove the result in the level of generality of multilinear operators. Given a cube P_0 on $\mathbb{R}^d$, we will denote by $\mathscr{D}(P_0)$ the dyadic lattice obtained by successive dyadic subdivisions of P_0. By a dyadic grid we will denote any dyadic lattice composed of cubes with sides parallel to the axis. A k-linear positive dyadic shift of complexity m is an operator of the form

$$\mathcal{A}^m_{P_0,\alpha}\vec{f}(x) = \mathcal{A}^m_{P_0,\alpha}(f_1, f_2, \cdots, f_k)(x) := \sum_{\substack{Q\in\mathscr{D}(P_0)\\ Q^{(m)}\subseteq P_0}} \alpha_Q \Big(\prod_{i=1}^{k} \langle f_i\rangle_{Q^{(m)}}\Big) \mathbb{1}_Q(x).$$

As a first step towards the proof of Theorem A, it is convenient to separate the scales of (or

Since we uploaded this document to *arXiv*, two other articles have appeared: [17] and [11], in which similar estimates are obtained.

slice) $\mathcal{A}^m_{P_0,\alpha}$ as follows:

$$\begin{aligned}\mathcal{A}^m_{P_0,\alpha}\vec{f}(x) &= \sum_{n=0}^{m-1}\sum_{j=1}^{\infty}\sum_{Q\in\mathscr{D}_{jm+n}(P_0)} \alpha_Q\Big(\prod_{i=1}^{k}\langle f_i\rangle_{Q^{(m)}}\Big)\mathbb{1}_Q(x)\\ &=: \sum_{n=0}^{m-1}\mathcal{A}^{m,n}_{P_0,\alpha}\vec{f}(x).\end{aligned}$$

Note that $\mathscr{D}_k(P_0)$ denotes the k-th generation of the lattice $\mathscr{D}(P_0)$. Now we rewrite $\mathcal{A}^{m;n}_{P_0,\alpha}$ as a sum of disjointly supported operators of the form $\mathcal{A}^{m;0}_{P,\alpha}$. Indeed,

$$\begin{aligned}\mathcal{A}^{m;n}_{P_0,\alpha}\vec{f}(x) &= \sum_{j=1}^{\infty}\sum_{Q\in\mathscr{D}_{jm+n}(P_0)} \alpha_Q\Big(\prod_{i=1}^{k}\langle f_i\rangle_{Q^{(m)}}\Big)\mathbb{1}_Q(x)\\ &= \sum_{P\in\mathscr{D}_n(P_0)}\sum_{j=1}^{\infty}\sum_{Q\in\mathscr{D}_{jm}(P)} \alpha_Q\Big(\prod_{i=1}^{k}\langle f_i\rangle_{Q^{(m)}}\Big)\mathbb{1}_Q(x)\\ &= \sum_{P\in\mathscr{D}_n(P_0)}\mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x),\end{aligned}$$

which leads to the expression

$$\mathcal{A}^m_{\alpha,P_0}\vec{f}(x) = \sum_{n=0}^{m-1}\sum_{P\in\mathscr{D}_n(P_0)}\mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x).$$

We say that a sequence $\{\alpha_Q\}_{Q\in\mathscr{D}(P_0)}$ is Carleson if its Carleson constant $\|\alpha\|_{\mathrm{Car}(P_0)} < \infty$, where

$$\|\alpha\|_{\mathrm{Car}(P_0)} = \sup_{P\in\mathscr{D}(P_0)}\frac{1}{|P|}\sum_{Q\in\mathscr{D}(P)}\alpha_Q|Q|.$$

The following intermediate step is the key to our approach:

Proposition 3.6. *Let $m \geq 1$ and α be a Carleson sequence. For integrable functions*

$f_1, \ldots, f_k \geq 0$ on P_0 there exists a sparse collection $\mathcal{S}$ of cubes in $\mathscr{D}(P_0)$ such that

$$\mathcal{A}^{m;0}_{P_0,\alpha}\vec{f}(x) \leq C_1\|\alpha\|_{\mathrm{Car}(P_0)} \sum_{Q\in\mathcal{S}} \Big(\prod_{i=1}^{k} \langle f_i\rangle_Q\Big) \mathbb{1}_Q(x),$$

where C_1 only depends on k and d, and in particular is independent of m.

To prove Proposition 3.6 we will proceed in three steps: we will first construct the collection $\mathcal{S}$, then show that we have the required pointwise bound, and finally that $\mathcal{S}$ is sparse. By homogeneity, we will assume that $\|\alpha\|_{\mathrm{Car}(P_0)} = 1$. Also, we will assume that the sequence α is finite, but our constants will be independent of the number of elements in the sequence.

Let $\Delta_{P_0} = 0$ and, for each $Q \in \mathscr{D}_{mj}(P_0)$ with $j \geq 0$, define the sequence $\{\gamma_Q\}_Q$ by

$$\gamma_Q = \max_{R\in\mathscr{D}_m(Q)} \alpha_R.$$

For each $Q \in \mathscr{D}_{mj}(P_0)$ with $j \geq 0$, we will inductively define the quantities Δ_Q and β_Q as follows:

$$\beta_Q = \begin{cases} 0 & \text{if } \Delta_Q - \Big(\prod_{i=1}^{k}\langle f_i\rangle_Q\Big)\gamma_Q \geq 0 \\ 2^{2(k+1)}C_W & \text{otherwise,} \end{cases}$$

where C_W is the boundedness constant of the unweighted endpoint weak-type of the operators $\mathcal{A}^m$ proved in Theorem 3.16 in the last section. Also, for every $R \in \mathscr{D}_m(Q)$ we define

$$\Delta_R = \Delta_Q + (\beta_Q - \alpha_R)\Big(\prod_{i=1}^{k}\langle f_i\rangle_Q\Big).$$

Note that the definition only applies to cubes in $\mathscr{D}_{mj}(P_0)$ for some j. For all other cubes in $\mathscr{D}_{P_0}$, we set $\beta_Q = \Delta_Q = 0$. The collection $\mathcal{S}$ consists of those cubes $Q \in \mathscr{D}(P_0)$ for which $\beta_Q \neq 0$. Note that, since $2^{2(k+1)}C_W > 1 = \|\alpha\|_{\mathrm{Car}(P_0)} \geq \alpha_R$ for all R and by the definition of γ_Q, we must have $\Delta_Q \geq 0$ for all Q. This can be easily seen by induction.

Remark 3.7. We are trying to construct a sparse operator of complexity 0 which dominates $\mathcal{A}^{m;0}_{P_0,\alpha}$. One way to achieve this is to let $\mathcal{S}$ be the collection of all dyadic subcubes of P_0, but of course this does not yield a sparse collection. A better way would be to let $\mathcal{S}$ consist of all dyadic cubes in P_0 for which at least one of its m-th generation children R satisfies $\alpha_R > 0$; unfortunately this yields a collection $\mathcal{S}$ which is not sparse, and in fact it can be seen that the Carleson sequence β associated with this collection can have a Carleson norm $\|\beta\|_{\mathrm{Car}(P_0)}$ which grows exponentially in m.

The main problem with this approach is that, when the time comes to decide whether a cube should be in $\mathcal{S}$ or not, we do not take into account which cubes have been selected in the previous steps. Note that whenever we add a cube Q to $\mathcal{S}$ we are not only "helping" to dominate the portion of $\mathcal{A}^{m;0}_{P_0,\alpha}$ coming from Q, but also what may come from any of its descendants.

One can account for this by having the algorithm use a sort of "memory" to, essentially, keep track of how many cubes in $\mathcal{S}$ (appropriately weighted with the averages of $\vec{f}$) lie above any particular cube. This is the purpose of Δ_Q. This can also be seen as the stopping time algorithm which selects a cube whenever the previously selected cubes do not provide enough height to dominate the operator until that point.

Lemma 3.8. *We have the pointwise bound*

$$\mathcal{A}^{m;0}_{P_0,\alpha}\vec{f}(x) \le \sum_{Q\in\mathscr{D}(P_0)} \beta_Q\Big(\prod_{i=1}^{k}\langle f_i\rangle_Q\Big)\mathbb{1}_Q(x). \tag{3.12}$$

Proof. We will prove by induction the following claim: if $P \in \mathscr{D}_{jm}(P_0)$ for some $j \ge 0$, then

$$\mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x) \le \Delta_P + \sum_{Q\in\mathscr{D}(P)} \beta_Q\Big(\prod_{i=1}^{k}\langle f_i\rangle_Q\Big)\mathbb{1}_Q(x). \tag{3.13}$$

Note that, when $P = P_0$, this is exactly (3.12). Since α is finite, there is a smallest $j_0 \in \mathbb{N}$ such that $\alpha_Q = 0$ for all cubes $Q \in \mathscr{D}_{\ge j_0 m}(P_0)$. Let Q be any cube in $\mathscr{D}_{j_0 m}(P_0)$, we obviously

We use $\mathscr{D}_{\ge k}(P)$ to denote those cubes Q in $\mathscr{D}(P)$ of generation at least k, so $|Q| \le 2^{-dk}|P|$.

have

$$\mathcal{A}^{m;0}_{Q,\alpha}\vec{f} \equiv 0 \quad \text{in } Q.$$

Since $\Delta_Q \geq 0$, the claim (3.13) is trivial for $P \in \mathscr{D}_{j_0 m}(P_0)$. Now, assume by induction that we have proved (3.13) for all cubes $P \in \mathscr{D}_{jm}(P_0)$ with $1 \leq j_1 \leq j$ and let P be any cube in $\mathscr{D}_{(j_1-1)m}(P_0)$. By definition,

$$\mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x) = \sum_{Q\in\mathscr{D}_m(P)} \Big(\alpha_Q\Big(\prod_{i=1}^{k}\langle f_i\rangle_P\Big)\mathbb{1}_Q(x) + \mathcal{A}^{m;0}_{Q,\alpha}\vec{f}(x)\Big).$$

Let $x \in Q \in \mathscr{D}_m(P)$, then by the induction hypothesis and the definition of Δ_Q:

$$\begin{aligned}
\mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x) &\leq \alpha_Q\Big(\prod_{i=1}^{k}\langle f_i\rangle_P\Big) + \Delta_Q + \sum_{R\in\mathscr{D}(Q)} \beta_R\Big(\prod_{i=1}^{k}\langle f_i\rangle_R\Big)\mathbb{1}_R(x)\\
&= \alpha_Q\Big(\prod_{i=1}^{k}\langle f_i\rangle_P\Big) + \Delta_P + (\beta_P - \alpha_Q)\Big(\prod_{i=1}^{k}\langle f_i\rangle_P\Big) + \sum_{R\in\mathscr{D}(Q)} \beta_R\Big(\prod_{i=1}^{k}\langle f_i\rangle_R\Big)\mathbb{1}_R(x)\\
&= \Delta_P + \beta_P\Big(\prod_{i=1}^{k}\langle f_i\rangle_P\Big) + \sum_{R\in\mathscr{D}(Q)} \beta_R\Big(\prod_{i=1}^{k}\langle f_i\rangle_R\Big)\mathbb{1}_R(x)\\
&= \Delta_P + \sum_{R\in\mathscr{D}(P)} \beta_R\Big(\prod_{i=1}^{k}\langle f_i\rangle_R\Big)\mathbb{1}_R(x),
\end{aligned}$$

which is what we wanted to show. □

Lemma 3.9. *The collection $\mathcal{S}$ is sparse.*

Proof. Let $P \in \mathcal{S}$, we have to show that the set

$$F := \bigcup_{Q\subsetneq P, Q\in\mathcal{S}} Q$$

satisfies $|F| \leq \frac{1}{2}|P|$. To this end, let $\mathcal{R}$ be the collection of maximal (strict) subcubes of P which are in $\mathcal{S}$, Note that for all $R \in \mathcal{R}$ we have $R \in \mathscr{D}_{N_R m}(P)$ for some $N_R \geq 1$. We thus have

$$F = \bigsqcup_{R \in \mathcal{R}} R.$$

By maximality, for all $R \in \mathcal{R}$ and dyadic cubes Q with $R \subsetneq Q \subsetneq P$ we have $\beta_Q = 0$. For all $R \in \mathcal{R}$ and $1 \leq j \leq N_R$ we now claim that

$$\Delta_{R^{((N_R-j)m)}} \geq \beta_P \Big(\prod_{i=1}^{k} \langle f_i \rangle_P \Big) - \sum_{\nu=1}^{j} \alpha_{R^{((N_R-\nu)m)}} \Big(\prod_{i=1}^{k} \langle f_i \rangle_{R^{((N_R-\nu+1)m)}} \Big). \tag{3.14}$$

Indeed, one can prove this by induction on j. If $j = 1$ then by definition we have

$$\begin{aligned}
\Delta_{R^{((N_R-1)m)}} &= \Delta_P + (\beta_P - \alpha_{R^{((N_R-1)m)}}) \Big(\prod_{i=1}^{k} \langle f_i \rangle_P \Big) \\
&\geq \beta_P \Big(\prod_{i=1}^{k} \langle f_i \rangle_P \Big) - \alpha_{R^{((N_R-1)m)}} \Big(\prod_{i=1}^{k} \langle f_i \rangle_P \Big),
\end{aligned}$$

since $\Delta_P \geq 0$.

