Singular integral operators

Hytönen, Tuomas; van Neerven, Jan; Veraar, Mark; Weis, Lutz

DOI
10.1007/978-3-031-46598-7_1

Publication date
2023

Document Version
Final published version

Published in
Ergebnisse der Mathematik und ihrer Grenzgebiete

11

Singular integral operators

Various operators of Analysis, many of them already encountered in these volumes, take the generic form

$$Tf(s) = \int_{\mathbb{R}^d} K(s,t) f(t) \, \mathrm{d}t. \tag{11.1}$$

The mapping properties of T will of course heavily depend on the assumptions made on the *kernel* K that we will discuss in more detail in this chapter. A general feature of the different conditions is that the kernel is allowed to blow up on the 'diagonal' $\{(x,x) : x \in \mathbb{R}^d\}$, so that its natural domain of definition is the set

$$\dot{\mathbb{R}}^{2d} := \{(s,t) \in \mathbb{R}^d \times \mathbb{R}^d : s \neq t\}.$$

This blow-up is one of the reasons for referring to (11.1) as a *singular integral*; in general this formula requires a careful interpretation and will only be meaningful under restrictions on f and s.

In the prominent special case of a *convolution kernel* $K(s,t) = \mathfrak{K}(s-t)$, the operator (11.1) takes (at least formally, and under reasonable assumptions also rigorously) a simple representation "on the Fourier transform side":

$$\widehat{Tf}(\xi) = \widehat{\mathfrak{K} * f}(\xi) = \widehat{\mathfrak{K}}(\xi)\widehat{f}(\xi) =: m(\xi)\widehat{f}(\xi);$$

thus $T = T_m$ can be identified with a *Fourier multiplier*; they have been studied extensively in Chapter 5 and Section 8.3.

The motivations to investigate *singular integral operators* in the non-transformed representation (11.1) are at least threefold. First, it allows for a wider class of examples beyond those of the convolution form. Second, even when the alternative Fourier multiplier representation is available in principle, an operator may naturally arise in the form (11.1), and identifying or estimating the corresponding multiplier explicitly may not be feasible in practise, as the Fourier transform is not isomorphic between the natural function spaces for the kernel $\mathfrak{K}$ and the multiplier m. Finally, and perhaps most importantly,

T. Hytönen et al., *Analysis in Banach Spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge / A Series of Modern Surveys in Mathematics 76,
https://doi.org/10.1007/978-3-031-46598-7_1

even for multiplier operators, the point-of-view of singular integrals gives us access to new methods and conclusions.

An overarching theme is of this chapter is *extrapolation*: As soon as an operator (11.1), with natural assumptions on the kernel K, is bounded on a single space $L^{p_0}(\mathbb{R}^d;X)$, it will be automatically bounded on several more spaces, including $L^p(\mathbb{R}^d;X)$ for other exponents $p\in(1,\infty)$ (with certain substitute results at the end-points $p\in\{1,\infty\}$), and even their weighted versions $L^p(w;X)$, where w is an arbitrary weight in the Muckenhoupt class A_p (see Appendix J). These results will be used to deduce analogous extrapolation results for *maximal* L^p*-regularity* of the *abstract Cauchy problem* in Chapter 17.

In terms of Banach spaces, this chapter deals with relatively general results, most of which are valid without restrictions of the class of admissible spaces. Such restrictions, and notably the ubiquitous UMD condition, will reappear in the subsequent chapters, when searching for conditions to verify the boundedness of (11.1) on just one $L^{p_0}(\mathbb{R}^d;X)$, to serve as an input to the extrapolation results that we develop in the chapter at hand.

11.1 Local oscillations of functions

A characteristic feature of singular integrals, the main topic of this chapter, is that their boundedness properties depend not only naive size estimates but on rather delicate cancellations between different oscillatory components. Before we dwell into a deeper study of there operators, we dedicate this section to a general treatment of oscillations of functions per se; this will streamline the subsequent discussion, where the results of this section will be put into action in the context of operator norm estimates.

Given $f\in L^0(\mathbb{R}^d;X)$ and $\lambda>0$, we define the following measure of oscillation of f on a cube Q,

$$\operatorname{osc}_\lambda(f;Q):=\inf_{c\in X}\inf_{|E|\leqslant\lambda|Q|}\|(f-c)\mathbf{1}_{Q\setminus E}\|_\infty.$$

Here, and in many occasions below where we will use the same notation, it is understood that the supremum is taken over all measurable subsets E of Q satisfying the stated requirement that $|E|\leqslant\lambda|Q|$. The idea is to quantify how much f deviates from a constant, if we ignore its (possibly wild) behaviour on an exceptional set of controlled proportion. The above way of measuring oscillations is essentially 'minimal' in that it can be controlled by average L^q oscillations for any $q>0$:

Lemma 11.1.1. *For any* $q\in(0,\infty)$*, we have*

$$\operatorname{osc}_\lambda(f;Q)\leqslant\inf_{c\in X}\frac{\|(f-c)\mathbf{1}_Q\|_{L^{q,\infty}}}{(\lambda|Q|)^{1/q}}.$$

Proof. For a fixed c, let $g := (f-c)\mathbf{1}_Q$. If we choose $t := \|g\|_{L^{q,\infty}}/(\lambda|Q|)^{1/q}$, then

$$|E_t| := |\{\|g\| > t\}| \leqslant \frac{\|g\|_{L^{q,\infty}}^q}{t^q} = \lambda|Q|$$

But then it is clear that

$$\inf_{|E|\leqslant\lambda|Q|} \|g\mathbf{1}_{Q\setminus E}\|_\infty \leqslant \|g\mathbf{1}_{Q\setminus E_t}\|_\infty \leqslant t,$$

which is precisely the claimed bound. □

Given a *real-valued* $f \in L^0(\mathbb{R}^d;\mathbb{R})$, any $m \in \mathbb{R}$ such that

$$|Q \cap \{f \leqslant m\}| \geqslant \tfrac{1}{2}|Q|, \qquad |Q \cap \{f \geqslant m\}| \geqslant \tfrac{1}{2}|Q|$$

is called a *median* of f on the cube (or more general set of finite positive measure) $Q \subseteq \mathbb{R}^d$. One routinely checks that a median always exists but may fail to be unique.

Lemma 11.1.2. *If $\lambda \in (0,\frac{1}{2})$ and $m_f \in \mathbb{R}$ is a median of $f \in L^0(Q;\mathbb{R})$ on Q, then*

$$\inf_{|E|\leqslant\lambda|Q|} \|(f-m_f)\mathbf{1}_{Q\setminus E}\|_\infty \leqslant 2\operatorname{osc}_\lambda(f;Q).$$

Proof. Let $c \in \mathbb{R}$ be arbitrary. Then $f - m_f = f - c - (m_f - c)$ and hence

$$\inf_{|E|\leqslant\lambda|Q|} \|(f-m_f)1_{Q\setminus E}\|_\infty \leqslant \inf_{|E|\leqslant\lambda|Q|} \|(f-c)1_{Q\setminus E}\|_\infty + |m_f - c|.$$

Note that $m_f - c$ is a median of $g := f - c$ on Q. Hence it suffices to check that the median m_g always satisfies

$$|m_g| \leqslant \|g1_{Q\setminus E}\|_\infty$$

whenever $|E| \leqslant \lambda|Q|$ and $\lambda < \frac{1}{2}$. If $m_g \geqslant 0$, then

$$|Q\cap\{|g|\geqslant|m_g|\}\setminus E| \geqslant |Q\cap\{g\geqslant m_g\}\setminus E| \geqslant \tfrac{1}{2}|Q| - |E| \geqslant (\tfrac{1}{2}-\lambda)|Q| > 0$$

and thus $\|g\mathbf{1}_{Q\setminus E}\|_\infty \geqslant |m_g|$. If $m_g < 0$, the argument is the same, just replacing the second step above by $|Q\cap\{g\leqslant m_g\}\setminus E|$. □

The previous lemma motivates the following:

Definition 11.1.3. *Let X be a Banach space and $f \in L^0(Q;X)$. A vector $m \in X$ is called a λ-pseudomedian of f on Q if*

$$\inf_{|E|\leqslant\lambda|Q|} \|(f-m)\mathbf{1}_{Q\setminus E}\|_\infty \leqslant 2\operatorname{osc}_\lambda(f;Q).$$

Indeed, Lemma 11.1.2 says that the usual median is a λ-pseudomedian for every $\lambda \in (0,\frac{1}{2})$. Concerning existence in the general case, we have:

Lemma 11.1.4. *Let X be a Banach space, $f \in L^0(Q;X)$ and $\lambda \in (0,\frac{1}{2})$. Then f has a λ-pseudomedian on Q.*

Proof. If $\operatorname{osc}_\lambda(f;Q) > 0$, this is obvious, since we can always come within any positive distance from the infimum. So only the case $\operatorname{osc}_\lambda(f;Q) = 0$ needs attention. In this case, there we can find a sequence of vectors $c_n \in X$ and sets $E_n \subseteq Q$ with $|E_n| \leqslant \lambda|Q|$ such that $\|(f - c_n)\mathbf{1}_{Q\setminus E_n}\|_\infty \to 0$. Since $|E_n \cup E_m| \leqslant 2\lambda|Q| < |Q|$, any $Q \setminus (E_n \cup E_m)$ has positive measure, and thus

$$\begin{aligned}\|c_n - c_m\| &= \|(c_n - c_m)\mathbf{1}_{Q\setminus(E_n\cup E_m)}\|_\infty \\ &\leqslant \|(f - c_n)\mathbf{1}_{Q\setminus E_n}\|_\infty + \|(f - c_m)\mathbf{1}_{Q\setminus E_m}\|_\infty \to 0.\end{aligned}$$

Thus $(c_n)_{n\geqslant 1}$ is a Cauchy sequence and hence convergent to some $c \in X$. But then

$$\begin{aligned}\inf_{|E|\leqslant\lambda|Q|} \|(f - c)\mathbf{1}_{Q\setminus E}\|_\infty &\leqslant \liminf_{n\to\infty} \|(f - c)\mathbf{1}_{Q\setminus E_n}\|_\infty \\ &\leqslant \liminf_{n\to\infty} \Big(\|(f - c_n)\mathbf{1}_{Q\setminus E_n}\|_\infty + \|c_n - c\|\Big) = 0,\end{aligned}$$

and thus this limit c is a λ-pseudomedian. □

Lemma 11.1.5. *Let X be a Banach space, let $f \in L^0(\mathbb{R}^d;X)$ and $\lambda \in (0,\frac{1}{2})$, and let $m_f(Q)$ be a λ-pseudomedian of f on Q. Then*

$$E^0 := Q \cap \{\|f - m_f(Q)\| > 2\operatorname{osc}_\lambda(f;Q)\}$$

satisfies $|E^0| \leqslant \lambda|Q|$.

Proof. Suppose for contradiction that $|E^0| > \lambda|Q|$. Denoting

$$E^\varepsilon := Q \cap \{\|f - m_f(Q)\| > 2\operatorname{osc}_\lambda(f;Q) + \varepsilon\}$$

we have $E^0 = \bigcup_{n=1}^\infty E^{1/n}$, so that by continuity of measure, we also have $|E^\varepsilon| > \lambda|Q|$ for some $\varepsilon = 1/n > 0$.

Let $|E| \leqslant \lambda|Q|$. Then

$$\|(f - m_f(Q))\mathbf{1}_{Q\setminus E}\|_\infty \geqslant (2\operatorname{osc}_\lambda(f;Q) + \varepsilon)\|\mathbf{1}_{E^\varepsilon\setminus E}\|_\infty = 2\operatorname{osc}_\lambda(f;Q) + \varepsilon,$$

since $|E^\varepsilon \setminus E| \geqslant |E^\varepsilon| - |E| > \lambda|Q| - \lambda|Q| = 0$. Taking the infimum over all $|E| \leqslant \lambda|Q|$, we contradict the definition of a λ-pseudomedian. □

11.1.a Sparse collections and Lerner's formula

Let us recall and expand the terminology related to dyadic cubes that we introduced in Chapter 3.