To prove the induction step, observe that (by the induction hypothesis) for $j > 1$

$$\begin{aligned}
\Delta_{R^{((N_R-j)m)}} &= \Delta_{R^{((N_R-j+1)m)}} + (\beta_{R^{((N_R-j+1)m)}} - \alpha_{R^{((N_R-j)m)}}) \Big(\prod_{i=1}^{k} \langle f_i \rangle_{R^{((N_R-j+1)m)}} \Big) \\
&= \Delta_{R^{((N_R-j+1)m)}} - \alpha_{R^{((N_R-j)m)}} \Big(\prod_{i=1}^{k} \langle f_i \rangle_{R^{((N_R-j+1)m)}} \Big) \\
&\geq \beta_P \Big(\prod_{i=1}^{k} \langle f_i \rangle_P \Big) - \sum_{\nu=1}^{j} \alpha_{R^{((N_R-\nu)m)}} \Big(\prod_{i=1}^{k} \langle f_i \rangle_{R^{((N_R-\nu+1)m)}} \Big).
\end{aligned}$$

From (3.14) with $j = N_R$, we have (since the terms are nonnegative)

$$\Delta_R \geq \beta_P \Big(\prod_{i=1}^{k} \langle f_i \rangle_P \Big) - \mathcal{A}_{P,\alpha}^{m;0} \vec{f}(x)$$

for all $x \in R$. Since $\beta_R \neq 0$, we must have

$$\Big(\prod_{i=1}^{k} \langle f_i \rangle_R \Big) \gamma_R - \Delta_R > 0,$$

i.e.:

$$\Big(\prod_{i=1}^{k}\langle f_i\rangle_R\Big)\gamma_R+\mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x)>2^{2(k+1)}C_W\Big(\prod_{i=1}^{k}\langle f_i\rangle_P\Big)$$

for all $x\in R$. Let $\mathcal{G}_P\vec{f}=\sum_{R\in\mathcal{R}}\gamma_R\Big(\prod_{i=1}^{k}\langle f_i\rangle_R\Big)\mathbb{1}_R$, then for all $x\in R$ we have

$$\mathcal{G}_P f(x)+\mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x)>2^{2(k+1)}C_W\Big(\prod_{i=1}^{k}\langle f_i\rangle_P\Big),$$

hence

$$\begin{aligned}|F|&\le\Bigg|\Bigg\{x\in P:\,\mathcal{G}_P\vec{f}(x)+\mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x)>2^{2(k+1)}C_W\Big(\prod_{i=1}^{k}\langle f_i\rangle_P\Big)\Bigg\}\Bigg|\\&\le\frac{\|\mathcal{G}_P+\mathcal{A}^{m;0}_{P,\alpha}\|^{1/k}_{L^1(P)\times\cdots\times L^1(P)\to L^{1/k,\infty}(P)}}{\Big(2^{2(k+1)}C_W\Big(\prod_{i=1}^{k}\langle f_i\rangle_P\Big)\Big)^{1/k}}\Big(\prod_{i=1}^{k}\|f_i\|_{L^1(P)}\Big)^{1/k}\\&=\frac{\|\mathcal{G}_P+\mathcal{A}^{m;0}_{P,\alpha}\|^{1/k}_{L^1(P)\times\cdots\times L^1(P)\to L^{1/k,\infty}(P)}}{(2^{2(k+1)}C_W)^{1/k}}|P|\end{aligned}$$

Let us compute the operator norm $\|\mathcal{G}_P\|_{L^1(P)\times\cdots\times L^1(P)\to L^{1/k,\infty}(P)}$. Observe that, since $\gamma_Q\le 1$ for all Q, the operator $\mathcal{G}$ is pointwise bounded by the multi-linear projection

$$\mathcal{P}_P\vec{f}(x)=\sum_{R\in\mathcal{R}}\Big(\prod_{i=1}^{k}\langle f_i\rangle_R\Big)\mathbb{1}_R(x)=\prod_{i=1}^{k}\Big(\sum_{R\in\mathcal{R}}\langle f_i\rangle_R\mathbb{1}_R(x)\Big).$$

For each $1\le i\le k$, we have $\|\sum_{R\in\mathcal{R}}\langle f_i\rangle_R\mathbb{1}_R\|_{L_1(P)}\le\|f_i\|_{L_1(P)}$. Therefore, by Hölder's inequality we get

$$\|\mathcal{P}_P\vec{f}\|_{L_{1/k,\infty}(P)}\le\prod_{i=1}^{k}\Big\|\sum_{R\in\mathcal{R}}\langle f_i\rangle_R\mathbb{1}_R\Big\|_{L_1(P)}\le\prod_{i=1}^{k}\|f_i\|_{L_1(P)}.$$

On the other hand we have

$$\|\mathcal{A}^{m;0}_{P,\alpha}\vec{f}\|_{L^{1/k,\infty}(P)}\le C_W\prod_{i=1}^{k}\|f_i\|_{L^1(P)}$$

by Theorem W.1. Combining these estimates we get

$$\|\mathcal{G}_P + \mathcal{A}^{m;0}_{P,\alpha}\|_{L^1(P)\times\cdots\times L^1(P)\to L^{1/k,\infty}(P)} \leq 2^{k+1}(1+C_W) \leq 2^{k+2}C_W$$

and the result follows. □

From lemmas 3.8 and 3.9 Proposition 3.6 follows at once. The proof shows that one can actually take $C_1 = 2^{2+k(7+d(2k-1))}$. We are now ready to finish the proof of Theorem A, which we state here in full generality:

Theorem 3.10. *Let α be a Carleson sequence and let P_0 be a dyadic cube. For every k-tuple of nonnegative integrable functions $f_1, \ldots, f_k$ on P there exists a sparse collection $\mathcal{S}$ of cubes in $\mathscr{D}(P)$ such that*

$$\mathcal{A}^m_{P,\alpha}\vec{f}(x) \leq C_2 \sum_{Q\in\mathcal{S}} \Big(\prod_{i=1}^k \langle f_i\rangle_Q\Big) \mathbb{1}_Q(x).$$

Proof. If $m = 0$ we can just apply Proposition 3.6 after noting that $\mathcal{A}^0_{P_0,\alpha}$ can be written as $\mathcal{A}^{1;0}_{P_0,\beta}$, where

$$\beta_Q = \alpha_{Q^{(1)}}.$$

One easily sees that $\|\alpha\|_{\mathrm{Car}(P_0)} = \|\beta\|_{\mathrm{Car}(P_0)}$. Hence, we may assume that $m \geq 1$. Recall the expression

$$\mathcal{A}^m_{P_0,\alpha}\vec{f}(x) = \sum_{n=0}^{m-1} \sum_{P\in\mathscr{D}_n(P_0)} \mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x).$$

from the beginning of the section. By Proposition 3.6, for each $0 \leq n \leq m-1$ and each $P \in \mathscr{D}_n(P_0)$ we can find a sparse collection of cubes $\mathcal{S}^n_P \subset \mathscr{D}(P)$ such that

$$\mathcal{A}^{m;0}_{P,\alpha}\vec{f}(x) \leq C_1\|\alpha\|_{\mathrm{Car}(P_0)} \sum_{Q\in\mathcal{S}^n_P} \Big(\prod_{i=1}^k \langle f_i\rangle_Q\Big) \mathbb{1}_Q(x).$$

Observe that the collection $\mathcal{S}^n = \cup_{P \in \mathscr{D}_n(P_0)} \mathcal{S}_P^n$ is also sparse, so

$$\mathcal{A}_{P_0,\alpha}^m \vec{f}(x) \leq C_1 \|\alpha\|_{\mathrm{Car}(P_0)} \sum_{n=0}^{m-1} \sum_{Q \in \mathcal{S}^n} \Big(\prod_{i=1}^k \langle f_i \rangle_Q \Big) \mathbb{1}_Q(x). \tag{3.15}$$

For $0 \leq n \leq m-1$ define

$$\mu_Q^n = \begin{cases} 1 & \text{if } Q \in \mathcal{S}^n \\ 0 & \text{otherwise.} \end{cases}$$

Since the collections $\mathcal{S}^n$ are sparse, the sequences μ^n are Carleson sequences with $\|\mu^n\|_{\mathrm{Car}(P_0)} \leq 2$, therefore the sequence

$$\mu_Q := \sum_{n=0}^{m-1} \mu_Q^n$$

is also Carleson with $\|\mu\|_{\mathrm{Car}(P_0)} \leq 2m$.

With this we can continue the argument using estimate (3.15) and the case $m = 0$:

$$\begin{aligned}
\mathcal{A}_{P_0,\alpha}^m \vec{f}(x) &\leq C_1 \|\alpha\|_{\mathrm{Car}(P_0)} \sum_{n=0}^{m-1} \sum_{Q \in \mathcal{S}^n} \Big(\prod_{i=1}^k \langle f_i \rangle_Q \Big) \mathbb{1}_Q(x) \\
&= C_1 \|\alpha\|_{\mathrm{Car}(P_0)} \sum_{n=0}^{m-1} \sum_{Q \in \mathscr{D}(P_0)} \mu_Q^n \Big(\prod_{i=1}^k \langle f_i \rangle_Q \Big) \mathbb{1}_Q(x) \\
&= C_1 \|\alpha\|_{\mathrm{Car}(P_0)} \sum_{Q \in \mathscr{D}(P_0)} \mu_Q \Big(\prod_{i=1}^k \langle f_i \rangle_Q \Big) \mathbb{1}_Q(x) \\
&= C_1 \|\alpha\|_{\mathrm{Car}(P_0)} \mathcal{A}_{P_0,\mu}^0 \vec{f}(x) \\
&\leq C_1 \|\alpha\|_{\mathrm{Car}(P_0)} C_1 2m \sum_{Q \in \mathcal{S}} \Big(\prod_{i=1}^k \langle f_i \rangle_Q \Big) \mathbb{1}_Q(x),
\end{aligned}$$

which yields the result with $C_2 = 2C_1^2$. □

Remark 3.11. The above procedure does not rely on any specific property of the Lebesgue measure. In fact, Theorem 3.1 also holds when we replace all averages –both in complexity 0 and complexity m operators– by averages with respect to any other locally finite Borel measure, because the proof is unaffected.

We now detail how to use Theorem 3.1 to derive the multilinear version of corollaries 3.2 and 3.3. For us, a multilinear Calderón-Zygmund operator will be an operator T satisfying

$$T(f_1, \dots, f_k) = \int_{\mathbb{R}^{dk}} K(x, y_1, \dots, y_k) f_1(y_1) \cdot \dots \cdot f_k(y_k) dy_1 \, \dots \, dy_k$$

for all $x \notin \cap_{i=1}^k \operatorname{supp} f_i$ for appropriate f_i. Also we will require that T extends to a bounded operator from $L^{q_1} \times \dots L^{q_k}$ to L^q where

$$\frac{1}{q} = \frac{1}{q_1} + \dots + \frac{1}{q_k},$$

and that it satisfies the size estimate

$$|K(y_0, \dots, y_k)| \leq \frac{A}{\left(\sum_{i,j=0}^k |y_i - y_j|\right)^{kd}}.$$

ω will be the modulus of continuity of the kernel of the operator i.e. a positive nondecreasing continuous and doubling function that satisfies

$$|K(y_0, \dots, y_j, \dots, y_k) - K(y_0, \dots, y_j', \dots, y_k)| \leq C\omega \left(\frac{|y_j - y_j'|}{\sum_{i,j=0}^k |y_i - y_j|} \right) \frac{1}{\left(\sum_{i,j=0}^k |y_i - y_j|\right)^{kd}}$$

for all $0 \leq j \leq k$, whenever $|y_j - y_j'| \leq \frac{1}{2} \max_{0 \leq i \leq k} |y_j - y_i|$. We can now prove Corollary 3.2:

Proof of Corollary 3.2. Fix a measurable f, and a cube $Q_0 \subset \mathbb{R}^d$. Our starting point is the formula

$$|T\vec{f}(x) - m_{T\vec{f}}(Q_0)| \lesssim \sum_{Q \in \mathcal{S}} \sum_{m=0}^{\infty} \omega(2^{-m}) \prod_{i=1}^{m} \langle |f_i| \rangle_{2^m Q} \mathbb{1}_Q(x),$$

which holds for a sparse subcollection $\mathcal{S} \subset \mathscr{D}(Q_0)$ (see [6] and [15], we are implicitly using a slight improvement of Lerner's formula which can be found in [12], Theorem 2.3). Here $m_f(Q)$ denotes the median of a measurable function f over a cube Q (see [23] for the precise definition), which satisfies

$$|m_f(Q)| \lesssim \frac{\|f\|_{L^{1,\infty}(Q)}}{|Q|}.$$

Hence we can just write

$$|T\vec{f}(x)| \lesssim \sum_{m=0}^{\infty} \omega(2^{-m}) \sum_{Q\in\mathcal{S}} \prod_{i=1}^{m} \langle |f_i| \rangle_{2^m Q} \mathbb{1}_Q(x), \tag{3.16}$$

By an elaboration of the well-known one-third trick, it was proven in [15] that there exist dyadic systems $\{\mathscr{D}^\rho\}_{\rho\in\{0,1/3,2/3\}^d}$ such that for every cube Q in $\mathbb{R}^d$ and every $m \geq 1$, there exists $\rho \in \{0, 1/3, 2/3\}^d$ and $R_{Q,m} \in \mathscr{D}^\rho$ such that

$$Q \subset R_{Q,m},\ 2^m Q \subset Q^{(m)},\ 3\ell(Q) < \ell(R_{Q,m}) \leq 6\ell(Q).$$

Also, we may assume that for each $\rho \in \{0, 1/3, 2/3\}^d$ there exists a cube $P(\rho)$ such that $Q_0 \subset P(\rho) \subset c_d P(\rho)$ for some dimensional constant c_d. Using this, we can further write (3.16) as

$$|T\vec{f}(x)| \lesssim \sum_{\rho\in\{0,\frac{1}{3},\frac{2}{3}\}^d} \sum_{m=0}^{\infty} \omega(2^{-m}) \sum_{\substack{Q\in\mathcal{S} \\ R_{Q,m}\in\mathscr{D}^\rho}} \Big(\prod_{i=1}^{k} \langle |f_i| \rangle_{R_{Q,m}^{(m)}} \Big) \mathbb{1}_{R_Q}.$$

Let $\mathcal{F}_m^\rho = \{R_{Q,m} :\ R_Q \in \mathscr{D}^\rho\} \subset \mathscr{D}(P(\rho))$. Then, we can estimate

$$|T\vec{f}(x)| \lesssim 6^d \sum_{\rho\in\{0,\frac{1}{3},\frac{2}{3}\}^d} \sum_{m=0}^{\infty} \omega(2^{-m}) \sum_{R\in\mathcal{F}_m^\rho} \Big(\prod_{i=1}^{k} \langle |f_i| \rangle_{R^{(m)}} \Big) \mathbb{1}_R,$$

since at most 6^d cubes Q in $\mathscr{D}$ are mapped to the same cube $R_{Q,m}$. Define the sequence

$$\alpha_Q^\rho = \begin{cases} 1 & \text{if } Q \in \mathcal{F}_m^\rho \\ 0 & \text{otherwise.} \end{cases}$$

The collections $\mathcal{F}_m^\rho$ are $2^{-1}\cdot 6^{-d}$-sparse, and hence Carleson with constant $2\cdot 6^d$. In order to apply Theorem 3.1, for each fixed $\rho \in \{0, \frac{1}{3}, \frac{2}{3}\}^d$, $m \geq 0$, we now split the sum as follows:

$$\begin{aligned}
\sum_{Q\in\mathscr{D}^\rho} \alpha_Q^\rho \Big(\prod_{i=1}^k \langle |f_i| \rangle_{Q^{(m)}}\Big) \mathbb{1}_Q(x) &= \sum_{Q\in\mathscr{D}_{\geq m}(P(\rho))} \alpha_Q^\rho \Big(\prod_{i=1}^k \langle |f_i| \rangle_{Q^{(m)}}\Big) \mathbb{1}_Q(x) \\
&+ \sum_{\ell=1}^\infty \sum_{Q\in\mathscr{D}_{m-\ell}(P(\rho))} \alpha_Q^\rho \Big(\prod_{i=1}^k \langle |f_i| \rangle_{Q^{(m)}}\Big) \mathbb{1}_Q(x) \\
&= \mathrm{I} + \mathrm{II}.
\end{aligned}$$

Now, since f_i is supported on $Q_0 \subset P(\rho)$ for $1 \leq i \leq k$ and all $\rho \in \{0, \frac{1}{3}, \frac{2}{3}\}^d$, we claim that $\mathrm{II} \leq \mathrm{I}$. Indeed, compute

$$\begin{aligned}
\sum_{\ell=1}^\infty \sum_{Q\in\mathscr{D}_{m-\ell}(P(\rho))} \alpha_Q^\rho \Big(\prod_{i=1}^k \langle |f_i| \rangle_{Q^{(m)}}\Big) \mathbb{1}_Q(x) &\leq \sum_{\ell=1}^\infty \sum_{Q\in\mathscr{D}_{m-\ell}(P(\rho))} \Big(\prod_{i=1}^k \langle |f_i| \rangle_{Q^{(m)}}\Big) \mathbb{1}_Q(x) \\
&= \sum_{\ell=1}^\infty \Big(\prod_{i=1}^k \langle |f_i| \rangle_{P(\rho)^{(\ell)}}\Big).
\end{aligned}$$

Now observe that, by the support condition on the tuple $\vec{f}$,

$$\prod_{i=1}^k \langle |f_i| \rangle_{P(\rho)^{(\ell)}} = 2^{-dk\ell} \prod_{i=1}^k \langle |f_i| \rangle_{P(\rho)},$$

which is enough to prove the claim. Therefore, we only need to work in the localized cubes $P(\rho)$, $\rho \in \{0, \frac{1}{3}, \frac{2}{3}\}^d$. Therefore, we can obtain the first assertion of Corollary 3.2 applying Theorem 3.1:

$$
\begin{aligned}
|T\vec{f}(x)| &\lesssim \sum_{\rho\in\{0,\frac{1}{3},\frac{2}{3}\}^d} \sum_{m=0}^{\infty} \omega(2^{-m}) \sum_{Q\in\mathscr{D}^\rho,\, Q\subset P(\rho)^{(m)}} \alpha_Q^\rho \Big(\prod_{i=1}^k \langle |f_i|\rangle_{Q^{(m)}}\Big) \mathbb{1}_Q(x) \\
&\lesssim \sum_{\rho\in\{0,\frac{1}{3},\frac{2}{3}\}^d} \sum_{m=0}^{\infty} \omega(2^{-m})(m+1) \sum_{Q\in\mathcal{S}_{m,\vec{f}}} \Big(\prod_{i=1}^k \langle |f_i|\rangle_Q\Big) \mathbb{1}_Q \\
&= \sum_{\rho\in\{0,\frac{1}{3},\frac{2}{3}\}^d} \sum_{m=0}^{\infty} \omega(2^{-m})(m+1) \mathcal{A}_{\mathcal{S}_{m,\vec{f}}}\vec{f}(x),
\end{aligned}
$$

for sparse collections $\mathcal{S}_{m,\vec{f}}$ that may depend both on m and $\vec{f}$ (and which are subfamilies of $\mathscr{D}(P(\rho))$ for each value of ρ). Now, reorganizing the sum above we obtain

$$
\begin{aligned}
|T\vec{f}(x)| &\lesssim \sum_{\rho\in\{0,\frac{1}{3},\frac{2}{3}\}^d} \sum_{\mathcal{S}_{m,\vec{f}}\subset\mathscr{D}^\rho} \omega(2^{-m})(m+1)\mathcal{A}_{\mathcal{S}_{m,\vec{f}}}\vec{f}(x) \\
&=: \sum_{\rho\in\{0,\frac{1}{3},\frac{2}{3}\}^d} \mathcal{A}_\rho \vec{f}(x).
\end{aligned}
$$

Now, by the logarithmic Dini condition, each of the operators $\mathcal{A}_\rho$ is bounded above by some absolute constant times a 0-shift whose associated sequence is 1-Carleson (and localized in $P(\rho)$) to which we can apply again Theorem 3.1. Therefore, we obtain

$$
|T\vec{f}(x)| \lesssim \sum_{\rho\in\{0,\frac{1}{3},\frac{2}{3}\}^d} \mathcal{A}_{\mathcal{S}_\rho}\vec{f}(x),
$$

for some sparse families $\mathcal{S}_\rho \subset \mathscr{D}^\rho$ which depend on $\vec{f}$. $\square$

We now introduce the notion of function quasi-norm. We say that $\|\cdot\|_{\mathbb{X}}$, defined on the set of measurable functions, is a function quasi-norm if:

(P1) There exists a constant $C > 0$ such that

$$
\|f+g\|_{\mathbb{X}} \le C\big(\|f\|_{\mathbb{X}} + \|g\|_{\mathbb{X}}\big),
$$

(P2) $\|\lambda f\|_{\mathbb{X}} = |\lambda| \|f\|_{\mathbb{X}}$ for all $\lambda \in \mathbb{C}$.

(P3) If $|f(x)| \leq |g(x)|$ almost-everywhere then $\|f\|_{\mathbb{X}} \leq \|g\|_{\mathbb{X}}$.

(P4) $\| \liminf_{n\to\infty} f_n \|_{\mathbb{X}} \leq \liminf_{n\to\infty} \|f_n\|_{\mathbb{X}}$

Fix some dyadic system $\mathscr{D}$ such that there exists an increasing sequence of dyadic cubes $\{P_\ell\}_\ell \subset \mathscr{D}$ whose union is the whole space $\mathbb{R}^d$, and denote $\mathbb{1}_{P_\ell} \vec{f} = (\mathbb{1}_{P_\ell} f_1, \ldots, \mathbb{1}_{P_\ell} f_k)$. Now, taking into account properties (P1) and (P3), if we take quasi-norms in the second assertion of Corollary 3.2, we have

$$\|\mathbb{1}_{P_\ell} T(\mathbb{1}_{P_\ell} \vec{f})\|_{\mathbb{X}} \lesssim \sup_{\mathscr{D},\mathcal{S}} \|\mathcal{A}_{\mathcal{S}}(\mathbb{1}_{P_\ell} \vec{f})\|_{\mathbb{X}} \ \forall \ell.$$

On the one hand, since $\vec{f}$ is integrable, $T(\mathbb{1}_{P_\ell} \vec{f})$ converges pointwise to $T(\vec{f})$. Therefore, we have

$$\mathbb{1}_{P_\ell} T(\mathbb{1}_{P_\ell} \vec{f}) \to T(\vec{f})$$

pointwise. Finally, we apply property (P4) and we get

$$\|T\vec{f}\|_{\mathbb{X}} = \left\| \liminf_{\ell} \mathbb{1}_{P_\ell} T(\mathbb{1}_{P_\ell} \vec{f}) \right\|_{\mathbb{X}} \leq \liminf_{\ell} \left\| \mathbb{1}_{P_\ell} T(\mathbb{1}_{P_\ell} \vec{f}) \right\|_{\mathbb{X}} \lesssim \sup_{\mathscr{D},\mathcal{S}} \left\| \mathcal{A}_{\mathcal{S}} \vec{f} \right\|_{\mathbb{X}}.$$

This is exactly Corollary 3.3.

Remark 3.12. We note that the dependence on m in the pointwise estimate of shifts of complexity m must be at least linear in m. To see this, let us work in dimension one and fix a large integer m. For any interval $I = [a, b)$ let I_j be the j-th interval of $\mathscr{D}_m(I)$:

$$I_j = a + |I|[j2^{-m}, (j+1)2^{-m}).$$

Define a tower over an interval I to be the collection of intervals

$$\mathcal{T}_I = \{[a, a + 2^{-k}|I|) : k \in \mathbb{N}\}.$$

The collection of intervals $\mathcal{S} = \bigcup_{J \in \mathscr{D}_m(I)} \mathcal{T}_J$ is a sparse collection. Now consider a function f on I which is defined by

$$f(x) = \begin{cases} 0 & \text{if } x \in I_j \text{ with } j \text{ even,} \\ 2 & \text{otherwise.} \end{cases}$$

Denote $\operatorname{gen}(J) = \log_2(\ell(I)\ell(J)^{-1})$ for cubes $J \in \mathscr{D}(I)$. Observe that for any dyadic interval $J \subseteq I$ with $\operatorname{gen}(J) \leq m - 1$ we have

$$\langle f \rangle_J = 1.$$

Consider now the action of $\mathcal{A}_{\mathcal{S}}^m$ on f. If $x \in (I_j)_0$ with j even then

$$\mathcal{A}_{\mathcal{S}}^m f(x) = m.$$

In order to construct a collection $\mathcal{S}'$ of intervals in I for which we have

$$\mathcal{A}_{\mathcal{S}}^m f(x) \leq C \mathcal{A}_{\mathcal{S}'}^0 f(x),$$

we would need to select every interval $J \subset I$ with $\operatorname{gen}(J) \geq m - 1$. Indeed, let $I^k(x)$ be the interval in $\mathscr{D}_k(I)$ which contains x and let α_J be 1 if $J \in \mathcal{S}'$ and 0 otherwise. Then

$$C \mathcal{A}_{\mathcal{S}'}^0 f(x) = C \sum_{k=0}^{m-1} \alpha_{I^k(x)} \geq m$$

for all $x \in (I_j)_0$ with j even. This implies that at least m/C of these intervals must be in

$\mathcal{S}'$. But this implies that the height

$$\sum_{J\in\mathcal{S}'} \alpha_J \mathbb{1}_J(x) \geq m/C$$

on half of the interval I, which contradicts the hypothesis of $\mathcal{S}'$ being sparse if m is large enough.

3.3 Applications

We are now ready to fully state and prove the applications of the pointwise bound as stated in the introduction. We begin with the multilinear sharp weighted estimates:

3.3.1 Multilinear A_2 theorem

We need some more definitions first. These were introduced in [27].

Definition 3.13 ($A_{\vec{P}}$ weights)**.** Let $\vec{P} = (p_1, \dots, p_k)$ with $1 \leq p_1, \dots, p_k < \infty$ and $\frac{1}{p} = \frac{1}{p_1} + \cdots + \frac{1}{p_k}$. Given $\vec{w} = (w_1, \dots, w_k)$, set

$$v_{\vec{w}} = \prod_{i=1}^{k} w_i^{p/p_i}.$$

We say that $\vec{w}$ satisfies the k-linear $A_{\vec{P}}$ condition if

$$[\vec{w}]_{A_{\vec{P}}} := \sup_Q \Big(\frac{1}{|Q|}\int_Q v_{\vec{w}}\Big) \prod_{i=1}^{k} \Big(\frac{1}{|Q|}\int_Q w_i^{1-p_i'}\Big)^{p/p_i}.$$

We call $[\vec{w}]_{A_{\vec{P}}}$ the $A_{\vec{P}}$ constant of $\vec{w}$. As usual, if $p_i = 1$ then we interpret $\frac{1}{|Q|}\int_Q w_i^{1-p_i'}$ to be $(\operatorname{ess\,inf}_Q w_i)^{-1}$.

The following theorem was proved in [28]:

Theorem 3.14. *Suppose* $1 < p_1, \dots, p_k < \infty$, $\frac{1}{p} = \frac{1}{p_1} + \cdots + \frac{1}{p_k}$ *and* $\vec{w} \in A_{\vec{P}}$. *Then*

$$\|\mathcal{A}_S \vec{f}\|_{L^p(v_{\vec{w}})} \lesssim [w]_{A_{\vec{P}}}^{\max(1, \frac{p_1'}{p}, \dots, \frac{p_k'}{p})} \prod_{i=1}^{k} \|f_i\|_{L^p(w_i)},$$

whenever $\mathcal{S}$ *is sparse.*

We can now use Corollary 3.3 to extend the above result to general k-linear Calderón-Zygmund operators:

Theorem 3.15. *Under the conditions of Theorem 3.14, for any k-linear Calderón-Zygmund operator T, we have*

$$\|T\vec{f}\|_{L^p(v_{\vec{w}})} \lesssim [\vec{w}]_{A_{\vec{P}}}^{\max(1, \frac{p_1'}{p}, \dots, \frac{p_k'}{p})} \prod_{i=1}^{k} \|f_i\|_{L^p(w_i)}.$$

Proof. We just need to apply Corollary 3.3 with $\|\cdot\|_{\mathbb{X}} := \|\cdot\|_{L^p(v_{\vec{w}})}$, which clearly is a function quasi-norm. The assumption of $\vec{f}$ being integrable is a qualitative one and can be trivially removed by the usual density arguments. □

3.3.2 Sharp aperture weighted Littlewood-Paley theorem

Here we follow Lerner [24], the reader can find a nice introduction and some references there. We begin with some definitions:

Let $\psi \in L^1(\mathbb{R}^d)$ with $\int_{\mathbb{R}^d} \psi(x)\, dx = 0$ satisfy

$$|\psi(x)| \lesssim \frac{1}{(1+|x|)^{d+\epsilon}} \tag{3.17}$$

$$\int_{\mathbb{R}^d} |\psi(x+h) - \psi(x)|\, dx \lesssim |h|^{\epsilon}. \tag{3.18}$$

We will denote the upper half-space $\mathbb{R}^d \times \mathbb{R}$ by $\mathbb{R}^{d+1}_+$ and the α-cone at x by

$$\Gamma_\alpha(x) = \{(y,t) \in \mathbb{R}^{d+1}_+ : |y - x| \leq \alpha t\}.$$

Let ψ_t be the dilation of ψ which preserves the L^1 norm, i.e.: $\psi_t(x) = t^{-d}\psi(x/t)$, then we can define the square function $S_{\alpha,\psi}f$ by

$$S_{\alpha,\psi}f(x) = \Big(\int_{\Gamma_\alpha(x)} |(f * \psi_t)(y)|^2 \, \frac{dy\,dt}{t^{d+1}}\Big)^{1/2}.$$

We will also need a regularized version. Let Φ be a Schwartz function such that

$$\mathbb{1}_{B(0,1)}(x) \leq \Phi(x) \leq \mathbb{1}_{B(0,2)}(x).$$

We define the regularized square function $\widetilde{S}_{\alpha,\psi}$ by

$$\widetilde{S}_{\alpha,\psi}f(x) = \Big(\int_{\mathbb{R}^{d+1}_+} \Phi\Big(\frac{x-y}{t\alpha}\Big)|(f * \psi_t)(y)|^2 \, \frac{dy\,dt}{t^{d+1}}\Big)^{1/2}.$$

The regularized version can be used instead of $S_{\alpha,\psi}$ in most cases since we have

$$S_{\alpha,\psi}f(x) \leq \widetilde{S}_{\alpha,\psi}f(x) \leq S_{2\alpha,\psi}f(x).$$

It was proved in [24] that

$$|(\widetilde{S}_{\alpha,\psi}f(x))^2 - (m_{Q_0}(\widetilde{S}_{\alpha,\psi}f)^2)| \lesssim \alpha^{2d} \sum_{m=0}^{\infty} 2^{-\delta m} \sum_{Q\in\mathcal{S}} \langle |f| \rangle_{2^m Q}^2 \mathbb{1}_Q(x)$$

By the same Theorem 3.1 in its bilinear formulation (with $f_1 = f_2 = f$), the last expression can be bounded, up to a constant, by an expression of the form

$$\alpha^{2d} \sum_{\rho\in\{0,\frac{1}{3},\frac{2}{3}\}^d} \sum_{m=0}^{\infty} 2^{-\delta m}(m+1) \sum_{Q\in\mathcal{S}^{\rho,m}} \langle |f| \rangle_Q^2 \mathbb{1}_Q(x).$$

As in [24], we know (a priori) that $m_{Q_0}(\widetilde{S}_{\alpha,\psi}f) \to 0$ as $|Q| \to \infty$ so by the triangle inequality and Fatou's lemma we can ignore that term (or by arguing as we did in the previous section). Finally, arguing as in the proof of corollaries 3.2 and 3.3, we arrive at

$$\|\widetilde{S}_{\alpha,\psi}f\|_{L^{p,\infty}(w)} \lesssim \alpha^d \sup_{\mathscr{D},\mathcal{S}} \|\mathcal{A}^0_{\mathcal{S}}(f,f)^{1/2}\|_{L^{p,\infty}(w)},$$

where the supremum is taken over all dyadic grids $\mathscr{D}$ and all sparse collections $\mathcal{S} \subset \mathscr{D}$. To finish the argument we recall the following result, which was shown in [19]:

$$\|\mathcal{A}^0_{\mathcal{S}}(f,f)^{1/2}\|_{L^{p,\infty}(w)} \lesssim [w]_{A_p}^{\max(\frac{1}{2},\frac{1}{p})} \Phi_p([w]_{A_p})\|f\|_{L^p(w)} \tag{3.19}$$

for $1 < p < 3$, where

$$\Phi_p(t) = \begin{cases} 1 & \text{if } 1 < p < 2 \\ 1 + \log t & \text{if } 2 \le p < 3. \end{cases}$$

We are thus able to extend Lerner's estimate to $1 < p \le 2$, obtaining

$$\|S_{\alpha,\psi}f\|_{L^{p,\infty}(w)} \lesssim \alpha^d [w]_{A_p}^{1/p} \|f\|_{L^p(w)} \quad \text{for } 1 < p < 2$$

and

$$\|S_{\alpha,\psi}f\|_{L^{2,\infty}(w)} \lesssim \alpha^d [w]_{A_2}^{1/2} (1 + \log[w]_{A_2})\|f\|_{L^2(w)}.$$

3.4 The weak-type estimate for multilinear m-shifts

Here we prove the weak-type estimate for k-linear m-shifts needed in section 3.2. Notice that the only important point of the estimates below is the independence of the constants from the parameter m. The proof could be more or less standard by now, but the authors have not been able to find it elsewhere. Therefore we include it for completeness.