Definition 11.1.6. *A dyadic system of cubes on $\mathbb{R}^d$ is a collection $\mathscr{D} = \bigcup_{j\in\mathbb{Z}} \mathscr{D}_j$, where*

(i) *each* $\mathscr{D}_j$ *is a partition of* $\mathbb{R}^d$ *of the form*

$$\mathscr{D}_j = \Big\{ s_j + 2^{-j}(m + [0,1)^d) : m \in \mathbb{Z}^d \Big\},$$

(ii) *each* $\mathscr{D}_{j+1}$ *refines the previous* $\mathscr{D}_j$*.*

When $s_j = 0$ *for all* $j \in \mathbb{Z}$*, we refer to the corresponding* $\mathscr{D}$ *as the* standard dyadic system*, and denote it by* $\mathscr{D}^0$*.*

Remark 11.1.7. One might like to replace (i) in Definition 11.1.6 by the "more intrinsic"

(iii) each $\mathscr{D}_j$ is a partition of $\mathbb{R}^d$ consisting of left-closed, right-open cubes of side-length 2^{-j}.

When $d = 1$, one can check that (i) and (iii) are equivalent. But, for $d > 1$, condition (iii) is strictly more general. For instance

$$\mathscr{D}_j := \big\{ 2^{-j}\big(m + [0,1)^2\big) + (0, \alpha \mathbf{1}_{[0,\infty)}(m_1)) : m \in \mathbb{Z}^2 \big\}, \quad \alpha \in \mathbb{R},$$

where all cubes in the right half-plane are shifted in the y-direction by a fixed amount $\alpha \in \mathbb{R}$ relative to the standard dyadic cubes, would qualify for (iii) but not for (i). The preference over one or the other definition may be a question of taste; we choose to work with Definition 11.1.6 as stated.

We will work be working with an arbitrary dyadic system as in Definition 11.1.6. For many purposes, the reader who so wishes may think of the *standard dyadic system*.

$$\mathscr{D}^0 := \bigcup_{j \in \mathbb{Z}} \mathscr{D}^0_j, \qquad \mathscr{D}^0_j := \{ 2^{-j}([0,1)^d + k) : k \in \mathbb{Z}^d \}, \quad j \in \mathbb{Z},$$

but here and there we will also make use of other systems, which makes it convenient to deal with a generic system from the beginning. For any given cube, we may speak of its *dyadic subcubes*, by which we understand all cubes obtained by repeatedly bisecting the edges of Q. We will use the notation $\mathscr{D}(Q)$ for the collection of all dyadic subcubes of a cube Q. If Q belongs to a dyadic system $\mathscr{D}$, then

$$\mathscr{D}(Q) = \{ Q' \in \mathscr{D} : Q' \subseteq Q \}.$$

Definition 11.1.8. *A* quadrant *of a dyadic system* $\mathscr{D}$ *of* $\mathbb{R}^d$ *is the union of any strictly increasing sequence* $Q_1 \subsetneq Q_2 \subsetneq Q_3 \subsetneq \ldots$ *of cubes* $Q_i \in \mathscr{D}$*.*

Remark 11.1.9. The standard dyadic system $\mathscr{D}^0$ has 2^d quadrants of the form $S_1 \times \cdots \times S_d$, where $S_i \in \{(-\infty, 0), [0, infty)\}$ for each $i \in \{1, \ldots, d\}$. It is also easy to construct dyadic systems, where $\mathbb{R}^d$ is the only quadrant.

The *dyadic Hardy–Littlewood maximal function* is defined by

$$M_{\mathscr{D}}f(x) = \sup_{Q\in\mathscr{D}:x\in Q} \langle \|f\| \rangle_Q, \qquad \langle f\rangle_Q := \fint_Q f := \frac{1}{|Q|}\int_Q f,$$

where the supremum is taken over all dyadic cubes containing x. Here, and throughout this chapter, unless indicated otherwise, integrals are taken with respect to Lebesgue measure and are abbreviated in the above way to unburden notation. Thus, when g is an integrable function, $\int_Q g$ is shorthand for $\int_Q g(x)\,\mathrm{d}x$. When integrating over all of $\mathbb{R}^d$ we will even write $\int g$ for $\int_{\mathbb{R}^d} g$.

Definition 11.1.10. *A collection $\mathscr{S}$ of sets $S \subseteq \mathbb{R}^d$ of finite measure is called γ-sparse, if each $S\in\mathscr{S}$ has a measurable subset $E(S)\subseteq S$ of size $|E(S)|\geqslant \gamma|S|$ such that the sets $E(S)$ are pairwise disjoint.*

While the definition can be made for general measurable sets, we will be mostly concerned with the case when $\mathscr{S}\subseteq\mathscr{D}$ is a subcollection of the dyadic cubes of $\mathbb{R}^d$.

A disjoint collection is obviously 1-sparse with $E(S) = S$. The usefulness of general γ-sparse collections comes from the fact that, on the one hand, they are easier to create than genuinely disjoint collections while, on the other hand, for the purposes of L^p estimates they are essentially as good as disjoint ones. This is quantified by the following:

Proposition 11.1.11. *Let $\mathscr{S}\subseteq\mathscr{D}$ be a γ-sparse collection of dyadic cubes S with disjoint subsets $|E(S)|\geqslant\gamma|S|$.*

(1) *If $a_S\geqslant 0$, then for all $p\in(0,\infty)$,*

$$\Big\|\sum_{S\in\mathscr{S}} a_S\mathbf{1}_S\Big\|_p \leqslant c_{p,\gamma}\Big\|\sum_{S\in\mathscr{S}} a_S\mathbf{1}_{E(S)}\Big\|_p, \quad \textit{where } c_{p,\gamma} = \begin{cases}\gamma^{-1}p, & p\in[1,\infty),\\ \gamma^{-1/p}, & p\in(0,1).\end{cases}$$

(2) *If $f\geqslant 0$, then for all $p\in(1,\infty)$,*

$$\Big(\sum_{S\in\mathscr{S}}\langle f\rangle_S^p|S|\Big)^{1/p} \leqslant \gamma^{-1/p}\,p'\|f\|_p.$$

Proof of Proposition 11.1.11. If $p\in[1,\infty)$, we dualise the left side of (1) against $\phi\in L^{p'}$:

$$\begin{aligned}\int\Big(\sum_{S\in\mathscr{S}}a_S\mathbf{1}_S\Big)\phi &= \sum_{S\in\mathscr{S}}a_S|S|\fint_S\phi \leqslant \frac{1}{\gamma}\sum_{S\in\mathscr{S}}a_S|E(S)|\inf_S M_{\mathscr{D}}\phi\\ &\leqslant \frac{1}{\gamma}\int\Big(\sum_{S\in\mathscr{S}}a_S\mathbf{1}_{E(S)}\Big)M_{\mathscr{D}}\phi \leqslant \frac{1}{\gamma}\Big\|\sum_{S\in\mathscr{S}}a_S\mathbf{1}_{E(S)}\Big\|_p\|M_{\mathscr{D}}\phi\|_{p'},\end{aligned}$$

where $\|M_{\mathscr{D}}\phi\|_{p'}\leqslant p\|\phi\|_{p'}$ by Doob's maximal inequality (Theorem 3.2.2; cf. the explanations preceding Theorem 3.2.27).

If $p \in (0,1)$, then the left side of (1) can be estimated by

$$\int \Big(\sum_{S \in \mathscr{S}} a_S \mathbf{1}_S \Big)^p \leqslant \int \sum_{S \in \mathscr{S}} a_S^p \mathbf{1}_S = \sum_{S \in \mathscr{S}} a_S^p |S| \leqslant \frac{1}{\gamma} \sum_{S \in \mathscr{S}} a_S^p |E(S)|$$
$$= \frac{1}{\gamma} \int \sum_{S \in \mathscr{S}} a_S^p \mathbf{1}_{E(S)} = \frac{1}{\gamma} \int \Big(\sum_{S \in \mathscr{S}} a_S \mathbf{1}_{E(S)} \Big)^p,$$

and taking the pth root completes the proof of (1).

For (2), we use $\langle f \rangle_S \leqslant \inf_{z \in S} M_{\mathscr{D}} f(z)$ and $|S| \leqslant \gamma^{-1} |E(S)|$ to find that

$$\sum_{S \in \mathscr{S}} \langle f \rangle_S^p |S| \leqslant \frac{1}{\gamma} \sum_{S \in \mathscr{S}} \int_{E(S)} (M_{\mathscr{D}} f)^p \, \mathrm{d}x \leqslant \frac{1}{\gamma} \| M_{\mathscr{D}} f \|_p^p \leqslant \frac{1}{\gamma} (p')^p \| f \|_p^p,$$

again by Doob's inequality in the last step. □

The different notions introduced above come together in the following useful estimate, which is the main result of this section:

Theorem 11.1.12 (Lerner's formula). *Let X be a Banach space, $Q^0 \subseteq \mathbb{R}^d$ be a cube and $f \in L^0(Q^0; X)$. Then there is a $\frac{1}{2}$-sparse subcollection $\mathscr{S} \subseteq \mathscr{D}(Q^0)$ such that, almost everywhere,*

$$\mathbf{1}_{Q^0} \| f - m_f(Q^0) \| \leqslant 4 \sum_{S \in \mathscr{S}} \operatorname{osc}_\lambda(f; S) \mathbf{1}_S, \qquad \lambda = 2^{-2-d},$$

where $m_f(Q^0)$ is any λ-pseudomedian of f on Q^0.

By Lemma 11.1.2, if $X = \mathbb{R}$, we can take $m_f(Q^0)$ to be a usual median of f.

Proof. We begin with a preliminary observation. For any collection of disjoint sets $Q_j \in \mathscr{D}(Q^0)$, we have the identity

$$\begin{aligned} \mathbf{1}_{Q^0}(f - m_f(Q^0)) &= \mathbf{1}_{Q^0 \setminus \bigcup_j Q_j}(f - m_f(Q^0)) \\ &\quad + \sum_j \mathbf{1}_{Q_j}(m_f(Q_j) - m_f(Q^0)) \\ &\quad + \sum_j \mathbf{1}_{Q_j}(f - m_f(Q_j)). \end{aligned} \tag{11.2}$$

Turning to the actual proof, let

$$E^0 := Q^0 \cap \Big\{ \| f - m_f(Q^0) \| > 2 \operatorname{osc}_\lambda(f; Q^0) \Big\}$$

so that $|E^0| \leqslant \lambda |Q^0|$ by Lemma 11.1.5.

For $\alpha \in (0,1)$ to be chosen, let Q_j^1 be the maximal cubes in $\mathscr{D}(Q^0)$ such that $|Q_j^1 \cap E^0| > \alpha |Q_j^1|$. Since any two dyadic cubes are either disjoint, or one is contained in the other, dyadic cubes that are maximal with respect to some property are necessarily disjoint; hence our preliminary observation applies to $Q_j = Q_j^1$. Moreover, by definition of the dyadic maximal operator, we have $M_{\mathscr{D}}\mathbf{1}_{E^0}(x) > \alpha$, if and only if x is contained in some dyadic Q with $\langle \mathbf{1}_{E^0}\rangle_Q > \alpha$, if and only if it is contained in a maximal dyadic cube with this property. Hence

$$\bigcup_j Q_j^1 = \{M_{\mathscr{D}}\mathbf{1}_{E^0} > \alpha\},$$

so that by Doob's inequality

$$\sum_j |Q_j^1| \leqslant \frac{1}{\alpha}\|\mathbf{1}_{E^0}\|_1 \leqslant \frac{\lambda}{\alpha}|Q^0|.$$

By Lebesgue's differentiation theorem, almost every point of E is contained in some Q_j^1, and hence

$$\mathbf{1}_{Q^0\setminus\bigcup_j Q_j^1}\|f - m_f(Q^0)\| \leqslant \mathbf{1}_{Q^0\setminus\bigcup_j Q_j^1} 2\operatorname{osc}_\lambda(f;Q^0)$$

almost everywhere.

By the maximality of the Q_j^1, their parent cubes $\widehat{Q}_j^1$ satisfy the opposite bound $|\widehat{Q}_j^1 \cap E| \leqslant \alpha|\widehat{Q}_j^1|$. Hence in particular

$$|Q_j^1 \cap E^0| \leqslant |\widehat{Q}_j^1 \cap E^0| \leqslant \alpha|\widehat{Q}_j^1| = 2^d\alpha|Q_j^1|.$$

Let also

$$E_j^1 = Q_j^1 \cap \Big\{\|f - m_f(Q_j^1)\| > 2\operatorname{osc}_\lambda(f;Q_j^1)\Big\}$$

so that $|E_j^1| \leqslant \lambda|Q_j^1|$ by Lemma 11.1.5. It follows that

$$|Q_j^1 \cap (E^0 \cup E_j^1)| \leqslant (2^d\alpha + \lambda)|Q_j^1|.$$

If $2^d\alpha + \lambda < 1$, then $Q_j^1 \setminus (E^0 \cup E_j^1)$ has positive measure, and for any x in this set, we have both

$$\|f(x) - m_f(Q^0)\| \leqslant 2\operatorname{osc}_\lambda(f;Q^0), \quad \|f(x) - m_f(Q_j^1)\| \leqslant 2\operatorname{osc}_\lambda(f;Q_j^1).$$

Since such points x exist, it follows in particular that

$$\|m_f(Q_j^1) - m_f(Q^0)\| \leqslant 2\operatorname{osc}_\lambda(f;Q^0) + 2\operatorname{osc}_\lambda(f;Q_j^1).$$

Substituting this to (11.2), we have

$$
\begin{aligned}
\mathbf{1}_{Q^0}\|f-m_f(Q^0)\| \leqslant \mathbf{1}_{Q^0\setminus\bigcup_j Q_j^1} 2\operatorname{osc}_\lambda(f;Q^0) \\
+\sum_j \mathbf{1}_{Q_j^1}\Big(2\operatorname{osc}_\lambda(f;Q^0)+2\operatorname{osc}_\lambda(f;Q^j)\Big) \\
+\sum_j \mathbf{1}_{Q_j^1}\|f-m_f(Q_j^1)\| \\
= \mathbf{1}_{Q^0} 2\operatorname{osc}_\lambda(f;Q^0)+\sum_j \mathbf{1}_{Q_j^1} 2\operatorname{osc}_\lambda(f;Q_j^1) \\
+\sum_j \mathbf{1}_{Q_j^1}\|f-m_f(Q_j^1)\|,
\end{aligned}
\tag{11.3}
$$

where each term in the last sum has exactly the same form as the left hand side and allows to iterate the same consideration.

Assuming that we have proved

$$
\begin{aligned}
\mathbf{1}_{Q^0}\|f-m_f(Q^0)\| \leqslant 4\sum_{n=0}^{N-1}\sum_j \mathbf{1}_{Q_j^n}\operatorname{osc}_\lambda(f;Q_j^n)+2\sum_j \mathbf{1}_{Q_j^N}\operatorname{osc}_\lambda(f;Q_j^N) \\
+\sum_j \mathbf{1}_{Q_j^N}\|f-m_f(Q_j^N)\|,
\end{aligned}
$$

where each Q_j^n is contained in some Q_i^{n-1} and

$$
\sum_{j:Q_j^n\subseteq Q_i^{n-1}} |Q_j^n| \leqslant \frac{\lambda}{\alpha}|Q_i^{n-1}|, \tag{11.4}
$$

applying (11.3) to each Q_j^N in place of Q^0 yields the analogue of the previous display with $N+1$ in place of N.

The support of the final error term has measure at most $\sum_j |Q_j^N| \leqslant (\lambda/\alpha)^N|Q^0|$, so if $\lambda/\alpha < 1$, this error term tends to zero pointwise almost everywhere. Hence, in the limit, we have

$$
\mathbf{1}_{Q^0}\|f-m_f(Q^0)\| \leqslant 4\sum_{n=0}^{\infty}\sum_j \mathbf{1}_{Q_j^n}\operatorname{osc}_\lambda(f;Q_j^n).
$$

Choosing $\alpha = 2\lambda$, (11.4) shows that the collection $\{Q_j^n\}_{n,j}$ is $\frac{1}{2}$-sparse, and with $\lambda = 2^{-2-d}$, we also have $2^d\alpha+\lambda = (2^{d+1}+1)\lambda = 2^{-1}+2^{-1-d} < 1$, as required. This concludes the proof. □

11.1.b Almost orthogonality in L^p

In a Hilbert space H such as $H = L^2(\mathbb{R}^d)$, orthogonality of elements h_i implies the fundamental Pythagorean identity

$$\Big\| \sum_i h_i \Big\|_H = \Big(\sum_i \|h_i\|_H^2 \Big)^{1/2}.$$

As we have seen in the previous Volumes, L^p analogues of this identity tend to either take the form of a one-sided estimate only, or, insisting in a two-sided equivalence, require the introduction of some randomised norm. In contrast to this, it may come as a surprise that sparse collections lead to relatively simple constructions that allow almost complete L^p analogues of the Pythagorean identity in certain situations.

We introduce some additional notation. The following definition is meaningful for any subcollection $\mathscr{S} \subseteq \mathscr{D}$ of the dyadic cubes, but it will prove itself particularly useful when $\mathscr{S}$ is sparse.

Definition 11.1.13. *For any subcollection $\mathscr{S} \subseteq \mathscr{D}$ of dyadic cubes, we have the following notions:*

(1) *For each $S \in \mathscr{S}$, let $\mathrm{ch}_{\mathscr{S}}(S) \subseteq \mathscr{S}$ (the $\mathscr{S}$-children of S) denote the collection of all maximal $S' \in \mathscr{S}$ such that $S' \subsetneq S$.*
(2) *For each $S \in \mathscr{S}$, let $E_{\mathscr{S}}(S) := S \setminus \bigcup_{S' \in \mathrm{ch}_{\mathscr{S}}(S)} S'$.*
(3) *For each $Q \in \mathscr{D}$, let $\pi_{\mathscr{S}}(Q)$ denote the minimal $S \in \mathscr{S}$ such that $S \supseteq Q$.*

When $\mathscr{S} = \mathscr{D}$, we reproduce the familiar notion $\mathrm{ch}_{\mathscr{D}} = \mathrm{ch}$ of dyadic children. The other two notions above are uninteresting in this special case, as we simply have $E_{\mathscr{D}}(Q) = \varnothing$ and $\pi_{\mathscr{D}}(Q) = Q$ for all $Q \in \mathscr{D}$.

We begin with a one-sided estimate:

Proposition 11.1.14. *Let X be a Banach space and $p \in [1, \infty)$. Let $\mathscr{S} \subseteq \mathscr{D}$ be a γ-sparse collection of dyadic cubes. For each $S \in \mathscr{S}$, let $f_S \in L^p(\mathbb{R}^d; X)$ be a function supported on S and constant on each $S' \in \mathrm{ch}_{\mathscr{S}}(S)$. Then*

$$\Big\| \sum_{S \in \mathscr{S}} f_S \Big\|_{L^p(\mathbb{R}^d;X)} \leqslant (1 + \gamma^{-1/p'} p) \Big(\sum_{S \in \mathscr{S}} \|f_S\|_{L^p(\mathbb{R}^d;X)}^p \Big)^{1/p}.$$

Proof. We assume that the right-hand side is finite, for otherwise there is nothing to prove. We then assume without loss of generality that $\mathscr{S}$ is finite. In fact, once we have proved the result for finite families, in the infinite case it follows easily that the partial sums of the series $\sum_{S \in \mathscr{S}} f_S$ (with arbitrary enumeration) form a Cauchy sequence in $L^p(\mathbb{R}^d; X)$, from which we deduce the (unconditional) convergence of this series and the asserted norm bound.

Concentrating on the finite case, by dualising with $g \in L^{p'}(\mathbb{R}^d; X^*)$, it is equivalent to the estimate

$$\int \sum_S \langle f_S, g \rangle \,\mathrm{d}x \leqslant (1 + \gamma^{-1/p'} p) \Big(\sum_S \|f_S\|_{L^p(\mathbb{R}^d;X)}^p \Big)^{1/p} \|g\|_{L^{p'}(\mathbb{R}^d;X^*)}.$$

Since f_S is supported on S and constant on each $S' \in \mathrm{ch}_{\mathscr{S}}(S)$, and since S is partitioned by $\mathrm{ch}_{\mathscr{S}}(S) \cup \{E_{\mathscr{S}}(S)\}$, we have

$$\begin{aligned}\int \sum_{S\in\mathscr{S}} \langle f_S, g\rangle \,\mathrm{d}x &= \sum_{S\in\mathscr{S}} \int_S \langle f_S, g\rangle \,\mathrm{d}x \\ &= \sum_{S\in\mathscr{S}} \sum_{S'\in \mathrm{ch}_{\mathscr{S}}(S)} \langle \langle f_S\rangle_{S'}, \langle g\rangle_{S'}\rangle |S'| + \int \sum_{S\in\mathscr{S}} \mathbf{1}_{E_{\mathscr{S}}(S)} \langle f_S, g\rangle \,\mathrm{d}x.\end{aligned}$$

We can estimate the second term by Hölder's inequality and the pairwise disjointness of the sets $E_{\mathscr{S}}(S)$,

$$\begin{aligned}\Big|\int \sum_{S\in\mathscr{S}} \mathbf{1}_{E_{\mathscr{S}}(S)} \langle f_S, g\rangle \,\mathrm{d}x\Big| &\leqslant \Big\| \sum_{S\in\mathscr{S}} \mathbf{1}_{E_{\mathscr{S}}(S)} f_S \Big\|_{L^p(\mathbb{R}^d;X)} \|g\|_{L^{p'}(\mathbb{R}^d;X^*)} \\ &= \Big(\sum_{S\in\mathscr{S}} \|\mathbf{1}_{E_{\mathscr{S}}(S)} f_S\|^p_{L^p(\mathbb{R}^d;X)} \Big)^{1/p} \|g\|_{L^{p'}(\mathbb{R}^d;X^*)}.\end{aligned}$$

For the first term we argue as follows.

$$\begin{aligned}&\Big| \sum_{S\in\mathscr{S}} \sum_{S'\in \mathrm{ch}_{\mathscr{S}}(S)} \langle \langle f_S\rangle_{S'}, \langle g\rangle_{S'}\rangle |S'| \Big| \\ &\leqslant \Big(\sum_{S\in\mathscr{S}} \sum_{S'\in \mathrm{ch}_{\mathscr{S}}(S)} \|\langle f_S\rangle_{S'}\|_X^p |S'| \Big)^{1/p} \Big(\sum_{S\in\mathscr{S}} \sum_{S'\in \mathrm{ch}_{\mathscr{S}}(S)} \|\langle g\rangle_{S'}\|_{X^*}^{p'} |S'| \Big)^{1/p'} \\ &\leqslant \Big(\sum_{S\in\mathscr{S}} \|f_S\|^p_{L^p(\mathbb{R}^d;X)} \Big)^{1/p} \Big(\sum_{S'\in\mathscr{S}} \langle \|g\|_{X^*}\rangle_{S'}^{p'} |S'| \Big)^{1/p'},\end{aligned}$$

where in the second factor we rearranged the double sum into a single sum, observing that every $S' \in \mathscr{S}$ is counted at most once as a child of a unique $S \in \mathscr{S}$. The second factor is bounded by $\gamma^{-1/p'} p\|g\|_{p'}$ thanks to Proposition 11.1.11(2). Summing up the bounds, we complete the proof of the direct estimate. □

The following lemma describes useful projections and also provides prominent examples of the functions f_S featuring in Proposition 11.1.14.

Lemma 11.1.15. *For $S \in \mathscr{S} \subseteq \mathscr{D}$ and $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d; X)$, let*

$$P_S f := \sum_{S'\in \mathrm{ch}_{\mathscr{S}}(S)} \mathbb{E}_{S'} f + \mathbf{1}_{E_{\mathscr{S}}(S)} f. \tag{11.5}$$

Then $\langle f\rangle_Q = \langle P_S f\rangle_Q$ for all $Q \in \mathscr{D}$ such that $\pi_{\mathscr{S}}(Q) = S$.