Theorem 3.16.

$$\sup_{\lambda>0} \lambda |\{x \in P_0 : \mathcal{A}^m_{P_0,\alpha}\vec{f}(x) > \lambda\}|^k \leq C_W \|\alpha\|_{\mathrm{Car}(P_0)} \prod_{i=1}^{k} \|f_i\|_{L^1(P_0)}, \tag{3.20}$$

where $C_W > 0$ only depends on k and d, and in particular is independent of m.

We will essentially follow Grafakos-Torres [10] and [14]. We first prove an L^2 bound and then apply a Calderón-Zygmund decomposition. For the L^2 bound we will use a multilinear Carleson embedding theorem by W. Chen and W. Damián [2], from which we only need the unweighted result:

$$\Big(\sum_{Q\in\mathscr{D}(P_0)} \alpha_Q \Big(\prod_{i=1}^{k} \langle f_i\rangle_Q\Big)^p\Big)^{\frac{1}{p}} \leq \|\alpha\|_{\mathrm{Car}(P_0)} \prod_{i=1}^{k} p_i' \|f_i\|_{L^{p_i}(P_0)} \tag{3.21}$$

whenever

$$\frac{1}{p} = \frac{1}{p_1} + \cdots + \frac{1}{p_k}.$$

Now we can prove

Proposition 3.17.

$$\|\mathcal{A}^m_{P_0,\alpha}\vec{f}\|_{L^2(P_0)} \leq 4\|\alpha\|_{\mathrm{Car}(P_0)} \prod_{i=1}^{k} \|f_i\|_{L^{2k}(P_0)}$$

Proof. We begin by using duality and homogeneity to reduce to showing

$$\int_{P_0} g(x)\mathcal{A}^m_{P_0,\alpha}\vec{f}(x)\,dx \leq 4$$

assuming that $\|f_i\|_{L^{2k}(P_0)} = \|g\|_{L^2(P_0)} = \|\alpha\|_{\mathrm{Car}(P_0)} = 1$ and $g \geq 0$. By definition and Cauchy-Schwarz, this is equivalent to

$$\Big(\sum_{Q\in\mathscr{D}_{\geq m}(P_0)} \alpha_Q \Big(\prod_{i=1}^{k} \langle f_i\rangle_{Q^{(m)}}\Big)^2 |Q|\Big)^{1/2} \Big(\sum_{Q\in\mathscr{D}_{\geq m}(P_0)} \alpha_Q \langle g\rangle_Q^2 |Q|\Big)^{1/2}.$$

The second term can be estimated, using (3.21) in the linear case, by

$$\Big(\sum_{Q\in\mathscr{D}_{\geq m}(P_0)} \alpha_Q \langle g\rangle_Q^2 |Q|\Big)^{1/2} \leq 2.$$

For the first term observe that the sequence β_Q defined by

$$\beta_Q = \frac{1}{2^{dm}} \sum_{R\in\mathscr{D}_m(Q)} \alpha_R$$

is a Carleson sequence adapted to P_0 of the same constant. Indeed:

$$\begin{aligned}\frac{1}{|Q|}\sum_{R\in\mathscr{D}(Q)} \beta_R|R| &= \frac{1}{|Q|}\sum_{R\in\mathscr{D}(Q)} |R|\frac{1}{2^{dm}}\sum_{T\in\mathscr{D}_m(R)} \alpha_T \\ &= \frac{1}{|Q|}\sum_{R\in\mathscr{D}(Q)}\sum_{T\in\mathscr{D}_m(R)} \alpha_T|T| \\ &= \frac{1}{|Q|}\sum_{R\in\mathscr{D}_{\geq m}(Q)} \alpha_R|R| \\ &\leq \|\alpha\|_{\mathrm{Car}(I)} \\ &= 1.\end{aligned}$$

Therefore, we can write the first term as

$$\Big(\sum_{Q\in\mathscr{D}(P_0)} \beta_Q\Big(\sum_{i=1}^{k}\langle f_i\rangle_Q\Big)^2|Q|\Big)^{1/2},$$

which can also be estimated by (3.21) as follows:

$$\Big(\sum_{Q\in\mathscr{D}(P_0)} \beta_Q\Big(\sum_{i=1}^{k}\langle f_i\rangle_Q\Big)^2|Q|\Big)^{1/2} \leq \Big(\frac{2k}{2k-1}\Big)^k \leq 2.$$

Combining both terms we arrive at

$$\int_{P_0} g(x)\mathcal{A}^m_{P_0,\alpha}\vec{f}(x)\,dx \leq 4$$

which is what we wanted. $\square$

Now we can prove Theorem 3.16.

Proof. By homogeneity we can assume $\|\alpha\|_{\mathrm{Car}(P_0)} = \|f_i\|_{L^1(P_0)} = 1$. We now follow the classical scheme which uses the L^2 bound and a standard Calderón-Zygmund decomposition, see for example Grafakos-Torres [10]. However, we need to be careful with the dependence on m, so we will adapt the proof in [14] to our operators.

Assume without loss of generality that $f_i \geq 0$. Define

$$\Omega_i = \{x \in P_0 : \mathcal{M}^d f_i(x) > \lambda^{1/k}\}.$$

If $\langle f_i \rangle_{P_0} > \lambda^{1/k}$ then by the homogeneity assumption

$$|P_0| < \lambda^{-1/k}$$

and the estimate follows. Therefore, we can assume $\langle f_i \rangle_{P_0} \leq \lambda^{1/k}$ for all $1 \leq i \leq k$ and hence we can write Ω_i as a union the cubes in a collection $\mathcal{R}_i$ consisting of pairwise disjoint dyadic (strict) subcubes of P_0 with the property

$$\langle f_i \rangle_R > \lambda^{1/k} \quad \text{and} \quad \langle f_i \rangle_{R^{(1)}} \leq \lambda^{1/k}.$$

For each $1 \leq i \leq k$ let $b_i = \sum_{R \in \mathcal{R}_i} b_i^R$, where

$$b_i^R(x) := \big(f_i(x) - \langle f_i \rangle_R\big) \mathbb{1}_R(x).$$

We now let $g_i = f_i - b_i$.

Observe that we have

$$|g_i(x)| \leq 2^d \lambda^{1/k},$$

as well as

$$|\Omega_i| = \sum_{R \in \mathcal{R}_i} |R| \leq \lambda^{-1/k}.$$

Define $\Omega = \cup_{i=1}^k \Omega_i$, then we have

$$\begin{aligned} |\{x \in P_0 : \mathcal{A}^m_{P_0,\alpha}\vec{f}(x) > \lambda\}| &\leq |\Omega| + |\{x \in P_0 \setminus \Omega : \mathcal{A}^m_{P_0,\alpha}\vec{f}(x) > \lambda\}| \\ &\leq k\lambda^{-1/k} + |\{x \in P_0 \setminus \Omega : \mathcal{A}^m_{P_0,\alpha}\vec{f}(x) > \lambda\}|. \end{aligned} \tag{3.22}$$

To estimate the second term observe that

$$\begin{aligned} \mathcal{A}^m_{P_0,\alpha}\vec{f}(x) &= \mathcal{A}^m_{P_0,\alpha}(\vec{g} + \vec{b})(x) \\ &= \mathcal{A}^m_{P_0,\alpha}\vec{g}(x) + \sum_{j=1}^{2^k-1} \mathcal{A}^m_{P_0,\alpha}(h_1^j, \dots, h_k^j)(x), \end{aligned}$$

where the functions h_i^j are either g_i or b_i and, furthermore, for each $1 \leq j \leq 2^k - 1$ there is at least one $1 \leq i \leq k$ such that $h_i^j = b_i$. Fix j and let i_j be such that $h_{i_j}^j = b_{i_j}$, then

$$\begin{aligned} \mathcal{A}^m_{P_0}(h_1^j, h_2^j, \dots, h_{i_j}^j, \dots, h_k^j)(x) &= \sum_{Q \in \mathscr{D}_{\geq m}(P_0)} \alpha_Q \Big(\prod_{i=1}^k \langle h_i^j \rangle_{Q^{(m)}} \Big) \mathbb{1}_Q(x) \\ &= \sum_{Q \in \mathscr{D}_{\geq m}(P_0)} \alpha_Q \langle b_{i_j} \rangle_{Q^{(m)}} \Big(\prod_{1 \leq i \leq k,\, i \neq i_j} \langle h_i^j \rangle_{Q^{(m)}} \Big) \mathbb{1}_Q(x) \\ &= \sum_{R \in \mathcal{R}_{i_j}} \sum_{Q \in \mathscr{D}_{\geq m}(P_0)} \alpha_Q \langle b_{i_j}^R \rangle_{Q^{(m)}} \Big(\prod_{1 \leq i \leq k,\, i \neq i_j} \langle h_i^j \rangle_{Q^{(m)}} \Big) \mathbb{1}_Q(x) \\ &= \sum_{R \in \mathcal{R}_{i_j}} \sum_{Q \in \mathscr{D}_{> m}(R)} \alpha_Q \langle b_{i_j}^R \rangle_{Q^{(m)}} \Big(\prod_{1 \leq i \leq k,\, i \neq i_j} \langle h_i^j \rangle_{Q^{(m)}} \Big) \mathbb{1}_Q(x). \end{aligned}$$

So we deduce that $\mathcal{A}^m_{P_0,\alpha}(h_1^j, \dots, h_k^j)(x) = 0$ for all $x \notin \Omega_{i_j}$. With this fact we can see that the second term in (3.22) is actually identical to

$$|\{x \in P_0 \setminus \Omega : \mathcal{A}^m_{P_0,\alpha}\vec{g}(x) > \lambda\}|.$$

Now we can use the L^2 bound as follows:

$$
\begin{aligned}
|\{x \in P_0 \setminus \Omega : \mathcal{A}^m_{P_0,\alpha}\vec{g}(x) > \lambda\}| &\le \frac{1}{\lambda^2}\|\mathcal{A}^m_{P_0,\alpha}\vec{g}\|^2_{L^2(P_0)} \\
&\le \frac{16}{\lambda^2}\prod_{i=1}^k \|g_i\|^2_{L^{2k}(P_0)} \\
&\le \frac{16}{\lambda^2}\prod_{i=1}^k \left(2^d\lambda^{1/k}\right)^{\frac{2k-1}{k}}\|g_i\|^{1/k}_{L^1(P_0)} \\
&= \frac{16}{\lambda^2}2^{d(2k-1)}\lambda^{2-1/k} \\
&= 2^{4+d(2k-1)}\lambda^{-1/k}.
\end{aligned}
$$

Putting both estimates together we arrive at

$$
|\{x \in P_0 : \mathcal{A}^m_{P_0,\alpha}\vec{f}(x) > \lambda\}| \le 2^{5+d(2k-1)}\lambda^{-1/k}
$$

which yields the result with $C_W = 2^{k(5+d(2k-1))}$. $\square$

4

On the embedding of A_1 into A_∞

Guillermo Rey

Submitted.

4.1 Introduction

The purpose of this article is to give a quantitative version of the classical embedding between Muckenhoupt classes

$$A_1 \hookrightarrow A_\infty. \tag{4.1}$$

The class A_1 is defined to be all weights $w \geq 0$ for which $Mw \leq Cw$ for some C, where

$$Mf(x) = \sup_{P \ni x} \frac{1}{|P|} \int_P |f(y)|\, dy$$

is the uncentered Hardy-Littlewood maximal operator (here the supremum is taken over cubes with sides parallel to the coordinate axes).

The class A_∞ is defined to be all weights $w \geq 0$ for which there exists a constant C and an exponent $\epsilon > 0$ such that

$$\frac{w(E)}{|P|} \leq C\Big(\frac{|E|}{|P|}\Big)^{\epsilon}$$

for all cubes P and all subsets $E \subseteq P$. See [8] for more equivalent definitions.

It is a well-known fact that every weight in A_1 is also in A_∞; here we give a quantitative version of this embedding.

We will actually work with a wider class of weights, the dyadic A_p weights. To state the result, let us fix a way to quantify exactly how a weight lies in dyadic A_1. Let P be a cube in $\mathbb{R}^d$, we define the $A_1^d(P)$ characteristic of a weight $w \geq 0$ to be

$$[w]_{A_1^d(P)} := \operatorname{ess\,sup}_{x\in P} \frac{M_P^{\text{dyadic}} w(x)}{w(x)},$$

where M_P^{dyadic} is the dyadic maximal operator localized to P:

$$M_P^d f(x) = \sup_{R\in\mathcal{D}(P)} \langle |f| \rangle_R \mathbb{1}_R(x).$$

Here we are denoting by $\mathcal{D}(P)$ the collection of all dyadic subcubes of P, and the average of a function f over a set E by

$$\langle f \rangle_E := \frac{1}{|E|}\int_E f(x)\,dx.$$

Also, we denote the characteristic function of a set E by $\mathbb{1}_E$.

We define the (non-dyadic) A_1 characteristic similarly:

$$[w]_{A_1(P)} := \operatorname{ess\,sup}_{x\in P} \frac{M_P w(x)}{w(x)},$$

where M_P is the uncentered Hardy-Littlewood maximal operator where the cubes are constrained to lie inside P.

The classical way to prove (4.1) proceeds by using the *reverse Hölder inequality* of Coifman-Fefferman [3] (see [13] for a recent sharp reverse Hölder inequality valid in a very general context): for any weight $w \in A_p$ we have

$$\langle w^q \rangle_P \leq C \langle w \rangle_P^q,$$

for some exponent $q > 1$ depending on w. Indeed, let C_{RH} be the best constant in the above inequality (which will depend on q and on how w lies in A_p), then:

$$\begin{aligned} w(E) &= \int_P w \mathbb{1}_E \\ &\leq \Big(\int_P w^q \Big)^{1/q} |E|^{1/q'} \\ &\leq C_{\mathrm{RH}}^{1/q} w(P) \Big(\frac{|E|}{|P|} \Big)^{1/q'}. \end{aligned}$$

For (non-dyadic) A_1 weights the most quantitative version of the reverse Hölder inequality was given by [47] in dimension one. Using the results of [47] one obtains

$$\frac{w(E)}{w(P)} \leq \frac{a}{a-1} \Big(\frac{|E|}{|P|} \Big)^{\frac{1}{a[w]_{A_1}(P)}}$$

for all $a > 1$, so one can get arbitrarily close to the exponent $\frac{1}{[w]_{A_1}}$ at the cost of a multiplicative constant. The results in [47] are, however, valid only for non-dyadic A_p weights, which behave much better in terms of sharp constants; also [47] is valid only in dimension 1.