Proof. From definition, we have

$$\langle P_S f\rangle_Q = \sum_{S'\in \mathrm{ch}_{\mathscr{S}}} \langle f\rangle_{S'} \frac{|S'\cap Q|}{|Q|} + \frac{1}{|Q|} \int_{Q\cap E_{\mathscr{S}}(S)} f \,\mathrm{d}x.$$

Since $\pi_{\mathscr{S}}(Q) = S$, we have $Q \subseteq S$ and it is not possible that $Q \subseteq S' \in \mathrm{ch}_{\mathscr{S}}(S)$. Hence $S' \cap Q \in \{\varnothing, S'\}$ for all $S' \in \mathrm{ch}_{\mathscr{S}}(S)$ and Q is exactly partitioned by $Q \cap E_{\mathscr{S}}(S)$ and those $S' \in \mathrm{ch}_{\mathscr{S}}(S)$ with $S' \subsetneq Q$. Thus

$$\sum_{S'\in \mathrm{ch}_{\mathscr{S}}} \langle f\rangle_{S'} \frac{|S'\cap Q|}{|Q|} = \frac{1}{|Q|}\sum_{S'\in \mathrm{ch}_{\mathscr{S}}, S'\subseteq Q} \langle f\rangle_{S'}|S'| = \frac{1}{|Q|}\sum_{S'\in \mathrm{ch}_{\mathscr{S}}, S'\subseteq Q}\int_{S'} f\,\mathrm{d}x,$$

and

$$\langle P_S f\rangle_Q = \frac{1}{|Q|}\sum_{S'\in \mathrm{ch}_{\mathscr{S}}, S'\subseteq Q}\int_{S'} f\,\mathrm{d}x + \frac{1}{|Q|}\int_{Q\cap E_{\mathscr{S}}(S)} f\,\mathrm{d}x = \frac{1}{|Q|}\int_Q f\,\mathrm{d}x,$$

confirming the lemma. □

A typical way in which a sparse collection arises is via the following basic construction:

Definition 11.1.16 (Principal cubes). *Let $Q_0 \in \mathscr{D}$ and $f \in L^1(Q_0; X)$. The collection of* principal cubes *of f in $\mathscr{D}(Q_0)$ with parameter $A > 1$ is the family $\mathscr{S} = \bigcup_{k=0}^{\infty}\mathscr{S}_k$ constructed as follows:*

(1) $\mathscr{S}_0 := \{Q_0\}$.
(2) *If $\mathscr{S}_k$ is already defined for some $k \in \mathbb{N}$, then*
   (a) *for each $S \in \mathscr{S}_k$ we let*

$$\mathrm{ch}_{\mathscr{S}}(S) := \Big\{S' \in \mathscr{D}(S) \text{ maximal with } \langle\|f\|_X\rangle_{S'} > A\langle\|f\|_X\rangle_S\Big\},$$

   (b) *and then*

$$\mathscr{S}_{k+1} := \bigcup_{S\in\mathscr{S}_k} \mathrm{ch}_{\mathscr{S}}(S).$$

The first instance of the interplay of a function and its principal cubes is the following:

Lemma 11.1.17. *Let $f \in L^1(Q_0; X)$ and $\mathscr{S}$ be the principal cubes of f with parameter $A > 1$. Then $\mathscr{S}$ is $(1 - A^{-1})$-sparse, and in fact*

$$|E_{\mathscr{S}}(S)| \geqslant (1 - \frac{1}{A})|S|. \tag{11.6}$$

If $P_S f$ is defined by (11.5), then $\|P_S f\|_{L^\infty(\mathbb{R}^d;X)} \leqslant 2^d A\langle\|f\|_X\rangle_S$.

Note that (11.6) is slightly more than the mere $(1 - A^{-1})$-sparseness of $\mathscr{S}$: it says that the disjoint subsets $E(S) \subseteq S$ in the definition of sparseness may be chosen as $E(S) = E_{\mathscr{S}}(S)$, which is not always the case for an arbitrary sparse family. For instance, $\mathscr{S} = \{[0,1), [0,\frac{1}{2}), [\frac{1}{2},1)\}$ is $\frac{1}{2}$-sparse, and one can take for instance $E([0,1)) = [\frac{1}{4},\frac{3}{4})$, $E([0,\frac{1}{2})) = [0,\frac{1}{4})$ and $E([\frac{1}{2},1)) = [\frac{3}{4},1)$, but $E_{\mathscr{S}}([0,1)) = \varnothing$ in this case.

Proof. By maximality, the cubes $S' \in \mathrm{ch}_{\mathscr{S}}(S)$ are pairwise disjoint. From the defining condition it follows that

$$\sum_{S' \in \mathrm{ch}_{\mathscr{S}}(S)} |S'| \leqslant \sum_{S' \in \mathrm{ch}_{\mathscr{S}}(S)} \frac{\int_{S'} \|f\|_X \, \mathrm{d}x}{A\langle \|f\|_X \rangle_S} \leqslant \frac{\int_S \|f\|_X \, \mathrm{d}x}{A\langle \|f\|_X \rangle_S} = \frac{|S|}{A}$$

and hence

$$|E_{\mathscr{S}}(S)| = |S| - \sum_{S' \in \mathrm{ch}_{\mathscr{S}}(S)} |S'| \geqslant (1 - \frac{1}{A})|S|.$$

If $x \in E_{\mathscr{S}}(S)$, then x is not contained in any $S' \in \mathrm{ch}_{\mathscr{S}}(S)$, and hence $\langle \|f\|_X \rangle_Q \leqslant A\langle \|f\|_X \rangle_S$ for all $Q \in \mathscr{D}(S)$ with $x \in Q$. As $\ell(Q) \to 0$, it follows from Lebesgue's Differentiation Theorem that $\|P_S f(x)\|_X = \|f(x)\|_X \leqslant A\langle \|f\|_X \rangle_S$ for almost every $x \in E_{\mathscr{S}}(S)$. If $x \in S' \in \mathrm{ch}_{\mathscr{S}}(S)$, then $f_S(x) = \langle f \rangle_{S'}$. By the maximality of S', its dyadic parent $\widehat{S}'$ satisfies the opposite inequality $\langle \|f\|_X \rangle_{\widehat{S}'} \leqslant A\langle \|f\|_X \rangle_S$, and hence

$$\begin{aligned} \|P_S f(x)\|_X \leqslant \langle \|f\|_X \rangle_{S'} &= \frac{1}{|S'|} \int_{S'} \|f\|_X \, \mathrm{d}x \\ &\leqslant \frac{2^d}{|\widehat{S}'|} \int_{\widehat{S}'} \|f\|_X \, \mathrm{d}x \leqslant 2^d A \langle \|f\|_X \rangle_S. \end{aligned}$$

These two cases confirm the upper bound $\|P_S f\|_{L^\infty(\mathbb{R}^d;X)} \leqslant 2^d A\langle \|f\|_X \rangle_S$. □

11.1.c Maximal oscillatory norms for L^p spaces

Based on the oscillations studied above, we introduce the related *John–Strömberg maximal operator*

$$M^{\#}_{0,\lambda} f(x) := \sup_{Q \ni x} \mathrm{osc}_\lambda(f; Q),$$

where the supremum is taken over all cubes containing $x \in \mathbb{R}^d$; a dyadic version $M^{\#,\mathscr{D}}_{0,\lambda}$ is obtained by restricting the supremum to dyadic cubes $Q \in \mathscr{D}$ only. Via this maximal operator we can obtain a useful oscillatory characterisation of $L^p(\mathbb{R}^d; X)$, which we will prove in the rest of this section:

Theorem 11.1.18. *Let X be a Banach space, $p \in (0, \infty)$, $\lambda = 2^{-2-d}$, and $f \in L^0(\mathbb{R}^d; X)$. Then there is a constant $c \in X$ such that $f - c \in L^p(\mathbb{R}^d; X)$ if and only if $M^{\#}_{0,\lambda} f \in L^p(\mathbb{R}^d)$, and in this case*

$$c_d^{-1/p} \|M^{\#}_{0,\lambda} f\|_{L^p(\mathbb{R}^d)} \leqslant \|f - c\|_{L^p(\mathbb{R}^d;X)} \leqslant c_p \|M^{\#}_{0,\lambda} f\|_{L^p(\mathbb{R}^d)},$$

where $c_p = 8p$ for $p \in [1, \infty)$ and $c_p = 2^{2+1/p}$ for $p \in (0, 1)$.

The result is also valid with $\mathbb{R}^d$ replaced by a cube $Q_0 \subseteq \mathbb{R}^d$ or a quadrant $S \subseteq \mathbb{R}^d$, and with the supremum in the maximal operator $M^{\#}_{0,\lambda}$ restricted to cubes $Q \subseteq Q_0$ or $Q \subseteq S$, respectively.

Remark 11.1.19. If we now *a priori* require that $f \in L^{p_0,\infty}(\mathbb{R}^d; X)$ for some $p_0 \in (0,\infty)$ (unrelated to the exponent p), then the constant $c \in X$ guaranteed by Theorem 11.1.18 is necessarily 0, and thus in fact $f \in L^p(\mathbb{R}^d; X)$.

Namely, if $f \in L^{p_0,\infty}(\mathbb{R}^d; X)$ and $f - c \in L^p(\mathbb{R}^d; X)$, it follows that $c = f - (f-c) \in L^{p_0,\infty}(\mathbb{R}^d; X) + L^p(\mathbb{R}^d; X)$, thus

$$|\{\|c\| > t\}| \leqslant |\{\|f\| > t/2\}| + |\{\|f - c\| > t/2\}| < \infty$$

for all $t > 0$, which would lead to a contradiction for $t \in (0, \|c\|)$.

By Lemma 11.1.1 for any $q \in (0,\infty)$, we have

$$\operatorname{osc}_\lambda(f;Q) \leqslant \inf_{c\in X} \frac{\|(f-c)\mathbf{1}_Q\|_{L^{q,\infty}}}{(\lambda|Q|)^{1/q}} \leqslant \frac{\|f\mathbf{1}_Q\|_{L^q}}{(\lambda|Q|)^{1/q}} = \lambda^{-1/q}\Big(\fint_Q \|f\|^q\Big)^{1/q}.$$

Taking the supremum over all cubes Q containing a given point, it follows that

$$M^{\#}_{0,\lambda} f \leqslant \lambda^{-1/q} M_q f, \qquad M_q f := \big(M(\|f\|^q)\big)^{1/q}, \tag{11.7}$$

where M is the Hardy–Littlewood maximal operator. The L^p boundedness of M_q is an easy combination of some estimates collected from Chapter 3:

Lemma 11.1.20. *For all* $0 < q < p < \infty$, *we have*

$$\max\Big(\|M_q\|_{L^p\to L^p}, \|M_q\|_{L^{p,\infty}\to L^{p,\infty}}\Big) \leqslant 3^{d/q+d/p}\Big(\frac{p}{p-q}\Big)^{1/q}.$$

Proof. The dyadic (in fact more general martingale) bounds for $M_q^{\mathscr{D}}$ on L^p and $L^{p,\infty}$ for $p \in (q,\infty)$, with norm bound $(p/(p-q))^{1/q}$ in each case, have been treated in Lemma 3.5.17. On the other hand, we recall from (3.36) that

$$Mf \leqslant 3^d \sup_{\alpha\in\{0,\frac13,\frac23\}^d} M^\alpha f,$$

thus

$$M_q f \leqslant 3^{d/q} \sup_{\alpha\in\{0,\frac13,\frac23\}^d} M_q^\alpha f \leqslant 3^{d/q}\Big(\sum_{\alpha\in\{0,\frac13,\frac23\}^d} [M_q^\alpha f]^p\Big)^{1/p}.$$

Hence

$$\|M_q f\|_p \leqslant 3^{d/q}\Big(\sum_{\alpha\in\{0,\frac13,\frac23\}^d} \|M_q^\alpha f\|_p^p\Big)^{1/p} \leqslant 3^{d/q+d/p}\Big(\frac{p}{p-q}\Big)^{1/q}\|f\|_p,$$

and, for every $\lambda > 0$,

$$\lambda|\{M_q f > \lambda\}|^{1/p} \leqslant \lambda\Big(\sum_\alpha |\{M_q^\alpha f > 3^{-d/q}\lambda\}|\Big)^{1/p} \leqslant 3^{d/q}\Big(\sum_\alpha \|M_q^\alpha f\|^p_{L^{p,\infty}}\Big)^{1/p},$$

after which the last step is exactly as in the strong-type case, now using the weak-type boundedness of the dyadic M_q^α instead. □

Proposition 11.1.21. *The operator $M^{\#}_{0,\lambda}$ is bounded from $L^p(\mathbb{R}^d;X)$ to $L^p(\mathbb{R}^d)$ and from $L^{p,\infty}(\mathbb{R}^d;X)$ to $L^{p,\infty}(\mathbb{R}^d)$, with norm at most $c_{d,\lambda}^{1/p}$, where $c_{d,\lambda}$ is a constant depending only on d and λ.*

The first half of Theorem 11.1.18 is immediate from this proposition (with the choice $\lambda = 2^{-2-d}$ so that $c_{d,\lambda} = c_d$), combined with the trivial observation that $M^{\#}_{0,\lambda} f = M^{\#}_{0,\lambda}(f-c)$ for any constant $c \in X$.

Proof. Let $Y \in \{L^p, L^{p,\infty}\}$. By (11.7) and Lemma 11.1.20, we have

$$\|M^{\#}_{0,\lambda} f\|_Y \leqslant \lambda^{-1/q} \|M_q f\|_Y \leqslant \lambda^{-1/q} 3^{d/q+d/p} \Big(\frac{p}{p-q}\Big)^{1/q} \|f\|_Y.$$

With, say, $q = \frac{1}{2}p$, the right hand side takes the form $(\lambda^{-2} 3^{3d} 2^2)^{1/p} \|f\|_Y$. □

Towards the deduction of a global L^p estimate from local ones, we record:

Lemma 11.1.22. *Let X be a Banach space and $p \in (0,\infty)$. Suppose that $f \in L^p_{\rm loc}(\mathbb{R}^d;X)$ satisfies*

$$\|\mathbf{1}_Q (f - c_Q)\|_p \leqslant K$$

for some constants $c_Q \in X$ and all cubes $Q \subseteq \mathbb{R}^d$. Then there is a constant $c \in X$ such that $f - c \in L^p(\mathbb{R}^d;X)$ and

$$\|f - c\|_p \leqslant K.$$

Proof. Consider an increasing sequence of cubes $Q_1 \subseteq Q_2 \subseteq \dots$ such that $\bigcup_{n=1}^\infty Q_n = \mathbb{R}^d$. If $m \leqslant n$, then

$$\begin{aligned}
\|c_{Q_m} - c_{Q_n}\| &= |Q_m|^{-1/p} \|\mathbf{1}_{Q_m}(c_{Q_m} - c_{Q_n})\|_p \\
&\leqslant |Q_m|^{-1/p} \Big(\|\mathbf{1}_{Q_m}(f - c_{Q_m})\|_p + \|\mathbf{1}_{Q_n}(f - c_{Q_n})\|_p \Big) \\
&\leqslant |Q_m|^{-1/p} 2K \to 0 \qquad \text{as} \quad m \to \infty.
\end{aligned}$$

Hence $(c_{Q_n})_{n\geqslant 1}$ is a Cauchy sequence and thus convergent to some $c \in X$. Now Fatou's lemma shows that

$$\int_{\mathbb{R}^d} \|f - c\|^p = \int_{\mathbb{R}^d} \lim_{n\to\infty} \mathbf{1}_{Q_n} \|f - c_{Q_n}\|^p \leqslant \liminf_{n\to\infty} \int_{Q_n} \|f - c_{Q_n}\|^p \leqslant K,$$

which completes the proof. □

We can now prove the remaining half of Theorem 11.1.18, which we restate as:

Proposition 11.1.23. *Let $f \in L^0(\mathbb{R}^d;X)$, $\lambda = 2^{-2-d}$, and suppose that $M^{\#}_{0,\lambda} f \in L^p(\mathbb{R}^d)$ for some $p \in (0,\infty)$. Then there is a constant $c \in X$ such that $f - c \in L^p(\mathbb{R}^d;X)$ and*

$$\|f-c\|_{L^p(\mathbb{R}^d;X)} \leqslant c_p \|M^{\#}_{0,\lambda} f\|_{L^p(\mathbb{R}^d)}, \qquad c_p = \begin{cases} 8p, & p \in [1,\infty), \\ 2^{2+1/p}, & p \in (0,1). \end{cases}$$

The result also holds with $\mathbb{R}^d$ replaced by a cube $Q_0 \subseteq \mathbb{R}^d$ or a quadrant $S \subseteq \mathbb{R}^d$, and with the supremum in the maximal operator $M^{\#}_{0,\lambda}$ restricted to cubes contained in Q_0 or S, respectively.

Proof. Consider a fixed cube $Q^0 \subseteq \mathbb{R}^d$. By Lerner's formula (Theorem 11.1.12), there is a $\frac{1}{2}$-sparse subcollection $\mathscr{S} \subseteq \mathscr{D}(Q^0)$ such that

$$\mathbf{1}_{Q^0} \|f - m_f(Q^0)\| \leqslant 4 \sum_{S \in \mathscr{S}} \mathbf{1}_S \operatorname{osc}_\lambda(f;S),$$

whenever $m_f(Q^0)$ is a λ-pseudomedian of f on Q^0. Taking L^p norms and using Proposition 11.1.11 (with $\gamma = \frac{1}{2}$), we get

$$\begin{aligned} \|\mathbf{1}_{Q^0}(f - m_f(Q^0))\|_p &\leqslant 4 \Big\| \sum_{S \in \mathscr{S}} \mathbf{1}_S \operatorname{osc}_\lambda(f;S) \Big\|_p \\ &\leqslant 4c_{p,\frac{1}{2}} \Big\| \sum_{S \in \mathscr{S}} \mathbf{1}_{E(S)} \operatorname{osc}_\lambda(f;S) \Big\|_p \leqslant 4c_{p,\frac{1}{2}} \|M^{\#}_{0,\lambda} f\|_p. \end{aligned}$$

This estimate is uniform with respect to the choice of $Q^0 \subseteq \mathbb{R}^d$; hence we can apply Lemma 11.1.22 with $c_Q = m_f(Q)$ to complete the proof.

The variant in the case of a cube or a quadrant in place of $\mathbb{R}^d$ is immediate by inspection of the argument. □

We conclude this section with an end-point analogue of Theorem 11.1.18 for the space $\mathrm{BMO}(\mathbb{R}^d;X)$ in place of $L^p(\mathbb{R}^d;X)$. Recall that we have previously defined the space $\mathrm{BMO}(\mathbb{R}^d;X)$ of functions of *bounded mean oscillation* as the class of functions $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d;X)$ such that

$$\|f\|_{\mathrm{BMO}(\mathbb{R}^d;X)} := \sup_Q \inf_{c \in X} \fint_Q \|f - c\| < \infty.$$

Proposition 11.1.24. *Let X be a Banach space, $\lambda = 2^{-2-d}$, and $f \in L^0(\mathbb{R}^d;X)$. Then $f \in \mathrm{BMO}(\mathbb{R}^d;X)$ if and only if $M^{\#}_{0,\lambda} f \in L^\infty(\mathbb{R}^d)$, and*

$$\lambda \|M^{\#}_{0,\lambda} f\|_{L^\infty(\mathbb{R}^d)} \leqslant \|f\|_{\mathrm{BMO}(\mathbb{R}^d;X)} \leqslant 8 \|M^{\#}_{0,\lambda} f\|_\infty.$$

Proof. From Lemma 11.1.1 it is immediate that

$$\operatorname{osc}_\lambda(f;Q) \leqslant \frac{1}{\lambda} \inf_{c \in X} \fint_Q \|f - c\|,$$

from which the first claimed inequality follows by taking the supremum over all cubes $Q \subseteq \mathbb{R}^d$.

In the other direction, given a cube $Q \subseteq \mathbb{R}^d$, Lerner's formula (Theorem 11.1.12) guarantees that

$$\fint_Q \|f - m_f(Q)\| \leqslant \frac{4}{|Q|} \sum_{S \in \mathscr{S}} |S| \operatorname{osc}_\lambda(f; S)$$
$$\leqslant \frac{4}{|Q|} \sum_{S \in \mathscr{S}} 2|E(S)| \|M^\#_{0,\lambda} f\|_\infty \leqslant 8\|M^\#_{0,\lambda} f\|_\infty,$$

and taking the supremum over all cubes Q proves the second bound. □

11.1.d The dyadic Hardy space and BMO

Often an efficient way of capturing the relevant local oscillations of a function is in terms of the following notion:

Definition 11.1.25 (Atom). *A function $a : \mathbb{R}^d \to X$ is called a (normalised)* $H^1_{\mathscr{D}}$-atom *if*

(i) $\operatorname{supp} a \subseteq Q$ *for some* $Q \in \mathscr{D}$*;*
(ii) $a \in L^\infty(\mathbb{R}^d; X)$ *(and* $\|a\|_\infty \leqslant 1/|Q|$*);*
(iii) $\int_Q a = 0$.

It is immediate that a normalised atom satisfies $\|a\|_1 \leqslant 1$. If $a \neq 0$ is an atom supported on $Q \in \mathscr{D}$, then $\dfrac{a}{|Q|\|a\|_\infty}$ is a normalised atom. Out of these atoms we can then construct a useful subspace of $L^1(\mathbb{R}^d; X)$:

Definition 11.1.26 (Atomic Hardy space). *The atomic Hardy space*

$$H^1_{\mathscr{D},\mathrm{at}}(\mathbb{R}^d; X)$$

consists of all $f \in L^1(\mathbb{R}^d; X)$ that admit a representation

$$f = \sum_{k=1}^\infty \alpha_k \Big(= \sum_{k=1}^\infty \lambda_k a_k \Big),$$

absolutely convergent in $L^1(\mathbb{R}^d; X)$, where each α_k is an $H^1_{\mathscr{D}}$-atom supported in some $Q_k \in \mathscr{D}$ (or each a_k is a normalised $H^1_{\mathscr{D}}$-atom and $\lambda_k \in \mathbb{K}$) with

$$\sum_{k=1}^\infty \|\alpha_k\|_\infty |Q_k| < \infty \quad \Big(\sum_{k=1}^\infty |\lambda_k| < \infty \Big).$$

The norm in this space is defined as

$$\|f\|_{H^1_{\mathscr{D},\mathrm{at}}} := \inf \sum_{k=1}^\infty \|\alpha_k\|_\infty |Q_k| \Big(= \inf \sum_{k=1}^\infty |\lambda_k| \Big)$$

where the infimum is taken over all such representations.

It is immediate that the two versions of the definition are equivalent via the correspondence $\lambda_k = \|\alpha_k\|_\infty |Q_k|$ and $a_k = \lambda_k^{-1}\alpha_k$.

A disadvantage of this definition is the difficulty of checking the membership of a given function in $H^1_{\mathscr{D},\mathrm{at}}(\mathbb{R}^d;X)$, as doing this via the definition would require one to construct the atomic decomposition, which might not be an easy task. The following notion is much more amenable to this:

Definition 11.1.27 (Maximal Hardy space). *The maximal Hardy space*

$$H^1_{\mathscr{D},\max}(\mathbb{R}^d;X)$$

consists of all $f \in L^1(\mathbb{R}^d;X)$ for which also the (cancellative) dyadic maximal function

$$M_{\mathscr{D}}f(x) := \sup_{Q\in\mathscr{D}} \mathbf{1}_Q(x)\|\langle f\rangle_Q\|_X$$

satisfies $M_{\mathscr{D}}f \in L^1(\mathbb{R}^d)$. The norm in this space is defined as

$$\|f\|_{H^1_{\mathscr{D},\max}} := \|M_{\mathscr{D}}f\|_{L^1(\mathbb{R}^d)}.$$

Theorem 11.1.28. *Let X be a Banach space. The spaces $H^1_{\mathscr{D},\mathrm{at}}(\mathbb{R}^d;X)$ and $H^1_{\mathscr{D},\max}(\mathbb{R}^d;X)$ are equal with equivalent norms; in fact*

$$\|h\|_{H^1_{\mathscr{D},\max}(\mathbb{R}^d;X)} \leqslant \|h\|_{H^1_{\mathscr{D},\mathrm{at}}(\mathbb{R}^d;X)} \leqslant 6\cdot 2^d \cdot \|h\|_{H^1_{\mathscr{D},\max}(\mathbb{R}^d;X)}.$$

Proof. Suppose first that $a \in L^\infty(\mathbb{R}^d;X)$ satisfies $\operatorname{supp} a \subseteq Q$ for some dyadic cube and $\int a = 0$. Then $\langle a\rangle_R \neq 0$ only if $R \subsetneq Q$, and hence $\operatorname{supp} M_{\mathscr{D}}a \subseteq Q$ as well. It follows that

$$\|M_{\mathscr{D}}a\|_1 \leqslant |Q|\|M_{\mathscr{D}}a\|_\infty \leqslant |Q|\|a\|_\infty.$$

If $h = \sum_{i=1}^\infty a_i$ is a series of such function on cubes Q_i, then by sublinearity

$$\|h\|_{H^1_{\mathscr{D},\max}(\mathbb{R}^d;E)} = \|M_{\mathscr{D}}h\|_1 \leqslant \sum_{i=1}^\infty \|M_{\mathscr{D}}a_i\|_1 \leqslant \sum_{i=1}^\infty |Q_i|\|a_i\|_\infty,$$

and taking the infimum over all such representations of h shows that

$$\|h\|_{H^1_{\mathscr{D},\max}(\mathbb{R}^d;X)} \leqslant \|h\|_{H^1_{\mathscr{D},\mathrm{at}}(\mathbb{R}^d;X)}.$$

In the other direction, suppose that $h \in H^1_{\max}(\mathbb{R}^d;X)$. Given $\lambda > 0$, let $\mathscr{Q}_\lambda$ be the collection of maximal dyadic cubes Q such that $\|\langle h\rangle_Q\|_X > \lambda$. Then

$$\sum_{Q\in\mathscr{Q}_\lambda} |Q| = |\{M_{\mathscr{D}}h > \lambda\}| \leqslant \frac{1}{\lambda}\|M_{\mathscr{D}}h\|_{L^1(\mathbb{R}^d)} = \frac{1}{\lambda}\|h\|_{H^1_{\mathscr{D},\max}(\mathbb{R}^d;X)}.$$