In [30] A. Melas showed that, for dyadic A_1 weights, one has

$$\Big\langle (M^{\mathrm{dyadic}} w)^p \Big\rangle_P \leq C(p, [w]_{A_1^d}) \langle w \rangle_P^p,$$

for all p such that

$$1 \leq p < \frac{\log(2^d)}{\log\Big(2^d - \frac{2^d-1}{[w]_{A_1^d}}\Big)},$$

and where $C(p,[w]_{A_1^d})$ is a constant which blows-up as p tends to the endpoint above.

Following the same steps as before, this implies an inequality of the form

$$\frac{w(E)}{w(P)} \leq C_\epsilon \Big(\frac{|E|}{|P|}\Big)^\epsilon$$

for all ϵ such that

$$0 \leq \epsilon < -\frac{\log\Big(1 - \frac{2^d-1}{2^d [w]_{A_1^d}}\Big)}{d\log 2} := \epsilon([w]_{A_1^d}, d),$$

and where C_ϵ is a constant which blows-up as ϵ tends to the endpoint $\epsilon([w]_{A_1^d}, d)$.

It was of interest whether one could achieve an estimate with the endpoint $\epsilon([w]_{A_1^d}, d)$, and this was answered positively by A. Osękowski in [41], where he proved the following weak-type estimate:

$$\frac{1}{|P|}\Big|\Big\{x \in P :\ M^{\mathrm{dyadic}} w(x) > 1\Big\}\Big| \leq \langle w \rangle_P^p \tag{4.2}$$

for all p such that

$$1 \leq p \leq \frac{\log(2^d)}{\log\Big(2^d - \frac{2^d-1}{[w]_{A_1^d}}\Big)}.$$

This estimate, coupled with Hölder's inequality for Lorentz spaces yields

$$\frac{w(E)}{w(P)} \leq C_{\epsilon(Q,d)} \Big(\frac{|E|}{|P|}\Big)^{\epsilon(Q,d)}$$

for all weights w with $[w]_{A_1^d} \leq Q$, thus settling the endpoint question of whether a decay rate of $(|E|/|P|)^{\epsilon(Q,d)}$ could be achieved. However, note that Hölder's inequality for Lorentz spaces (when used in this way) has a constant which explodes when $p \to 1$ which in this case implies that the constant $C_{\epsilon(Q,d)}$ will blow-up as $Q \to \infty$.

In this article we improve this conclusion by directly computing the function

$$\mathbb{B}(x, y, m) = \sup \frac{w(E)}{|P|},$$

where the supremum is taken over all sets $E \subseteq P$ with $|E|/|P| = x$, and all dyadic A_1 weights w with $[w]_{A_1^d(P)} \leq Q$, $\langle w \rangle_P = y$ and $\operatorname{ess\,inf}_{z \in P} w(z) = m$.

The expression for $\mathbb{B}$ is a little involved and we refer the reader to section 4.3 for its full form, but we can already give an upper bound for $\mathbb{B}(\cdot, Q, 1)$:

$$\mathbb{B}(x, Q, 1) \leq \widetilde{f}(x) := Q x^{\epsilon(Q,d)}. \tag{4.3}$$

This shows that the decay rate deduced from Osękowski's estimate can be achieved with a uniform constant as $Q \to \infty$ (note that the constant Q cancels when estimating $\frac{w(E)}{w(P)}$). Observe also that this recovers the result of Osękowski when one takes w instead of its maximal function in (4.2), which can be interpreted as a weak-type reverse Hölder inequality. Indeed, assume without loss of generality that $|P| = \operatorname{ess\,inf} w = 1$ and let $E_\lambda = \{x \in P : w(x) > \lambda\}$, then our estimate will show (see (4.12)) that

$$w(E_\lambda) \leq Q\Big(\frac{w(P) - 1}{Q - 1}\Big)\Big(|E_\lambda| \frac{Q - 1}{w(P) - 1}\Big)^{\epsilon(Q,d)}.$$

So integrating w over this set yields

$$\lambda |E_\lambda|^{1-\epsilon(Q,d)} \leq (\langle w \rangle_P - 1)^{1-\epsilon(Q,d)} \Big(\frac{Q}{(Q-1)^{\epsilon(Q,d)}}\Big) \leq \langle w \rangle_P.$$

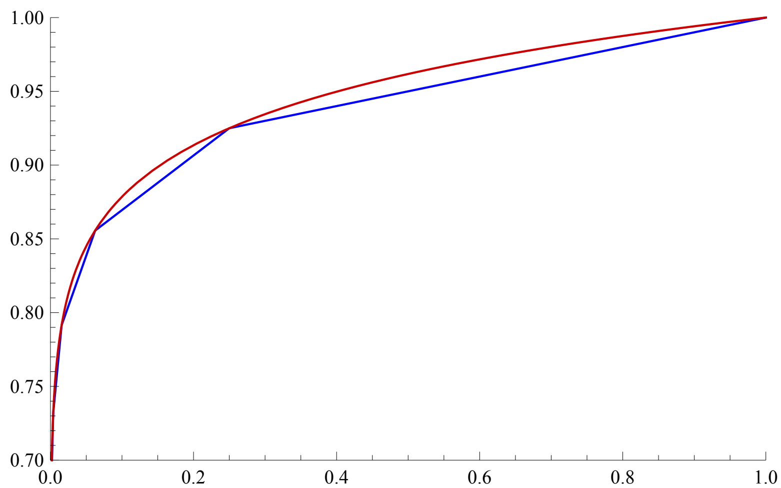
Or, in other words,

$$\|w\|_{L^{p,\infty}} \leq \int_P w(x)\, dx$$

for the same p's as in (4.2).

However, the function $\mathbb{B}(\cdot, Q, 1)$ is, surprisingly, slightly better. Indeed if we define $f(x) = \mathbb{B}(x, Q, 1)$, then our main result shows that f is the piecewise-linear interpolation of the function $\widetilde{f}$ evaluated at the points 2^{-dk} for $k \in \mathbb{N}$.

Figure 4.1 Plots of f and $\widetilde{f}$



In Figure 4.1 we show a normalized section of the plot (the values are divided by Q) of the functions f and $\widetilde{f}$ with $Q = 10$ and in dimension two.

4.1.1 Organization

The article is organized as follows: in section 4.2 we cast the problem as one of finding a certain Bellman function, then in section 4.3 we give a lower bound for the Bellman function; we also describe the structure of the maximizers. In section 4.4 we show that the lower bound found in the previous section is also an upper bound, hence showing that the function found is the actual Bellman function.

4.2 The Bellman function approach

Define, as in the introduction, the function

$$\mathbb{B}(x,y,m)=\sup\Big\{\frac{w(E)}{|P|}: E\subseteq P,\, [w]_{A_1^d(P)}\leq Q \text{ such that } |E|=x|P|,\, \langle w\rangle_P=y,\, m=\operatorname{ess\,inf} w\Big\}.$$

By translation and dilation invariance, the function $\mathbb{B}$ is independent of P.

The domain, which will be denoted by $\Omega_{\mathbb{B}}$ is:

$$0\leq x\leq 1$$
$$0<m\leq y\leq Qm.$$

In this section we cast finding $\mathbb{B}$ as a minimization problem. We will follow the Bellman function method, see for example [36], [47] or [45], and [41] for an approach closer to ours.

The function $\mathbb{B}$ satisfies the following *Main Inequality*

$$\mathbb{B}(x,y,m)\geq \Big\langle \mathbb{B}(x_i,y_i,m_i)\Big\rangle, \tag{4.4}$$

where $\langle x_i\rangle=x$, $\langle y_i\rangle=y$, $\min m_i=m$, $(x_i,y_i,m_i)\in\Omega$, and $(x,y,m)\in\Omega$. In inequality (4.4), and for the rest of the article, we use the notation

$$\langle \xi_i\rangle := \frac{1}{n}\sum_{i=1}^{n}\xi_i,$$

whenever $\{\xi_i\}$ is a discrete sequence of n numbers; usually n will be obvious from the context so we will omit its dependence.

We can see (4.4) by combining almost-extremizers defined on the first-generation dyadic subcubes of P into one on the whole cube P.

We also have the *obstacle condition*

$$\mathbb{B}(1, y, y) = y,$$

which is just the observation that if $E = P$ almost everywhere, then $\langle \mathbb{1}_E w \rangle_P = \langle w \rangle_P$.

From the definition of $\mathbb{B}$ we have the homogeneity property

$$\mathbb{B}(x, \lambda y, \lambda m) = \lambda \mathbb{B}(x, y, m). \tag{4.5}$$

If we find a nonnegative function B defined in $\Omega_{\mathbb{B}}$ and which satisfies the main inequality and the obstacle condition above, then $\mathbb{B} \leq B$. This is a typical fact whose proof we omit, but the reader can consult [41] for a proof in a similar case.

The homogeneity condition will let us assume that $m = 1$ in (4.4):

Proposition 4.1. *If a function B defined on $\Omega_{\mathbb{B}}$ satisfies the main inequality* (4.4) *with $m = 1$ and the homogeneity property* (4.5)*, then it must also satisfy the main inequality for all $m > 0$.*

Proof. This is just the observation that the domain of B is invariant under simultaneous dilations of the variables y and m. □

We want to find a set of necessary and sufficient conditions for B to satisfy the main inequality, but which are simpler to verify. To this end, let us first prove *necessary* conditions that any such B must satisfy.

The following Lemma is simple but important in what follows. It tells us that, in order to exploit (4.4), we should strive to minimize the variables m_i as much as possible. We will let $N := 2^d$ for the rest of the article.

Lemma 4.2. *Any function B satisfying* (4.4) *is decreasing in m. More precisely: assume (x, y, m_1) and (x, y, m_2) are two points in $\Omega_{\mathbb{B}}$ with $m_1 \leq m_2$, then*

$$B(x, y, m_1) \geq B(x, y, m_2). \tag{4.6}$$

Proof. Let $x_i = x$ and $y_i = y$ for all $1 \le i \le 2^d := N$. Also, let

$$\widetilde{m}_i = \begin{cases} m_1 & \text{if } i = 1 \\ m_2 & \text{if } i > 1. \end{cases}$$

Then the points $(x_i, y_i, \widetilde{m}_i)$ are all in Ω. Also, $\langle x_i \rangle = x$ and $\langle y_i \rangle = y$. Since $m_1 \le m_2$ we also have that $\min(\widetilde{m}_i) = m_1$, so using (4.4):

$$B(x, y, m_1) \ge \frac{1}{N} B(x, y, m_1) + \frac{N-1}{N} B(x, y, m_2),$$

which after rearranging yields (4.6). $\square$

The following Lemma follows directly from the main inequality (4.4).

Lemma 4.3. *For any fixed $m > 0$, the function $(x, y) \mapsto B(x, y, m)$ is concave.*

Proof. This is just the observation that the domain Ω is convex, together with (4.4) with $m_i = m$. $\square$

Now we are able to make the first reduction in (4.4) (after the trivial one of setting $m = 1$):

Proposition 4.4. *Suppose B is a nonnegative function defined in $\Omega_{\mathbb{B}}$ and which satisfies the obstacle condition,* (4.5), *and* (4.6). *If B satisfies*

$$B(x, y, 1) \ge \left\langle B\Big(x_i, y_i, \max\Big(1, \frac{y_i}{Q}\Big)\Big)\right\rangle \tag{4.7}$$

for all N-tuples of points (x_i, y_i) satisfying

$$0 \le x_i \le 1, \quad \textit{and} \quad \langle x_i \rangle = x, \tag{4.8}$$

$$1 \le y_i, \quad \min(y_i) \le Q, \quad \textit{and} \quad \langle y_i \rangle = y, \tag{4.9}$$

then we must have that $B = \mathbb{B}$.

Proof. The above conditions make (4.7) certainly necessary. To see that it is sufficient, take any N-tuple (x_i, y_i, m_i) of points in $\Omega_{\mathbb{B}}$ satisfying

$$\langle x_i \rangle = x, \quad \langle y_i \rangle = y \quad \text{and} \quad \min(m_i) = 1.$$

Consider now the alternative N-tuple formed by $(x_i, y_i, \widetilde{m}_i)$, where

$$\begin{aligned}\widetilde{m}_i &= \begin{cases} \frac{y_i}{Q} & \text{if } y_i \geq Q \\ 1 & \text{otherwise.} \end{cases} \\ &= \max\Big(1, \frac{y_i}{Q}\Big).\end{aligned}$$

These points all lie in $\Omega_{\mathbb{B}}$ and moreover they still satisfy the condition

$$\min(\widetilde{m}_i) = 1.$$

However, by inequality (4.6) we have

$$B\Big(x_i, y_i, \max\Big(1, \frac{y_i}{Q}\Big)\Big) \geq B(x_i, y_i, m_i).$$

□

This proposition is useful because it allows us to "almost" eliminate the third variable from our analysis. The reason that we used the word "almost" comes from the fact that we still have the extraneous condition that $\min(y_i) \leq Q$, which is an effect of having $\min(m_i) = 1$. We now proceed to eliminate this condition too.

Suppose that of the N points (x_i, y_i), there are exactly $N - k$ of them for which $y_i \geq Q$. Then, after possibly reordering the inequality (which we can do without loss of generality), the right hand side of (4.7) becomes

$$\frac{1}{N}\Big(\sum_{i=1}^{k} B(x_i, y_i, 1) + \sum_{i=k+1}^{N} B\Big(x_i, y_i, \max\Big(\frac{y_i}{Q}\Big)\Big)\Big)$$

which can be written, after applying the homogeneity property (4.5), as

$$\frac{1}{N}\Big(\sum_{i=1}^{k} B(x_i, y_i, 1) + \sum_{i=k+1}^{N} \frac{y_i}{Q} B(x_i, Q, 1)\Big).$$

So, verifying (4.7) reduces to just showing that B is concave in (x, y), decreasing in m, and that for each $1 \leq k \leq N-1$

$$B(x, y, 1) \geq \frac{1}{N}\Big(\sum_{i=1}^{k} B(x_i, y_i, 1) + \sum_{i=k+1}^{N} \frac{y_i}{Q} B(x_i, Q, 1)\Big) \tag{4.10}$$

for all (x, y) and all (x_i, y_i) as in Proposition (4.4), with the additional assumption that $y_i \geq Q$ for $k \geq k+1$.

The next proposition allows us to just consider the case where $k = N-1$ in the above inequality.

Proposition 4.5. *Let M be a nonnegative function defined on Ω and which satisfies that*

1. *M is concave.*

2. *The function $t \mapsto tM(x, y/t)$ is decreasing.*

3. *For all (x, y) and all $(\widetilde{x}, \widetilde{y})$ in Ω we have*

$$M(x, y) \geq \frac{N-1}{N} M(\widetilde{x}, \widetilde{y}) + \frac{Ny - (N-1)\widetilde{y}}{QN} M(Nx - (N-1)\widetilde{x}, Q) \tag{4.11}$$

whenever $Nx - (N-1)\widetilde{x} \geq 0$ and $Ny - (N-1)\widetilde{y} \geq Q$.

Then, defining B by homogeneity as in (4.5)*:*

$$B(x, y, m) = mM(x, y/m),$$

yields a function which satisfies the conditions of Proposition 4.4

Proof. First of all note that, by the above discussion, we just need to find M satisfying the conditions (1), (2) and

$$M(x,y) \geq \frac{1}{N}\Big(\sum_{i=1}^{k} M(x_i,y_i,1) + \sum_{i=k+1}^{N} \frac{y_i}{Q} M(x_i,Q,1)\Big),$$

where the average of x_i is x, the average of y_i is y and all $y_i \geq Q$ for $i \geq k+1$.