Let $\widehat{\mathscr{Q}}_\lambda$ be the collection of maximal dyadic cubes that have a child in $\mathscr{Q}_\lambda$. Thus these cubes do not belong to $\mathscr{Q}_\lambda$ themselves. Hence $\|\langle h\rangle_Q\|_E \leqslant \lambda$ for $Q \in \widehat{\mathscr{Q}}_\lambda$, and also

$$\sum_{Q\in\widehat{\mathscr{D}}_\lambda}|Q|\leqslant\sum_{Q\in\mathscr{D}_\lambda}|\widehat{Q}|=\sum_{Q\in\mathscr{D}_\lambda}2^d|Q|=2^d|\{Mh>\lambda\}|.$$

Let then

$$g_\lambda:=1_{\complement(\bigcup\widehat{\mathscr{D}}_\lambda)}h+\sum_{Q\in\widehat{\mathscr{D}}_\lambda}1_Q\langle h\rangle_Q,\qquad b_\lambda:=\sum_{Q\in\widehat{\mathscr{D}}_\lambda}1_Q(h-\langle h\rangle_Q).$$

By definition of $M_{\mathscr{D}}$, we have $\|\langle h\rangle_Q\|_X\leqslant M_{\mathscr{D}}h(x)$ whenever $x\in Q\in\mathscr{D}$. As $\ell(Q)\to 0$, this gives $\|f(x)\|_X\leqslant M_{\mathscr{D}}h(x)$ at a.e. x by the Lebesgue Differentiation Theorem. Thus $\|g_\lambda(x)\|_\leqslant M_{\mathscr{D}}h(x)$ almost everywhere. On the other hand, we have $\|\langle h\rangle_Q\|_X\leqslant\lambda$ for $Q\in\widehat{\mathscr{D}}_\lambda$, and $M_{\mathscr{D}}h(x)\leqslant\lambda$ for $x\in\complement(\bigcup\widehat{\mathscr{D}}_\lambda)$; thus in fact $\|g_\lambda\|_X\leqslant\min(\lambda,M_{\mathscr{D}}h)$ almost everywhere, where $M_{\mathscr{D}}h\in L^1(\mathbb{R}^d)$. Moreover, $g_\lambda=h$ on $\{M_{\mathscr{D}}h\leqslant\lambda\}\to\mathbb{R}^d$ as $\lambda\to\infty$, and hence

$$g_\lambda\to\begin{cases}h, & \lambda\to\infty,\\ 0, & \lambda\to 0,\end{cases}$$

pointwise, and by dominated convergence also in $L^1(\mathbb{R}^d;X)$. Thus

$$\begin{aligned}h&=\sum_{k\in\mathbb{Z}}(g_{2^{k+1}}-g_{2^k})=\sum_{k\in\mathbb{Z}}(b_{2^k}-b_{2^{k+1}})\\
&=\sum_{k\in\mathbb{Z}}\Big(\sum_{Q\in\widehat{\mathscr{D}}_{2^k}}1_Q(h-\langle h\rangle_Q)-\sum_{R\in\widehat{\mathscr{D}}_{2^{k+1}}}1_R(h-\langle h\rangle_R)\Big)\\
&=\sum_{k\in\mathbb{Z}}\sum_{Q\in\widehat{\mathscr{D}}_{2^k}}\Big(1_{Q\setminus\bigcup\widehat{\mathscr{D}}_{2^{k+1}}}(h-\langle h\rangle_Q)+\sum_{\substack{R\in\widehat{\mathscr{D}}_{2^{k+1}}\\ R\subseteq Q}}1_R(\langle h\rangle_R-\langle h\rangle_Q)\Big)\\
&=:\sum_{k\in\mathbb{Z}}\sum_{Q\in\widehat{\mathscr{D}}_{2^k}}a_{k,Q}.\end{aligned}$$

Here $\operatorname{supp}a_Q\subseteq Q$, $\int a_Q=0$ and $\|a_Q\|_\infty\leqslant 2^{k+1}+2^k=3\cdot 2^k$. Hence

$$\begin{aligned}\|h\|_{H^1_{\mathscr{D},\mathrm{at}}}&\leqslant\sum_{k\in\mathbb{Z}}\sum_{Q\in\widehat{\mathscr{D}}_{2^k}}|Q|\|a_{k,Q}\|_\infty\leqslant\sum_{k\in\mathbb{Z}}3\cdot 2^k\sum_{Q\in\widehat{\mathscr{D}}_{2^k}}|Q|\\
&\leqslant\sum_{k\in\mathbb{Z}}3\cdot 2^k\cdot 2^d|\{M_{\mathscr{D}}h>2^k\}|\\
&\leqslant\sum_{k\in\mathbb{Z}}3\cdot 2\cdot 2^d\int_{2^{k-1}}^{2^k}|\{M_{\mathscr{D}}h>t\}|\,\mathrm{d}t\\
&=6\cdot 2^d\|M_{\mathscr{D}}h\|_{L^1(\mathbb{R}^d)}=6\cdot 2^d\|h\|_{H^1_{\mathscr{D},\max}(\mathbb{R}^d;X)}.\end{aligned}$$

□

Corollary 11.1.29. *The space* $H^1_{\mathscr{D},\mathrm{at}}(\mathbb{R}^d;X)=H^1_{\mathscr{D},\max}(\mathbb{R}^d;X)$ *is complete.*

Proof. It in enough to prove this for $H^1_{\mathscr{D},\max}(\mathbb{R}^d;X)$. Since $\|f(x)\|_X \leqslant M_{\mathscr{D}}f(x)$ at a.e. $x \in \mathbb{R}^d$, we have $\|f\|_{L^1(\mathbb{R}^d;X)} \leqslant \|f\|_{H^1_{\mathscr{D},\max}(\mathbb{R}^d;X)}$. Hence, if $(f_n)_{n=1}^\infty$ is a Cauchy sequence in $H^1_{\mathscr{D},\max}(\mathbb{R}^d;X)$, it is also a Cauchy sequence in $L^1(\mathbb{R}^d;X)$ and thus $\|f_n - f\|_1 \to 0$ for some $f \in L^1(\mathbb{R}^d;X)$. Since $\langle\ \rangle_Q$ is continuous from $L^1(\mathbb{R}^d;X)$ to X, we have for all $x \in Q \in \mathscr{D}$ we have, for each $h \in H^1_{\mathscr{D},\max}(\mathbb{R}^d;X)$,

$$\|\langle f-h\rangle_Q\|_X = \lim_{n\to\infty} \|\langle f_n - h\rangle_Q\|_X \leqslant \liminf_{n\to\infty} M_{\mathscr{D}}(f_n-h)(x);$$

hence $M_{\mathscr{D}}(f-h)(x) \leqslant \liminf_{n\to\infty} M_{\mathscr{D}}(f_n-h)(x)$, and thus by Fatou's lemma

$$\|M_{\mathscr{D}}(f-h)\|_{L^1(\mathbb{R}^d)} \leqslant \liminf_{n\to\infty} \|M_{\mathscr{D}}(f_n-h)\|_{L^1(\mathbb{R}^d)}.$$

With $h=0$, this shows that $f \in H^1_{\mathscr{D},\max}(\mathbb{R}^d;X)$. With $h = f_m$, we find that

$$\lim_{m\to\infty} \|M_{\mathscr{D}}(f-f_m)\|_{L^1(\mathbb{R}^d)} \leqslant \lim_{m\to\infty} \liminf_{n\to\infty} \|M_{\mathscr{D}}(f_n-f_m)\|_{L^1(\mathbb{R}^d)} = 0,$$

and hence $f_m \to f$ in $H^1_{\mathscr{D},\max}(\mathbb{R}^d;X)$. □

Theorem 11.1.30. *Let X be a Banach space. The duality*

$$\langle b,h\rangle := \lim_{N\to\infty} \int \langle b_N, h\rangle = \sum_{i=1}^\infty \int \langle b, a_i\rangle, \quad b_N := \min\Big\{1, \frac{N}{\|b\|_{X^*}}\Big\} b$$

between $b \in \mathrm{BMO}_{\mathscr{D}}(\mathbb{R}^d;X^*)$ and $h \in H^1_{\mathscr{D},\mathrm{at}}(\mathbb{R}^d;X)$ is well defined, and realises $\mathrm{BMO}_{\mathscr{D}}(\mathbb{R}^d;X^*)$ with the norm

$$\|b\|_{\mathrm{BMO}_{\mathscr{D}}(\mathbb{R}^d;X)} := \sup_{Q\in\mathscr{D}} \inf_{c\in X} \fint_Q \|b-c\|_X$$

as an isometric subspace of $(H^1_{\mathscr{D},\mathrm{at}}(\mathbb{R}^d;X))^*$.

Proof. Since all norms BMO norms appearing in this proof are dyadic, we drop the subscript $\mathscr{D}$ for the benefit of slightly lighter notation.

Part 1: Estimating the dual norm by the BMO *norm*

If $\operatorname{supp} a_i \subseteq Q_i \in \mathscr{D}$ and $\int a_i = 0$, we have

$$\Big|\int \langle b, a_i\rangle\Big| = \Big|\int_{Q_i} \langle b-c, a_i\rangle\Big| \leqslant \fint_{Q_i} \|b-c\|_{X^*} |Q_i| \|a_i\|_\infty$$

for all $c \in E^*$. Taking the infimum over $c \in E^*$ it follows that

$$\Big|\int \langle b, a_i\rangle\Big| \leqslant \|b\|_{\mathrm{BMO}} |Q_i| \|a_i\|_\infty$$

and hence $\sum_{i=1}^{\infty}\int\langle b,a_i\rangle$ converges for $b\in \mathrm{BMO}(\mathbb{R}^d;E^*)$ and $\sum_{i=1}^{\infty}a_i\in H^1_{\mathrm{at}}(\mathbb{R}^d;E)$.

One checks that $\|b_N-c_N\|_{X^*}\leqslant 2\|b-c\|_{X^*}$, whence

$$\inf_{c\in E^*}\fint_Q\|b_N-c\|_{E^*}\leqslant\inf_{c\in E^*}\fint_Q\|b_N-c_N\|_{E^*}\leqslant 2\inf_{c\in E^*}\fint_Q\|b-c\|_{E^*},$$

so that $b_N\in(\mathrm{BMO}\cap L^\infty)(\mathbb{R}^d;X^*)$ and

$$\Big|\int\langle b_N,a_i\rangle\Big|\leqslant\|b_N\|_{\mathrm{BMO}}|Q_i|\|a_i\|_\infty\leqslant 2\|b\|_{\mathrm{BMO}}|Q_i|\|a_i\|_\infty.$$

Thus

$$\sum_{i=1}^{\infty}\int\langle b,a_i\rangle=\sum_{i=1}^{\infty}\lim_{N\to\infty}\int\langle b_N,a_i\rangle=\lim_{N\to\infty}\sum_{i=1}^{\infty}\int\langle b_N,a_i\rangle=\lim_{N\to\infty}\int\langle b_N,h\rangle,$$

where the first two identities use dominated convergence in $L^1(Q_i)$ and in ℓ^1, respectively, and the last one follows from the convergence of the series $h=\sum_{i=1}^{\infty}a_i$ in $L^1(\mathbb{R}^d;E)$, and the fact that $b_N\in L^\infty(\mathbb{R}^d;E^*)\subseteq(L^1(\mathbb{R}^d;E))^*$. This shows in particular that the pairing of $\langle b,h\rangle$ is independent of the particular series representation of h, and hence well defined. Taking the infimum over all representations in the estimate

$$|\langle b,h\rangle|\leqslant\sum_{i=1}^{\infty}\|b\|_{\mathrm{BMO}}|Q_i|\|a_i\|_\infty,$$

we find that

$$\|b\|_{(H^1_{\mathrm{at}}(\mathbb{R}^d;X))^*}\leqslant\|b\|_{\mathrm{BMO}(\mathbb{R}^d;X^*)}. \tag{11.8}$$

Part 2: Estimating the BMO *norm by the dual norm*

For the converse estimate, consider a cube Q and suppose first that $s\in L^1(Q;X^*)$ is a simple function, thus measurable with respect to a finite σ-algebra $\mathscr{F}$ of Q. The advantage of this setting is that, for a finite σ-algebra, we have the duality $(L^p(\mathscr{F};X))^*=L^{p'}(\mathscr{F};X^*)$ for an arbitrary Banach space X and for every $p\in[1,\infty]$, including in particular $p=\infty$. Now $\inf_{c\in E^*}\|s-c\|_{L^1(Q;X^*)}$ is the norm of the equivalence class $[s]\in L^1(\mathscr{F};X^*)/X^*$, where $L^1(\mathscr{F};X^*)=(L^\infty(\mathscr{F};X))^*$.

We claim that the quotient space above is the dual of the subspace $L^\infty_0(\mathscr{F};X)\subseteq L^\infty(\mathscr{F};X)$ of functions with mean zero. In fact, recall from Proposition B.1.4 that for any subspace $Y\subseteq Z$, we have the identification $Y^*=Z^*/Y^\perp$, the quotient of Z^* with the annihilator $Y^\perp$ of Y in Z^*. Now $Z=L^\infty(\mathscr{F};X)$ for a finite σ-algebra $\mathscr{F}$, in which case $Z^*=L^1(\mathscr{F};X^*)$. To identify $Y^\perp$ for $Y=L^\infty_0(\mathscr{F};X)$, it is easy to check that the only functions $f\in L^1(\mathscr{F};X^*)$ for which $\int\langle f,g\rangle=0$ for all $g\in L^\infty_0(\mathscr{F};X)$ are the constant functions. Thus indeed $L^1(\mathscr{F};X^*)/X^*=(L^\infty_0(\mathscr{F};X))^*$, and hence

$$\inf_{c\in X^*} \|s-c\|_{L^1(Q;X^*)} = \|[s]\|_{L^1(\mathscr{F};X^*)/X^*} = \sup_{\substack{g\in L^\infty_0(\mathscr{F};X)\\ \|g\|_\infty\leqslant 1}} \Big|\int \langle s,g\rangle\Big|.$$

Now, given $b\in \mathrm{BMO}(\mathbb{R}^d;X^*)$ and a cube Q, we choose a simple $s\in L^1(Q;X^*)$ such that $\|b-s\|_{L^1(Q;X^*)}\leqslant\varepsilon$. Then

$$\begin{aligned}\inf_{c\in X^*} \|b-c\|_{L^1(Q;X^*)} &\leqslant \inf_{c\in X^*} \|s-c\|_{L^1(Q;X^*)} + \varepsilon \\ &\leqslant \sup_{\substack{g\in L^\infty_0(Q;X)\\ \|g\|_\infty\leqslant 1}} \Big|\int \langle s,g\rangle\Big| + \varepsilon \leqslant \sup_{\substack{g\in L^\infty_0(Q;X)\\ \|g\|_\infty\leqslant 1}} \Big|\int \langle b,g\rangle\Big| + 2\varepsilon.\end{aligned}$$

But each $g\in L^\infty_0(Q;X)$ is an $H^1_{\mathscr{D}}$-atom, and hence

$$\Big|\int\langle b,g\rangle\Big| \leqslant \|b\|_{(H^1_{\mathscr{D},\mathrm{at}}(\mathbb{R}^d;X))^*}\|g\|_{H^1_{\mathscr{D},\mathrm{at}}(\mathbb{R}^d;X)} \leqslant \|b\|_{(H^1_{\mathscr{D},\mathrm{at}}(\mathbb{R}^d;X))^*}\|g\|_\infty|Q|.$$

Dividing by $|Q|$ and letting $\varepsilon\to 0$, we obtain

$$\inf_{c\in X^*} \fint_Q \|b-c\|_{X^*} \leqslant \|b\|_{(H^1_{\mathscr{D},\mathrm{at}}(\mathbb{R}^d;X))^*},$$

and hence the estimate converse to (11.8). □

11.2 Singular integrals and extrapolation of L^{p_0} bounds

In this section we study a fairly broad class of kernels satisfying a relatively general integrability condition first introduced by Hörmander. Nevertheless, this condition turns out to be strong enough to yield a fundamental extrapolation property of singular integral operators: once bounded on one L^{p_0} space, they remain bounded on the full scale of L^p spaces for $p\in(1,\infty)$, together with appropriate end-point estimates for $p=1$ and $p=\infty$.

The precise classes of kernels relevant are described in the following definition. We recall that $\dot{\mathbb{R}}^{2d} = \mathbb{R}^{2d}\setminus\{(t,t):\, t\in\mathbb{R}^d\}$.

Definition 11.2.1. *Let X and Y be Banach spaces, $p_0\in[1,\infty]$, and consider*

$$K:\dot{\mathbb{R}}^{2d}\to\mathscr{L}(X,Y),\qquad T\in\mathscr{L}(L^{p_0}(\mathbb{R}^d;X),L^{p_0,\infty}(\mathbb{R}^d;Y)).$$

(1) *We say that T* has kernel K*, or that K* is the kernel of T*, if for every $f\in L^{p_0}_{\mathrm{c}}(\mathbb{R}^d;X)$ and almost every s at a positive distance from* $\operatorname{supp} f$ *the following holds: for every functional $y^*\in Y^*$, the function $t\mapsto\langle K(s,t)f(t),y^*\rangle$ is integrable, and*

$$\langle Tf(s),y^*\rangle = \int\langle K(s,t)f(t),y^*\rangle\,\mathrm{d}t.$$

(2) *We say that* K *is a* Hörmander (resp. operator-Hörmander) kernel*, or satisfies the* Hörmander (resp. operator-Hörmander) condition*, if the following estimate holds for all* $x \in X$ *and* $t, t' \in \mathbb{R}^d$ *with a fixed constant* c *independent of these quantities:*

$$\begin{aligned} &\int_{|s-t|>2|t-t'|} \|[K(s,t) - K(s,t')]x\|_Y \, \mathrm{d}s \leqslant c\|x\|_X \\ \Big(\textit{resp.} \quad &\int_{|s-t|>2|t-t'|} \|K(s,t) - K(s,t')\|_{\mathscr{L}(X,Y)} \, \mathrm{d}x \leqslant c\Big). \end{aligned} \tag{11.9}$$

The smallest admissible c *is denoted by* $\|K\|_{\text{Hör}}$ *(resp.* $\|K\|_{\text{Hör}_{\text{op}}}$*).*

(3) *We say that* K *is a* dual Hörmander (resp. dual operator-Hörmander) kernel*, or satisfies the* dual Hörmander (resp. dual operator-Hörmander) condition*, if the following estimate holds for every* $y^* \in Y^*$ *and* $s, s' \in \mathbb{R}^d$ *with a fixed constant* c' *independent of these quantities:*

$$\begin{aligned} &\int_{|t-s|>2|s-s'|} \|[K(s,t)^* - K(s',t)^*]y^*\|_{X^*} \, \mathrm{d}t \leqslant c'\|y^*\|_{Y^*} \\ \Big(\textit{resp.} \quad &\int_{|t-s|>2|s-s'|} \|K(s,t) - K(s',t)\|_{\mathscr{L}(X,Y)} \, \mathrm{d}t \leqslant c'\Big) \end{aligned} \tag{11.10}$$

The smallest admissible c' *is denoted by* $\|K\|_{\text{Hör}^*}$ *(resp.* $\|K\|_{\text{Hör}^*_{\text{op}}}$*).*

(4) *If* $Q \subseteq \mathbb{R}^d$ *is a cube or a quadrant, we make analogous definitions with each occurrence of* $\mathbb{R}^d$ *replaced by* Q*; in particular, with* $\dot{\mathbb{R}}^{2d}$ *by* $\{(s,t) \in Q \times Q : s \neq t\}$*, and the integrals extended over* Q *only, while keeping the other integrations conditions in force. In this situation, we say that* K *is a (dual/operator) Hörmander kernel on* Q*, respectively.*

Remark 11.2.2. If K is a (dual/operator) Hörmander kernel, then its restriction to $\{(s,t) \in Q \times Q : s \neq t\}$ is a (dual/operator) Hörmander kernel on Q.

Example 11.2.3. A kernel $K(x,y)$ that only depends on the difference $x - y$, i.e., $K(x-y) = k(x-y)$ for some function k, is called a *convolution kernel.* For such kernels, after simple changes of variables, the Hörmander and dual Hörmander conditions take the forms

$$\int_{|s|>2|t|} \|[k(s-t) - k(s)]x\|_Y \, \mathrm{d}s \leqslant c\|x\|_X,$$

$$\int_{|s|>2|t|} \|[k(s-t)^* - k(s)^*]y^*\|_{X^*} \, \mathrm{d}s \leqslant c'\|y^*\|_{Y^*},$$

and similar reformulations of the operator Hörmander conditions are obvious.

The role of these conditions in the extrapolation of L^p-boundedness is summarised in the next theorem. Before stating the result, we make a remark concerning the extension of the action of operators from $L^{p_0}(\mathbb{R}^d; X)$ to $L^\infty(\mathbb{R}^d; X)$. An inherent obstacle here is that the intersection $L^{p_0}(\mathbb{R}^d; X) \cap L^\infty(\mathbb{R}^d; X)$ is not dense in $L^\infty(\mathbb{R}^d; X)$. As a substitute we have:

Lemma 11.2.4. *Let X be a Banach space. The closure of $L^p(\mathbb{R}^d;X)\cap L^\infty(\mathbb{R}^d;X)$ in $L^\infty(\mathbb{R}^d;X)$ is independent of $p\in(0,\infty)$, and it coincides with*

$$\bar{L}^\infty_{\text{fin}}(\mathbb{R}^d;X) := \overline{L^\infty_{\text{fin}}(\mathbb{R}^d;X)}^{L^\infty(\mathbb{R}^d;X)}, \quad \textit{where}$$
$$L^\infty_{\text{fin}}(\mathbb{R}^d;X) := \{f\in L^\infty(\mathbb{R}^d;X) : |\{f\neq 0\}|<\infty\}.$$

Proof. It is clear that $L^\infty_{\text{fin}}(\mathbb{R}^d;X)\subseteq L^p(\mathbb{R}^d;X)\cap L^\infty(\mathbb{R}^d;X)$, and taking the closures of both sides proves one side of the claim.

Conversely, let $p\in(0,\infty)$, a function $f\in L^p(\mathbb{R}^d;X)\cap L^\infty(\mathbb{R}^d;X)$, and $\varepsilon>0$ be given. Now

$$F_\varepsilon := \{\|f(\cdot)\|_X>\varepsilon\}\leqslant \varepsilon^{-p}\|f\|^p_{L^p(\mathbb{R}^d;X)}<\infty,$$

and hence $f_\varepsilon := \mathbf{1}_{F_\varepsilon} f\in L^\infty_{\text{fin}}(\mathbb{R}^d;X)$. On the other hand, it is clear that

$$\|f-f_\varepsilon\|_{L^\infty(\mathbb{R}^d;X)} = \|1_{\complement F_\varepsilon} f\|_{L^\infty(\mathbb{R}^d;X)}\leqslant\varepsilon.$$

Since this can be done for any $\varepsilon>0$, we find that f belongs to the $L^\infty(\mathbb{R}^d;X)$-closure of $L^\infty_{\text{fin}}(\mathbb{R}^d;X)$. Since $f\in L^p(\mathbb{R}^d;X)\cap L^\infty(\mathbb{R}^d;X)$, this whole intersection belongs to the said closure, and then so does the closure of this intersection. This completes the proof. □

Theorem 11.2.5 (Calderón–Zygmund). *Let X and Y be Banach spaces and $p_0\in[1,\infty]$. Let*

$$T\in\mathscr{L}(L^{p_0}(\mathbb{R}^d;X),L^{p_0,\infty}(\mathbb{R}^d;Y))$$

(where $L^{\infty,\infty}:=L^\infty$) with norm $A_0:=\|T\|_{\mathscr{L}(L^{p_0}(\mathbb{R}^d;X),L^{p_0,\infty}(\mathbb{R}^d;Y))}$.