Also, note that (4.11) is just the case of (4.7) with $k = N-1$. So, in what follows we assume $k < N-1$.

Fix all points (x_i, y_i) for $i \leq k$ and consider the collection $\mathcal{V}$ of all vectors $\vec{y} = (y_{k+1}, \dots, y_N)$ with $y_i \geq Q$ for $k \geq k+1$ and satisfying.

$$\frac{1}{N}\sum_{i=k+1}^{N} y_i + \frac{1}{N}\sum_{i=1}^{k} y_i = y.$$

We can write this condition as

$$\widehat{y} := \frac{1}{N-K}\sum_{i=k+1}^{N} y_i = \frac{Ny - \sum_{i=1}^{k} y_i}{N-k} = \frac{Ny - k\widetilde{y}}{N-k},$$

where we have defined $\widetilde{y} = \frac{1}{k}\sum_{i=1}^{k} y_i$.

It is an easy exercise to verify that

$$\frac{1}{N}\sum_{i=k+1}^{N} \frac{y_i}{Q} M(x_i,Q) \leq \frac{1}{QN}\sum_{i=k+1}^{N} b_i M(x_i,Q),$$

where b_i are defined by

$$b_i = \begin{cases} Q & \text{if } i \neq i_{\max} \\ (N-k)\widehat{y} - Q(N-k-1) & \text{if } i = i_{\max}, \end{cases}$$

and where $i_{\max}$ is defined to be the index which maximizes $M(x_i, Q)$ for $i \geq k+1$.

Observe that the vector $(b_{k+1}, \ldots, b_N) \in \mathcal{V}$, so we can assume that $y_i = b_i$ for $i \geq k+1$. But then, we can reorganize the inequality to put all of the terms except one (the one with $i_{\max}$) on the first summation. Writing it this way makes it evident that it really was a particular example of the inequality with $k = N - 1$. □

4.3 Finding the Bellman function

In this section we give a lower bound M for $\mathbb{M}$, and in the next section we will show that this lower bound is also an upper bound and hence that $M = \mathbb{M}$.

First recall that

$$t \mapsto \mathbb{M}(x, y/t)$$

is non-increasing and therefore that $\mathbb{M}(1, y) \geq y$ (here we are using the obstacle $\mathbb{M}(1,1) = 1$. Since $\mathbb{M}(0,1) \geq 0$, we now can extend this bound to the subdomain $y \leq 1 + (Q-1)x$ to get:

$$\mathbb{M}(x, y) \geq x + y - 1 \quad \forall (x, y) \in \Omega : \; y \leq 1 + (Q-1)x.$$

We will now give a lower bound for $\mathbb{M}$ in the rest of the domain. The idea is to use inequality (4.11) setting the number $Nx - (N-1)\widetilde{x}$ to be as large as possible, within the domain that we know, and then iterate.

More precisely let $x_0 = 1$, observe that if $Nx - (N-1)\widetilde{x} = x_0$, then

$$\widetilde{x} = \frac{Nx - x_0}{N - 1}.$$

Clearly we need $x \geq 1/N$ for $\widetilde{x}$ to be in the domain, so we set $x = \frac{1}{N}$. We will also make $\widetilde{y}$ as small as possible, which means $\widetilde{y} = 1$.

Putting it all together we obtain, using (4.11) with $x = \frac{1}{N}$ and $y = Q$:

$$\mathbb{M}\Big(\frac{1}{N}, Q\Big) \geq \frac{NQ - (N-1)}{NQ}\mathbb{M}(x_0, Q) = Q\Big(1 - \frac{N-1}{NQ}\Big).$$

Now we iterate this procedure. Set $x = x_{k+1} = \frac{x_k}{N}$, $y = Q$, $\widetilde{y} = 1$ and $\widetilde{x} = 0$, then (4.11) gives

$$\mathbb{M}(x_{k+1}, Q) \geq (1 - \frac{N-1}{NQ}\Big)\mathbb{M}(x_k, Q),$$

so

$$\mathbb{M}(N^{-k}, Q) \geq Q\Big(1 - \frac{N-1}{NQ}\Big)^k.$$

Between x_{k+1} and x_k we know that $M(\cdot, Q)$ is concave, so $\mathbb{M}$ must certainly be at least linear in these intervals. Now, since $\mathbb{M}(0,1) \geq 0$, we can also extend this bound by homogeneity and get the upper bound

$$\mathbb{M}(x,y) \geq \frac{y-1}{Q-1}\mathbb{M}\Big(x\frac{Q-1}{y-1}, Q\Big) \geq \frac{y-1}{Q-1}f\Big(x\frac{Q-1}{y-1}\Big)$$

for $y - 1 \geq x$. Here, f is the piecewise linear function defined on $[0,1]$ by linearly interpolating the points

$$f(x_k) = Q\Big(1 - \frac{N-1}{NQ}\Big)^k$$

between x_{k+1} and x_k, Figure 4.1 shows what f typically looks like.

Putting it all together, we get

$$\mathbb{M}(x,y) \geq \left\{\begin{array}{ll} x + y - 1 & \text{if } y \leq 1 + (Q-1)x \\ \frac{y-1}{Q-1}f\Big(x\frac{Q-1}{y-1}\Big) & \text{if } y \geq 1 + (Q-1)x. \end{array}\right\} =: M(x,y). \tag{4.12}$$

The way we proved these bounds also shows how one would construct pairs of weights w and sets E showing that $\mathbb{M}$ is at least the promised lower bound. We now give a detailed description of these examples.

4.3.1 Explicit extremizers

Let's start with examples corresponding to the line $(1, y)$ with $y \in [1, Q]$. To get the bound $\mathbb{M}(1, y) \geq y$ we used the main inequality keeping all the parameters fixed except one of the m_i's. So let us repeat the proof, but now with actual weights. Fix a cube P and let $P_1, \ldots P_N$ be its dyadic children. Define $w_i(x) = 1$ for all i and all $x \in P_i$ except for $i = N$, for which we define $w_i(x) = 1 + N(y - 1)$ for all $x \in P_N$. Now define $w(x) = w_i(x)$ for all $x \in P_i$; clearly $\operatorname{ess\,inf}_{x \in P} w(x) = 1$ and $\langle w \rangle_P = y$. Now, since $x = 1$, we should set $E = P$. The pair (w, E) is clearly contained in the supremum in the definition of $\mathbb{B}(1, y, 1) = \mathbb{M}(1, y)$ and so

$$\mathbb{M}(1, y) \geq \frac{w(P)}{|P|} = y \tag{4.13}$$

for this particular choice of w. Of course, any weight with $\langle w \rangle_P = y$ would also have been sufficient since $x = 1$.

Examples for weights and sets corresponding to points (x, y) on the rest of the domain are more complicated. We will start by constructing examples along the line $y = Q$.

The way we proved that $\mathbb{M}(\frac{1}{N}, Q) \geq Q(1 - \frac{N-1}{NQ})$ was by using (4.11) with $\widetilde{x} = 0$, $\widetilde{y} = 1$, $x = \frac{1}{N}$ and $y = Q$. Similarly, we got the bound $\mathbb{M}(x_{k+1}, Q) \geq (1 - \frac{N-1}{NQ})\mathbb{M}(x_k, Q)$ by using (4.11) with $\widetilde{x} = 0$, $\widetilde{y} = 1$, $x = \frac{1}{N^{k+1}}$ and $y = Q$. Looking back at how we got (4.11), we see that we combined $N - 1$ trivial weight-set pairs (the pairs $(w \equiv 1, E = \emptyset)$) with an example coming from

$$\mathbb{B}\Big(\frac{1}{N^k}, N(Q - 1) + 1, N - \frac{N-1}{Q}\Big).$$

We then used homogeneity to translate this to an example which would extremize

$$\mathbb{M}\Big(\frac{1}{N^k}, Q\Big),$$

but having lost a factor slightly larger than one.

We can trace back these steps with the following lemma:

Lemma 4.6. *Let P be a cube in $\mathbb{R}^d$. Given a pair (w, E) where w is a dyadic A_1 weight with $[w]_{A_1} \leq Q$ and with $\langle w\rangle_P = Q$, $\operatorname{ess\,inf}_{z\in P} w(z) = 1$, and $\langle \mathbb{1}_E\rangle_P = x$, there exists a pair $(\widetilde{w}, \widetilde{E})$ where $\widetilde{w}$ is another dyadic A_1 weight with $[w]_{A_1^d} \leq Q$ and with $\langle \widetilde{w}\rangle_P$, $\operatorname{ess\,inf}_{z\in P} \widetilde{w}(z) = 1$, and $\langle \mathbb{1}_E\rangle_P = x/N$ for which*

$$\frac{\widetilde{w}(\widetilde{E})}{|P|} \geq \Big(1 - \frac{N-1}{NQ}\Big)\frac{w(E)}{|P|}.$$

Moreover, the set $\widetilde{E}$ is entirely contained in one of the dyadic subcubes of P and $\widetilde{w}$ is identically 1 on the complement of $\widetilde{E}$.

Proof. As before, enumerate the children of P by $P_1, \ldots, P_N$. We start by translating and dilating (w, E) to the subcube P_1, we do this with the obvious linear change of variables. We then multiply the weight we just constructed by the constant $\frac{NQ-(N-1)}{Q}$. Let us call this new weight w_1. Clearly $\operatorname{ess\,inf}_{z\in P_1} w_1(z) = \frac{NQ-(N-1)}{Q} \geq 1$ and $\langle w_1\rangle_{P_1} = NQ - (N-1)$. Now define $w_i(z) = 1$ for all $z \in P_i$ and each $i \geq 2$ and combine all of these weights into one: $\widetilde{w}(z) = w_i(z)$, for all $z \in P_i$. This new weight is a dyadic A_1 weight with $[\widetilde{w}]_{A_1^d} \leq Q$.

With E we do the same: we translate and dilate it to P_1; let us call this new set E_1. This new set has $\langle \mathbb{1}_{E_1}\rangle = x$. Define $\widetilde{E}$ to be just E_1 (so $\mathbb{1}_{\widetilde{E}}(z) = \mathbb{1}_{E_1}(z)$).

We assert that this new pair $(\widetilde{w}, \widetilde{E})$ satisfies the promised estimate. Indeed (assuming without loss of generality that $|P| = 1$):

$$\begin{aligned}\widetilde{w}(\widetilde{E}) &= \frac{1}{N}\Big((N-1)w_2(\widetilde{E}) + w_1(\widetilde{E})\Big)\\ &= \frac{1}{N}w_1(\widetilde{E})\\ &= \Big(1 - \frac{N-1}{NQ}\Big)w(E),\end{aligned}$$

which is what we wanted. $\square$

Given a cube P and a pair (w, E) as in Lemma 4.6, we define

$$T(w) = \widetilde{w},$$

where $\widetilde{w}$ is the weight constructed in the proof of Lemma 4.6. Similarly, we define

$$S(E) = \widetilde{E}.$$

With this lemma at hand we can now describe the structure of the examples which show that $\mathbb{M}(N^{-k}, Q) \geq Q(1 - \frac{N-1}{NQ})^k$.

Lemma 4.7. *Let P be any cube and let w_0 be the weight constructed when proving* (4.13) *(but any weight with $\langle w_0 \rangle_P = Q$, $\operatorname{ess\,inf}_{z\in P} w_0(z) = 1$, and with $[w]_{A_1^d} = Q$ will work as well). Define the weights w_k and the sets E_k inductively by*

$$w_{k+1} = Tw_k \quad \textit{and} \quad E_{k+1} = SE_k,$$

where $E_0 = P$.

Then w_k is an A_1^d weight with $[w]_{A_1^d} = Q$, $\langle w_k \rangle_P = Q$, $\operatorname{ess\,inf}_{z\in P} w_k(z) = 1$, $\langle \mathbb{1}_{E_k} \rangle_P = N^{-k}$ and

$$\frac{w_k(E_k)}{|P|} = Q\Big(1 - \frac{N-1}{NQ}\Big)^k.$$

Proof. The proof is just to iteratively apply Lemma 4.6. $\square$

It remains to extend the examples to the rest of the domain. But recall that the bound we gave for $\mathbb{M}$ on the rest of the domain was obtained by linear interpolation, so we just need to combine examples that have already been constructed.

The following lemma shows how to combine two pairs (w_0, E_0) and (w_1, E_1) into one:

Lemma 4.8. *Let P be a cube and let (w_0, E_0) and (w_1, E_1) be two pairs. Assume w_0 and w_1 are both dyadic A_1 weights with $[w_i]_{A_1^d} \leq Q$, and also:*

$$\langle \mathbb{1}_{E_i} \rangle_P = x_i, \quad \langle w_i \rangle_P = y_i, \quad \operatorname{ess\,inf}_{z \in P} w_i(z) = 1.$$

Then, for any $\lambda \in [0,1]$ we can construct a pair $\mathcal{C}_\lambda((w_0, E_1), (w_1, E_1)) = (w, E)$, where w is a dyadic A_1 weight with $[w]_{A_1^d} \leq Q$,

$$\langle \mathbb{1}_E \rangle_P = x, \quad \langle w \rangle_P = y, \quad \operatorname{ess\,inf}_{z \in P} w(z) = 1,$$

and

$$\frac{w(E)}{|P|} = (1-\lambda)\frac{w_0(E_0)}{|P|} + \lambda \frac{w_1(E_1)}{|P|},$$

where

$$x = (1-\lambda)x_0 + \lambda x_1 \quad \textit{and} \quad y = (1-\lambda)y_0 + \lambda y_1.$$

Proof. Note that, at least when λ is a dyadic rational, repeated applications of the Main Inequality give exactly these dynamics. So we should follow the proof of the Main Inequality, whose meaning is to show what happens when one combines pairs (w_i, E_i) defined on the dyadic children of a cube into one pair (w, E) on the whole cube.

There is a slight technicality: if one applies this combination procedure a finite number of times, one can only prove this lemma in the case where λ is a dyadic rational, but we can still prove this lemma with a limiting argument.

Let b_i be the digits of λ when written in binary:

$$\lambda = \sum_{i=1}^{\infty} b_i 2^{-i}$$

(it does not matter which of the possible binary representations one uses).

Fix the cube P and let R be any of its dyadic subcubes. Define $S_{P\to R}$ to be the linear change of variables which maps P to R.

Given a cube P let $P_1, \dots P_N$ be a fixed enumeration of its first-generation children, this ordering will be fixed throughout the proof (in the sense that we will use the same ordering on every other cube, which we obtain by translating and dilating the original ordering).

The idea is to split the subcubes of P and on half of them put a translated and dilated copy of either (w_0, E_0) or (w_1, E_1), depending on the binary digit of the current step. We apply the same procedure on each of the remaining cubes (but now with the next digit).

More precisely, let $\mathrm{ch}(P)$ be the first-generation dyadic subcubes of P and define $\mathcal{H}^1_{\pm}(P)$ to be the subset of $\mathrm{ch}(P)$ consisting of the first or second half the dyadic children, i.e.:

$$\mathcal{H}^1_{-}(P) = \{P_1, \dots, P_{2^{d-1}}\} \quad \text{and} \quad \mathcal{H}^1_{+}(P) = \{P_{2^{d-1}+1}, \dots, P_{2^d}\}.$$

We inductively define $\mathcal{H}^{j+1}_{\pm}(P)$ as follows:

$$\mathcal{H}^{j+1}_{\pm}(P) = \bigcup_{R\in\mathcal{H}^j_{+}(P)} \mathcal{H}_{\pm}(R).$$

We define the weight w by

$$w(x) = \sum_{j=1}^{\infty} \sum_{R\in\mathcal{H}^j_{-}(P)} \Big((1-b_j)S_{P\to R}w_0(x) + b_j S_{P\to R}w_1(x)\Big).$$

Similarly, we define the set E by

$$\mathbb{1}_E(x) = \sum_{j=1}^{\infty} \sum_{R\in\mathcal{H}^j_{-}(P)} \Big((1-b_j)S_{P\to R}\mathbb{1}_{E_0}(x) + b_j S_{P\to R}\mathbb{1}_{E_1}(x)\Big).$$

One can now check that this pair satisfies the required properties; see [43] for a very similar construction. $\square$

With this Lemma, we can now express the structure of the examples on the line $y = Q$ of Ω which lie between the points with coordinates $x = N^{-k}$. Indeed, let (w_k, E_k) be the weight-set pair constructed by Lemma 4.7. Then for any $x \in (N^{-k-1}, N^{-k})$ we have

$$(w_x, E_x) := \mathcal{C}_\lambda((w_{k+1}, E_{k+1}), (w_k, E_k)),$$

where

$$x = (1-\lambda)N^{-k-1} + \lambda N^{-k}.$$

To extend to the rest of Ω, let $(x, y) \in \Omega$ with $y < Q$. First assume that $y \leq 1 + (Q-1)x$; then we should use the previous Lemma with boundary on $x = 1$. Indeed let

$$(w, E) = \mathcal{C}_\lambda((\mathbb{1}, \emptyset), (w^y, P)),$$

where $\lambda = 1 + \frac{y-1}{x}$ and where w^y is any dyadic A_1 weight with $[w]_{A_1^d} \leq Q$, $\langle w^y \rangle_P = y$ and $\operatorname{ess\,inf}_{z \in P} w(z) = 1$. This pair clearly satisfies all the required estimates.

Now suppose that $y \geq 1 + (Q-1)x$ and let (w_*, E_*) be the pair we just constructed on the line $y = Q$ with x-coordinate $x\frac{Q-1}{y-1}$. Then

$$(w, E) = \mathcal{C}_\lambda\big((\mathbb{1}, \emptyset), (w_*, E_*)\big),$$

with $\lambda = x\frac{Q-1}{y-1}$ also satisfies all the required estimates.

4.4 Verifying the Main Inequality

We now have to show that the function M that we found in the previous section satisfies all the required conditions which, we recall, are:

1. M is concave.

2. The function $t \mapsto tM(x, y/t)$ is nonincreasing.

3. For all $(x, y) \in \Omega$ and all $(\widetilde{x}, \widetilde{y})$ in Ω with $\widetilde{x} \leq x$ and $Ny - (N-1)\widetilde{y} \geq Q$, we have

$$M(x, y) \geq \frac{N-1}{N} M(\widetilde{x}, \widetilde{y}) + \frac{Ny - (N-1)\widetilde{y}}{NQ} M(Nx - (N-1)\widetilde{x}, Q). \tag{4.14}$$

It will be convenient to examine the function f, in particular observe that

$$f'(x) = (N\eta)^k,$$

where $\eta = 1 - \frac{N-1}{NQ}$, whenever $x \in (N^{-k-1}, N^{-k})$.

The ratio $\eta N > 1$ whenever $Q > (N-1)/N$, which is always the case since $Q \geq 1$, hence f is concave. Since f is concave, it follows that M must also be concave, since M is just the extension of f by homogeneity to the subdomain of Ω which lies above the diagonal $y = 1 + (Q-1)x$, and below this line the function is just the plane $z = x + y - 1$. A brief check now shows that M is indeed concave in Ω. This proves (1).

Now we will show that the function

$$t \mapsto tM(x, y/t)$$

is decreasing, thus proving (2).

To show this, note that we just need to prove $yM_y \geq M$ wherever M is differentiable. This obviously holds for $y < 1 + (Q-1)x$, so it suffices to assume $y > 1 + (Q-1)x$. By homogeneity, we can translate this condition to one for f:

$$\frac{y}{Q-1} f\Big(x\frac{Q-1}{y-1}\Big) - \frac{xy}{y-1} f'\Big(x\frac{Q-1}{y-1}\Big) \geq \frac{y-1}{Q-1} f\Big(x\frac{Q-1}{y-1}\Big).$$

Let $u = x\frac{Q-1}{y-1}$, then this inequality becomes

$$\frac{1}{u} f(u) - yf'(u) \geq 0$$

for all $u \in [0,1]$ and all $y \in [1, Q]$. Since f is increasing, this inequality is strongest when $y = Q$, so it suffices to show

$$f(u) \geq Quf'(u).$$

Recall that f is piecewise linear, so let $u_0 = N^{-k-1}$ and $u_1 = N^{-k}$ and assume $u \in (u_0, u_1)$. The above inequality now becomes

$$f(u_0) + (u - u_0)f'(u_0+) \geq Quf'(u_0+).$$

Thus, we can reduce to showing

$$\frac{f(u_0)}{f'(u_0+)} \geq u_0 + (Q-1)u_1.$$

But an easy computation, using the value of f' computed before, yields that this inequality is equivalent to

$$\eta \geq 1 - \frac{N-1}{NQ},$$

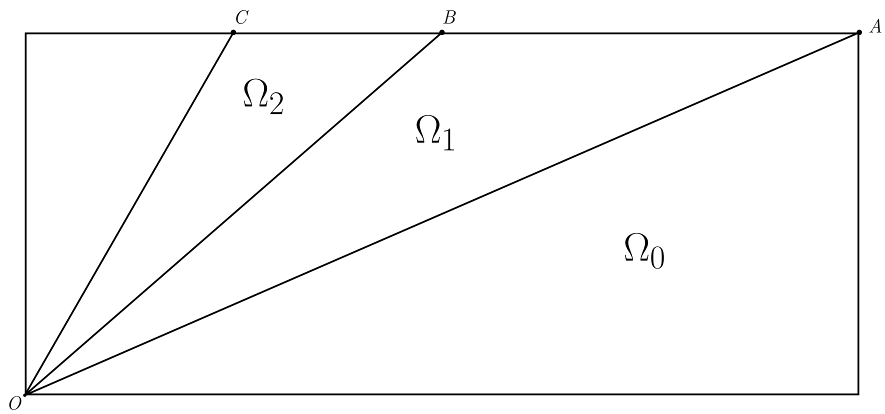
which is precisely the value of η so we are done. This shows (2).

Finally, we are left with verifying (3). To do this we will construct a sequence of functions M_k defined on Ω, all of which satisfy (3) on a specific subset of Ω. Define

$$\Omega_k = \{(x,y) \in \Omega : \, y \leq 1 + (Q-1)N^k x\}.$$

Figure 4.2 represents the first three of these domains (again, the diagram is not to scale). For example Ω_2 is the subdomain of Ω which lies to the right of the line joining O and C.

Figure 4.2 Domains Ω_k



We define M_k to be the wedge formed by the k-th plane of M on $\Omega \setminus \Omega_{k-1}$ and the $(k-1)$-th plane of M on Ω_{k-1}, that is:

$$M_k(x,y) = \begin{cases} a_k x + b_k(y-1) & \text{if } (x,y) \in \Omega \setminus \Omega_{k-1} \\ a_{k-1}x + b_{k-1}(y-1) & \text{if } (x,y) \in \Omega_{k-1}. \end{cases}$$

where $M(x,y) = a_k x + b_k(y-1)$ on $\Omega_k \setminus \Omega_{k-1}$. One can give the explicit formulas for a_k and b_k:

$$a_k = (N\eta)^k, \quad b_k = \eta^k.$$

Obviously M_0 satisfies (3).

Fix any $(x,y) \in \Omega$, we can assume without loss of generality that $(x,y) \in \Omega_k$ for some k. Introduce the notation

$$x = \frac{N-1}{N}\widetilde{x} + \frac{1}{N}\widehat{x} \quad \text{and} \quad y = \frac{N-1}{N}\widetilde{y} + \frac{1}{N}\widehat{y}.$$

Since M is concave, we have that $M_k \geq M$ on Ω (M_k is a "supporting wedge" of the graph of M). Instead of (3) we will prove (under the same hypotheses)

$$M_k(x,y) \geq \frac{N-1}{N} M_k(\widetilde{x},\widetilde{y}) + \frac{1}{N}\frac{\widehat{y}}{Q} M_k(\widehat{x},Q), \tag{4.15}$$

which, by the above remark, is a stronger statement.

We will first show that we can assume the point $(\widehat{x},Q)$ to be in Ω_k. Indeed, suppose that $\widetilde{x}$ is so small that $(\widehat{x},Q) \notin \Omega_k$, then

$$\begin{aligned}
\frac{\partial}{\partial \widetilde{x}}\Big(\text{Right hand side of (4.15)}\Big) &= \frac{N-1}{N} a_k - \frac{N-1}{N}\frac{\widehat{y}}{Q} a_{k-1} \\
&= \Big(\frac{N-1}{N}\Big)\Big(a_k - \frac{\widehat{y}}{Q} a_{k-1}\Big) \\
&\geq \Big(\frac{N-1}{N}\Big)\Big(a_k - \frac{Ny-(N-1)}{Q} a_{k-1}\Big) \\
&\geq \Big(\frac{N-1}{N}\Big)\Big(a_k - \frac{NQ-(N-1)}{Q} a_{k-1}\Big).
\end{aligned}$$

Now recall that $a_k = (N\eta)^k$, so the partial derivative of the right hand side of equation (4.15) is at least

$$\frac{N-1}{N}(N\eta)^{k-1}\Big(N\eta - \frac{NQ-(N-1)}{Q}\Big) = 0,$$

so the right hand side is increasing, at least as long as $(\widehat{x},Q) \in \Omega_{k-1}$.

This allows us to assume that $\widetilde{x}$ is large enough to make $(\widehat{x},Q) \in \Omega_k$ (by continuity). Under this assumption the inequality becomes much easier since M_k is now being evaluated always on Ω_k, and hence we can assume that M_k itself is a plane. Now it is easy to check that the inequality is indeed true under these conditions.

To see this, observe that inequality (4.15) can be written as:

$$ax + b(y-1) \geq \frac{N-1}{N}\big(a\widetilde{x} + b(\widetilde{y}-1)\big) + \frac{1}{N}\frac{\widehat{y}}{Q}\big(a\widehat{x} + b(Q-1)\big).$$

We can reorganize this as:

$$a\Big(x - \frac{N-1}{N}\widetilde{x} - \frac{1}{N}\frac{\widehat{y}}{Q}\widehat{x}\Big) + b\Big(y - 1 - \frac{N-1}{N}\widetilde{y} + \frac{N-1}{N} - \frac{1}{N}\frac{\widehat{y}}{Q}(Q-1)\Big) \geq 0.$$

This simplifies to showing

$$a\Big(\frac{\widehat{x}}{N} - \frac{\widehat{x}}{N}\frac{\widehat{y}}{Q}\Big) + b\Big(\frac{\widehat{y}}{NQ} - \frac{1}{N}\Big) \geq 0,$$

which is equivalent to

$$\Big(\frac{\widehat{y}}{Q} - 1\Big)\big(b - a\widehat{x}\big) \geq 0.$$

Since the assumptions force $\widehat{y}$ to be at least Q, we just need to check that $\widehat{x} \leq \frac{b}{a}$. But this is exactly the bound that is guaranteed from the considerations above since $\frac{b}{a} = N^{-k}$.

5

Borderline weak-type bounds for singular integrals

Carlos Domingo-Salazar, Michael Lacey, and Guillermo Rey

Bulletin of the London Mathematical Society, October 2015.

5.1 Introduction

The purpose of this chapter is to show some applicaitons of the techniques developed so far. The results in this chapter use the pointwise domination of singular integrals from Chapter 3, as well as several facts about sparse families and Muckenhoupt weights.

The first theorem is in the context of linear singular integral operators:

Theorem 5.1. *Let T be a Calderón-Zygmund operator on $\mathbb{R}^d$ and w an A_1 weight, then*

$$\|Tf\|_{L^{1,\infty}(w)} \lesssim_{T,d} [w]_{A_1}\big(1 + \log[w]_{A_\infty}\big)\|f\|_{L^1(w)}.$$

It is unknown whether the logarithmic term is sharp, but a power is necessary, see [35].

We can also state a very similar theorem for square functions:

Theorem 5.2. *Let G be a square function as in 3.5, then*

$$\|Gf\|_{L^{2,\infty}(w)} \lesssim_{G,d} \sqrt{[w]_{A_2}(1+\log[w]_{A_\infty})}\|f\|_{L^2(w)}.$$

Theorem 5.1 was already known, see for example [26], but here we give an elementary proof which uses the machinery developed in the previous chapters. Theorem 5.2 was obtained in [7]. Since the proofs are very similar, here we just prove Theorem 5.1. The proof follows the steps in [19], as developed in [7].

5.2 Proof

By the pointwise domination of Calderón-Zygmund operators proved in Chapter 3 it suffices to prove

$$\|\mathcal{A}_{\mathcal{S}}f\|_{L^{1,\infty}(w)} \lesssim_d [w]_{A_1}\big(1+\log[w]_{A_\infty}\big)\|f\|_{L^1(w)}, \tag{5.1}$$

where $\mathcal{S}$ is a sparse family of cubes, and f is nonnegative. See Chapter 4 for precise definitions of A_1 and A_∞.

After possibly splitting the family into several subfamilies, we can assume that $\mathcal{S}$ is $\frac{1}{4}$-sparse, that is:

$$\Big|\bigcup_{R\subsetneq Q, R\in\mathcal{S}} R\Big| \leq \frac{1}{4}|Q| \quad \forall Q\in\mathcal{S}.$$

Now, by homogeneity, it suffices to show

$$w\Big(x:\, \mathcal{A}_{\mathcal{S}}f(x) > 3\Big) \lesssim_d [w]_{A_1}\big(1+\log[w]_{A_\infty}\big),$$

for all nonnegative functions f with $\|f\|_{L^1(w)} = 1$, and for all weights w.

It will be convenient to split the family $\mathcal{S}$ into better-behaved subfamilies:

$$\mathcal{S}_m = \{Q \in \mathcal{S} :\ 2^{-m-1} < \langle f \rangle_Q \leq 2^{-m}\}$$

and

$$\mathcal{S}_- = \{Q \in \mathcal{S} :\ \langle f \rangle_Q > 1\}.$$

We have

$$w\Big(x :\ \mathcal{A}_{\mathcal{S}} f(x) > 3\Big) \leq w\Big(x :\ \mathcal{A}_{\mathcal{S}_-} f(x) > 1\Big) + w\Big(x :\ \mathcal{A}_{\mathcal{S}_+} f(x) > 2\Big),$$

where we have defined $\mathcal{S}_+$ to be the union of all the families $\mathcal{S}_m$ for $m \geq 0$.

We begin estimating the first term. Note that $\mathcal{A}_{\mathcal{S}_-} f$ is supported on $\{x : M^d f(x) > 1\}$, where M^d is the dyadic maximal function, so

$$w\Big(x :\ \mathcal{A}_{\mathcal{S}_-} f(x) > 1\Big) \leq \|M^d\|_{L^1(w) \to L^{1,\infty}(w)} \leq [w]_{A_1},$$

so this deals with the first summand.

The second summand will be split into two:

$$\begin{aligned} w\Big(x :\ \mathcal{A}_{\mathcal{S}_+} f(x) > 2\Big) &\leq w\Big(x :\ \sum_{m=0}^{m_0-1} \mathcal{A}_{\mathcal{S}_m} f(x) > 1\Big) + w\Big(x :\ \sum_{m=m_0}^{\infty} \mathcal{A}_{\mathcal{S}_m} f(x) > 1\Big) \\ &= I + II. \end{aligned}$$

The way we defined the subfamilies $\mathcal{S}_m$ gives us very good control of the averages of f. Indeed:

Lemma 5.3. *For each $m \in \mathbb{N}$ define*

$$E_m(Q) = Q \setminus \bigcup_{R \subsetneq Q,\, R \in \mathcal{S}_m} R,$$

then

$$\langle f \mathbb{1}_{E_m(Q)} \rangle_Q \sim \langle f \rangle_Q \tag{5.2}$$

for all $Q \in \mathcal{S}_m$.

Proof. Indeed: if we let $R_1, R_2, \ldots$ be the maximal subcubes of Q en $\mathcal{S}_m$ then

$$\begin{aligned}
\frac{1}{|Q|}\int_Q f \mathbb{1}_{E_m(Q)} &= \langle f \rangle_Q - \sum_i \frac{1}{|Q|}\int_{R_i} f \\
&> 2^{-m-1} - \sum_i \frac{|R_i|}{|Q|}\frac{1}{|R_i|}\int_{R_i} f \\
&\geq 2^{-m-1} - 2^{-m}\sum_i \frac{|R_i|}{|Q|} \\
&\geq 2^{-m-1} - 2^{-m}\frac{1}{4} \\
&\gtrsim \langle f \rangle_Q.
\end{aligned}$$

□

The reason this is useful is because the sets $\{E_m(Q)\}$ are pairwise disjoint when Q runs over $\mathcal{S}_m$. We can use this to deal with I:

$$\begin{aligned}
w\Big(x : \sum_{m=0}^{m_0-1} \mathcal{A}_{\mathcal{S}_m} f(x) > 1\Big) &\leq \int \sum_{m=0}^{m_0-1} \mathcal{A}_{\mathcal{S}_m} f(x)\, w(x)dx \\
&= \sum_{m=0}^{m_0-1} \int \sum_{Q\in\mathcal{S}_m} \langle f\rangle_Q \mathbb{1}_Q\, w(x)dx \\
&\lesssim \sum_{m=0}^{m_0-1} \sum_{Q\in\mathcal{S}_m} \langle f\mathbb{1}_{E_m(Q)}\rangle_Q \mathbb{1}_Q(x)\, w(x)dx \\
&= \sum_{m=0}^{m_0-1} \sum_{Q\in\mathcal{S}_m} \int f\mathbb{1}_{E_m(Q)} \frac{w(Q)}{|Q|} \\
&\leq [w]_{A_1} \sum_{m=0}^{m_0-1} \int fw \\
&= m_0[w]_{A_1}.
\end{aligned}$$

Finally, to estimate II, let $\{a_m\}_{m=m_0}^{\infty}$ be a sequence of nonnegative numbers such that

$$\sum_{m=m_0}^{\infty} a_m = 1.$$

Then

$$\begin{aligned}
II = w\Big(x : \sum_{m=m_0}^{\infty} \mathcal{A}_{\mathcal{S}_m} f(x) > \sum_{m=m_0}^{\infty} a_m\Big) \\
\leq \sum_{m=m_0}^{\infty} w\Big(x : \mathcal{A}_{\mathcal{S}_m} f(x) > a_m\Big) \\
\leq \sum_{m=m_0}^{\infty} w\Big(x : \sum_{Q\in\mathcal{S}_m} \langle f\rangle_Q \mathbb{1}_Q(x) > a_m\Big) \\
\leq \sum_{m=m_0}^{\infty} w\Big(x : \sum_{Q\in\mathcal{S}_m} \mathbb{1}_Q(x) > 2^m a_m\Big).
\end{aligned}$$

Call

$$b_m(x) := \sum_{Q\in\mathcal{S}_m} \mathbb{1}_Q(x).$$

Since $\mathcal{S}_m$ is sparse, this function is almost, but not quite, uniformly bounded; it is actually in BMO. In fact, for each m there exist a collection of maximal dyadic cubes $Q_1^m, Q_2^m, \cdots \in \mathcal{S}_m$ such that b_m is supported in the union of these cubes and

$$\frac{1}{|Q_i^m|}\Big|\Big\{x \in Q_i^m : b_m(x) > \lambda\Big\}\Big| \leq e^{-C\lambda}$$

for all $\lambda \geq 1$ and all $i \geq 1$.

Now we can use that every A_1 weights is also in A_∞ (see also Chapter 4) to obtain

$$\frac{w\Big(\Big\{x \in Q_i^m : b_m(x) > \lambda\Big\}\Big)}{w(Q_i^m)} \leq \exp\Big(-\frac{c}{[w]_{A_\infty}}\lambda\Big).$$

After summing in i, this yields

$$w\Big(\Big\{x : b_m(x) > \lambda\Big\}\Big) \leq \exp\Big(-\frac{c}{[w]_{A_\infty}}\lambda\Big)\sum_i w(Q_i^m).$$

Since all of the cubes Q_i^m are contained in the set

$$\{x : M^d f(x) > 2^{-m}\},$$

we can use the (weighted) weak-type boundedness of the maximal function to give the estimate

$$w\Big(\Big\{x : b_m(x) > \lambda\Big\}\Big) \leq [w]_{A_1} 2^m \exp\Big(-\frac{c}{[w]_{A_\infty}}\lambda\Big).$$

Now, plugging this estimate back, we have

$$II \leq [w]_{A_1} \sum_{m=m_0}^{\infty} 2^m \exp\Big(-\frac{c}{[w]_{A_\infty}} 2^m a_m\Big).$$

We should choose a_m so that it "looses" against 2^m (since we want exponential growth), while still summing to 1. A possible choice is

$$a_m = \xi^{-m}(1-\xi^{-1})\xi^{m_0}$$

for some $1 < \xi < 2$, like for example $3/2$.

Plugging this in the previous inequality, and estimating the sum by an integral we finally obtain

$$II \lesssim [w]_{A_1} \int_{m_0}^{\infty} 2^x \exp\Big(-\frac{c}{[w]_{A_\infty}} 2^x \xi^{-x}(1-\xi^{-1})\xi^{m_0}\Big)\, dx.$$

Calling $\eta = 2/\xi$ we see that $1 < \eta < 2$ and:

$$\begin{aligned} II &\lesssim [w]_{A_1} \int_0^{\infty} 2^x \exp\Big(-\frac{c'}{[w]_{A_\infty}} \eta^x \xi^{m_0}\Big)\, dx \\ &\lesssim [w]_{A_1} \int_0^{\infty} y^{\alpha} e^{-y}\, \frac{dy}{y} \\ &\lesssim [w]_{A_1} \end{aligned}$$

provided we choose $m_0 \sim \log[w]_{A_\infty}$, and where $\alpha > 0$.

With this choice of m_0 we obtain

$$w\Big(\mathcal{A}_{\mathcal{S}} f > 3\Big) \lesssim [w]_{A_1}(1 + \log[w]_{A_\infty}),$$

which is what we needed to prove the theorem.

BIBLIOGRAPHY

BIBLIOGRAPHY

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

[2] W. Chen and W. Damián. Weighted estimates for the multisublinear maximal function. *Rend. Circ. Mat. Palermo (2)*, 62(3):379–391, 2013.

[3] R. R. Coifman and C. Fefferman. Weighted norm inequalities for maximal functions and singular integrals. *Studia Math.*, 51:241–250, 1974.

[4] D. Cruz-Uribe, J. M. Martell, and C. Pérez. Sharp weighted estimates for approximating dyadic operators. *Electron. Res. Announc. Math. Sci.*, 17:12–19, 2010.

[5] D. Cruz-Uribe, J. M. Martell, and C. Pérez. Sharp weighted estimates for classical operators. *Adv. Math.*, 229(1):408–441, 2012.

[6] W. Damián, A. K. Lerner, and C. Pérez. Sharp weighted bounds for multilinear maximal functions and Calderón-Zygmund operators. *arXiv:1211.5115*, 2012.

[7] C. Domingo-Salazar, M. T. Lacey, and G. Rey. Borderline Weak Type Estimates for Singular Integrals and Square Functions. *ArXiv e-prints*, May 2015.

[8] J. Duoandikoetxea. *Fourier analysis*, volume 29 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001. Translated and revised from the 1995 Spanish original by David Cruz-Uribe.

[9] J. B. Garnett and P. W. Jones. BMO from dyadic BMO. *Pacific J. Math.*, 99(2):351–371, 1982.

[10] L. Grafakos and R. H. Torres. Multilinear Calderón-Zygmund theory. *Adv. Math.*, 165(1):124–164, 2002.

[11] T. S. Hänninen. Remark on dyadic pointwise domination and median oscillation decomposition. *ArXiv e-prints*, Feb. 2015.

[12] T. Hytönen. A_2 theorem: remarks and complements. *preprint*, 2012.

[13] T. Hytönen, C. Pérez, and E. Rela. Sharp reverse Hölder property for A_∞ weights on spaces of homogeneous type. *J. Funct. Anal.*, 263(12):3883–3899, 2012.

[14] T. P. Hytönen. The sharp weighted bound for general Calderón-Zygmund operators. *Ann. of Math. (2)*, 175(3):1473–1506, 2012.

[15] T. P. Hytönen, M. T. Lacey, and C. Pérez. Sharp weighted bounds for the q-variation of singular integrals. *Bull. Lond. Math. Soc.*, 45(3):529–540, 2013.

[16] M. T. Lacey. An elementary proof of the A_2 Bound. *ArXiv e-prints*, Jan. 2015.

[17] M. T. Lacey. Weighted Weak Type Estimates for Square Functions. *ArXiv e-prints*, Jan. 2015.

[18] M. T. Lacey, E. T. Sawyer, and I. Uriarte-Tuero. Two Weight Inequalities for Discrete Positive Operators. *ArXiv e-prints*, Nov. 2009.

[19] M. T. Lacey and J. Scurry. Weighted Weak Type Estimates for Square Functions. *ArXiv e-prints*, Nov. 2012.

[20] A. K. Lerner. A pointwise estimate for the local sharp maximal function with applications to singular integrals. *Bull. Lond. Math. Soc.*, 42(5):843–856, 2010.

[21] A. K. Lerner. A simple proof of the A_2 conjecture. *International Mathematics Research Notices*, 2012.

[22] A. K. Lerner. On an estimate of Calderón-Zygmund operators by dyadic positive operators. *J. Anal. Math.*, 121:141–161, 2013.

[23] A. K. Lerner. A simple proof of the A_2 conjecture. *Int. Math. Res. Not. IMRN*, 90(14):3159–3170, 2013.

[24] A. K. Lerner. On sharp aperture-weighted estimates for square functions. *J. Fourier Anal. Appl.*, 20(4):784–800, 2014.

[25] A. K. Lerner and F. Nazarov. personal communication.

[26] A. K. Lerner, S. Ombrosi, and C. Pérez. A_1 bounds for Calderón-Zygmund operators related to a problem of Muckenhoupt and Wheeden. *Math. Res. Lett.*, 16(1):149–156, 2009.

[27] A. K. Lerner, S. Ombrosi, C. Pérez, R. H. Torres, and R. Trujillo-González. New maximal functions and multiple weights for the multilinear Calderón-Zygmund theory. *Adv. Math.*, 220(4):1222–1264, 2009.

[28] K. Li, K. Moen, and W. Sun. The sharp weighted bound for multilinear maximal functions and Calderón-Zygmund operators. *J. Fourier Anal. Appl.*, 20(4):751–765, 2014.

[29] A. D. Melas. The Bellman functions of dyadic-like maximal operators and related inequalities. *Adv. Math.*, 192(2):310–340, 2005.

[30] A. D. Melas. A sharp L^p inequality for dyadic A_1 weights in $\mathbb{R}^n$. *Bull. London Math. Soc.*, 37(6):919–926, 2005.

[31] A. D. Melas. Dyadic-like maximal operators on $L \log L$ functions. *J. Funct. Anal.*, 257(6):1631–1654, 2009.

[32] A. D. Melas. Sharp general local estimates for dyadic-like maximal operators and related Bellman functions. *Adv. Math.*, 220(2):367–426, 2009.

[33] A. D. Melas and E. Nikolidakis. On weak type inequalities for dyadic maximal functions. *J. Math. Anal. Appl.*, 348(1):404–410, 2008.

[34] F. Nazarov, A. Reznikov, V. Vasyunin, and A. Volberg. A_1 conjecture: Weak norm estimates of weighted singular integral operators and Bellman functions. (Preprint), 2013.

[35] F. Nazarov, A. Reznikov, V. Vasyunin, and A. Volberg. A Bellman function counterexample to the A_1 conjecture: the blow-up of the weak norm estimates of weighted singular operators. *ArXiv e-prints*, June 2015.

[36] F. Nazarov, S. Treil, and A. Volberg. The Bellman functions and two-weight inequalities for Haar multipliers. *J. Amer. Math. Soc.*, 12(4):909–928, 1999.

[37] F. Nazarov, S. Treil, and A. Volberg. Bellman function in stochastic control and harmonic analysis. In *Systems, approximation, singular integral operators, and related topics (Bordeaux, 2000)*, volume 129 of *Oper. Theory Adv. Appl.*, pages 393–423. Birkhäuser, Basel, 2001.

[38] F. Nazarov and A. Volberg. Random walks and Bellman functions for weak estimate blow-ups with A_1 and A_2 weights. (Preprint), 2013.

[39] F. L. Nazarov and S. R. Treil. The hunt for a Bellman function: applications to estimates for singular integral operators and to other classical problems of harmonic analysis. *Algebra i Analiz*, 8(5):32–162, 1996.

[40] A. Osękowski. *Sharp martingale and semimartingale inequalities*, volume 72 of *Instytut Matematyczny Polskiej Akademii Nauk. Monografie Matematyczne (New Series) [Mathematics Institute of the Polish Academy of Sciences. Mathematical Monographs (New Series)]*. Birkhäuser/Springer Basel AG, Basel, 2012.

[41] A. Osękowski. Sharp inequalities for dyadic A_1 weights. *Arch. Math. (Basel)*, 101(2):181–190, 2013.

[42] S. Petermichl. The sharp bound for the Hilbert transform on weighted Lebesgue spaces in terms of the classical A_p characteristic. *Amer. J. Math.*, 129(5):1355–1375, 2007.

[43] G. Rey and A. Reznikov. Extremizers and sharp weak-type estimates for positive dyadic shifts. *Adv. Math.*, 254:664–681, 2014.

[44] A. Reznikov, V. Vasyuinin, and A. Volberg. Extremizer sequences and Bellman function for the weak type estimate of the martingale transform. (Preprint), 2013.

[45] L. Slavin, A. Stokolos, and V. Vasyunin. Monge-Ampère equations and Bellman functions: the dyadic maximal operator. *C. R. Math. Acad. Sci. Paris*, 346(9-10):585–588, 2008.

[46] V. Vasyunin and A. Volberg. Monge-Ampère equation and Bellman optimization of Carleson embedding theorems. In *Linear and complex analysis*, volume 226 of *Amer. Math. Soc. Transl. Ser. 2*, pages 195–238. Amer. Math. Soc., Providence, RI, 2009.

[47] V. I. Vasyunin. The exact constant in the inverse Hölder inequality for Muckenhoupt weights. *Algebra i Analiz*, 15(1):73–117, 2003.