(1) *If T has a Hörmander kernel K, then*
   (a) *T extends uniquely to $T\in\mathscr{L}(L^p(\mathbb{R}^d;X),L^p(\mathbb{R}^d;Y))$ for all $p\in(1,p_0)$, and*

$$\|T\|_{\mathscr{L}(L^p(\mathbb{R}^d;X),L^p(\mathbb{R}^d;Y))}\leqslant c_d\Big(\frac{p_0-1}{(p_0-p)(p-1)}\Big)^{1/p}(A_0+\|K\|_{\text{Hör}});$$

   (b) *T extends uniquely to $T\in\mathscr{L}(L^1(\mathbb{R}^d;X),L^{1,\infty}(\mathbb{R}^d;Y))$ and*

$$\|T\|_{\mathscr{L}(L^1(\mathbb{R}^d;X),L^{1,\infty}(\mathbb{R}^d;Y))}\leqslant c_d(A_0+\|K\|_{\text{Hör}}).$$

(2) *If T has a dual Hörmander kernel K, then*
   (a) *T extends uniquely to $T\in\mathscr{L}(L^p(\mathbb{R}^d;X),L^p(\mathbb{R}^d;Y))$ for all $p\in(p_0,\infty)$, and*

$$\|T\|_{\mathscr{L}(L^p(\mathbb{R}^d;X),L^p(\mathbb{R}^d;Y))}\leqslant c_d p\Big(\frac{p_0}{p-p_0}\Big)^{1/p}(A_0+\|K\|_{\text{Hör}^*});$$

(b) *T extends uniquely to $T \in \mathscr{L}(\bar{L}^\infty_{\mathrm{fin}}(\mathbb{R}^d;X), \mathrm{BMO}(\mathbb{R}^d;Y))$, where the space $\bar{L}^\infty_{\mathrm{fin}}(\mathbb{R}^d;X)$ is as in Lemma 11.2.4, and*

$$\|T\|_{\mathscr{L}(\bar{L}^\infty_{\mathrm{fin}}(\mathbb{R}^d;X),\mathrm{BMO}(\mathbb{R}^d;Y))} \leqslant c_d(A_0 + \|K\|_{\mathrm{Hör}})\|f\|_{L^\infty(\mathbb{R}^d;X)}$$

for all f in this space.

(3) *If T has a kernel K that satisfies both the Hörmander and the dual Hörmander conditions, then for all $p \in (1,\infty)$, T extends uniquely to $T \in \mathscr{L}(L^p(\mathbb{R}^d;X), L^p(\mathbb{R}^d;Y))$, and*

$$\|T\|_{\mathscr{L}(L^p(\mathbb{R}^d;X),L^p(\mathbb{R}^d;Y))} \leqslant c_d \cdot pp' \cdot (A_0 + \|K\|_{\mathrm{Hör}} + \|K\|_{\mathrm{Hör}^*}).$$

(4) *All claims remain valid when $\mathbb{R}^d$ is replaced either by a cube or a quadrant throughout. In this case, it suffices to relax the Hörmander conditions accordingly, as in Definition 11.2.1(4).*

The rest of this section is dedicated to a case-by-case proof of the different assertions of Theorem 11.2.5. For the proof of (1), we introduce the fundamental *Calderón–Zygmund decomposition* in Proposition 11.2.6. The proof of (2), in turn, depends on the notion of local oscillations developed in Section 11.1. The result of (2b) does not directly allow the extension of T to all of $L^\infty(\mathbb{R}^d;X)$ since $L^{p_0}(\mathbb{R}^d;X) \cap L^\infty(\mathbb{R}^d;X)$ is not dense in this space; see Theorem 11.2.9 for results in this direction. The proof of (3) is essentially a combination of (1) and (2), but note that this case provides additional information about $p = p_0$ (bootstrapping the initial weak-type bound into a strong-type one) and improves the quantitative estimates for p close to p_0, where the bounds provided by (1) and (2) blow up as $p \to p_0$. Finally, the claims (4) will be dealt with by indicating the relevant modifications in the proofs of (1) through (3). As it turns out, these modifications are fairly minor, although in the case of (1) they might not be entirely obvious.

11.2.a Calderón–Zygmund decomposition and case $p \in (1, p_0)$

The key to extrapolating in this range is the following classical result:

Proposition 11.2.6 (Calderón–Zygmund decomposition). *Let X be a Banach space. Given $f \in L^1(\mathbb{R}^d;X)$ and $\lambda > 0$, there exists a decomposition $f = g + b$, where*

$$\|g\|_\infty \leqslant 2^d\lambda, \qquad \|g\|_1 \leqslant \|f\|_1,$$

and $b = \sum_i b_i$, where

$$\operatorname{supp} b_i \subseteq Q_i, \quad \int b_i = 0, \quad \sum_i |Q_i| \leqslant \frac{1}{\lambda}\|f\|_1, \quad \sum_i \|b_i\|_1 \leqslant 2\|f\|_1$$

for some disjoint dyadic cubes Q_i. If f is simple, then all b_i are also simple.

If $f \in L^1(Q_0;X)$ for some cube $Q_0 \subseteq \mathbb{R}^d$ and $\lambda \geqslant 2^{-d} \fint_{Q_0} \|f\|$, then the cubes Q_i can be chosen as dyadic subcubes of the initial Q_0, and the function g to be supported on Q_0.

If $f \in L^1(S;X)$ for some quadrant of $\mathbb{R}^d$, then we have $Q_i \subseteq S$.

Proof. Let $Q_i \in \mathscr{D}$ be the maximal dyadic cubes such that $\fint_{Q_i} \|f\| > \lambda$. Then they are pairwise disjoint, and

$$\sum_i |Q_i| = |\{M_{\mathscr{D}} f > \lambda\}| \leqslant \frac{1}{\lambda}\|f\|_1.$$

We define $b_i := 1_{Q_i}(f - \langle f \rangle_{Q_i})$ (which is clearly simple if f is), whence the first two properties of b_i are clear, and it remains to estimate

$$\sum_i \|b_i\|_1 \leqslant \sum_i (\|1_{Q_i} f\|_1 + |Q_i| \|\langle f \rangle_{Q_i}\|) \leqslant \sum_i 2 \int_{Q_i} \|f\| \leqslant 2\|f\|_1$$

by the disjointness of the cubes. To ensure that $f = g + b$, we must then define

$$g := 1_{\complement(\bigcup_i Q_i)} f + \sum_i 1_{Q_i} \langle f \rangle_{Q_i},$$

where the terms are disjointly supported. If $x \in \complement(\bigcup_i Q_i)$, then all dyadic cubes $Q \ni x$ satisfy $\fint_Q |f| \leqslant \lambda$, and thus

$$\|g(x)\| = \|f(x)\| = \lim_{\substack{Q \ni x \\ \ell(Q) \to 0}} \fint_Q \|f\| \leqslant \lambda$$

at almost every such x by the Lebesgue Differentiation Theorem 2.3.4 (or in fact just the scalar-valued version, since we apply it to the function $\|f(\cdot)\|$ rather than f itself). On the other hand, the maximality of Q_i implies that its dyadic parent $\widehat{Q}_i$ satisfies the opposite inequality, $\fint_{\widehat{Q}_i} |f| \leqslant \lambda$. Thus

$$\|g(x)\|_X = \|\langle f \rangle_{Q_i}\|_X \leqslant \frac{1}{|Q_i|} \int_{Q_i} \|f\|_X \leqslant \frac{|\widehat{Q}_i|}{|Q_i|} \cdot \frac{1}{|\widehat{Q}_i|} \int_{\widehat{Q}_i} \|f\|_X \leqslant 2^d \cdot \lambda$$

for $x \in Q_i$, and we see that $\|g(x)\| \leqslant 2^d \lambda$ in both cases. Moreover,

$$\|g\|_1 = \int_{\complement(\bigcup_i Q_i)} \|f\| + \sum_i |Q_i| \|\langle f \rangle_{Q_i}\| \leqslant \int_{\complement(\bigcup_i Q_i)} \|f\| + \sum_i \int_{Q_i} \|f\| = \|f\|_1$$

by the disjointness of the cubes.

If $f \in L^1(Q_0; X)$ and $\lambda \geqslant \fint_{Q_0} \|f\|$, then the maximal dyadic subcubes Q_i of Q_0 with $\fint_{Q_i} \|f\| > \lambda$, are necessarily strict subcubes of Q_0, and the same proof produces a decomposition with the claimed additional properties. If $\lambda \in [2^{-d}, 1)\fint_{Q_0} \|f\|$, then we let the family $\{Q_i\}_i$ consist of the initial cube Q_0 only, so that $g := \langle f \rangle_{Q_0} \mathbf{1}_{Q_0}$ and $b := (f - \langle f \rangle_{Q_0})\mathbf{1}_{Q_0}$. Then $\|g\|_\infty = \|\langle f \rangle_{Q_0}\| \leqslant 2^d \lambda$ and $\sum_i |Q_i| = |Q_0| \leqslant \lambda^{-1}\|f\|_1$ by the two assumed bounds on λ. The last claim of the theorem is obvious. □

We can now give:

Proof of Theorem 11.2.5(1). Our plan is to first prove the weak-type result (1b), and then obtain the strong-type bound (1a) via the Marcinkiewicz Interpolation Theorem 2.2.3.

For $f \in L^{p_0}(\mathbb{R}^d; X) \cap L^1(\mathbb{R}^d; X)$ and $\lambda > 0$, we estimate $\lambda|\{\|Tf\| > \lambda\}|$.

Let $f = g + b$ the Calderón–Zygmund decomposition of f at level $\alpha\lambda$ (instead of λ), where α is to be determined. Then

$$\|g\|_{p_0} \leqslant \|g\|_\infty^{1/p_0'} \|g\|_1^{1/p_0} \leqslant (2^d \alpha\lambda)^{1/p_0'} \|f\|_1^{1/p_0},$$

so in particular $g \in L^{p_0}(\mathbb{R}^d; X)$, and thus $b = f - g \in L^{p_0}(\mathbb{R}^d; X)$. Since $b = \sum_i b_i$ and the b_i are disjointly supported, it follows that each b_i also belongs to $L^{p_0}(\mathbb{R}^d; X)$ and the identity $b = \sum_i b_i$ also holds in the sense of convergence in $L^{p_0}(\mathbb{R}^d; X)$. The assumption that $T \in \mathscr{L}(L^{p_0}(\mathbb{R}^d; X), L^{p_0,\infty}(\mathbb{R}^d; Y))$ then implies that

$$Tf = T(g + b) = Tg + Tb, \qquad Tb = T\sum_i b_i = \sum_i Tb_i.$$

If Q_i are the corresponding cubes, let B_i be the concentric ball of twice the diameter and $O^* := \bigcup_i B_i$. Then

$$|\{\|Tf\| > \lambda\}| \leqslant |\{\|Tg\| > \lambda/2\}| + |\{\|Tb\| > \lambda/2\} \setminus O^*| + |O^*|, \tag{11.11}$$

where the last term satisfies

$$|O^*| \leqslant \sum_i |B_i| = \sum_i c_d |Q_i| \leqslant \frac{c_d}{\alpha\lambda} \|f\|_1.$$

For the middle term, we have

$$|\{\|Tb\| > \lambda/2\} \setminus O^*| \leqslant \int_{\complement O^*} \frac{\|Tb\|}{\lambda/2} \leqslant \frac{2}{\lambda} \sum_i \int_{\complement O^*} \|Tb_i\| \leqslant \frac{2}{\lambda} \sum_i \int_{\complement B_i} \|Tb_i\|.$$

In order to estimate the ith term here, we denote by z_i the common centre of the cube Q_i and the ball B_i. Now the integral representation of $Tb_i(s)$ is available at $s \in \complement B_i$. Explicitly, for each $y^* \in Y^*$,

$$\langle Tb_i(s), y^* \rangle = \int \langle K(s,t) b_i(t), y^* \rangle \,\mathrm{d}t = \int \langle [K(s,t) - K(s,z_i)] b_i(t), y^* \rangle \,\mathrm{d}t,$$

where the last step follows from the fact that $\int b_i(t) \,\mathrm{d}t = 0$. Thus

$$\|Tb_i(s)\|_Y \leqslant \int_{Q_i} \|[K(s,t) - K(s,z_i)] b_i(t)\|_Y \,\mathrm{d}t$$

and hence

$$\int_{\complement B_i} \|Tb_i(s)\|_Y \,\mathrm{d}s \leqslant \int_{Q_i} \int_{\complement B_i} \|[K(s,t) - K(s,z_i)] b_i(t)\|_Y \,\mathrm{d}s \,\mathrm{d}t$$

$$\leqslant \int_{Q_i} \|K\|_{\text{Hör}} \|b_i(t)\|_X \, \mathrm{d}t,$$

since $|s - z_i| \geqslant 2 \operatorname{diam}(Q_i) \geqslant 2|t - z_i|$ for $s \in \complement B_i$ and $t \in Q_i$. Substituting back, it follows that

$$\frac{2}{\lambda} \sum_i \int_{\complement B_i} \|Tb_i\| \leqslant \frac{2}{\lambda} \|K\|_{\text{Hör}} \sum_i \int_{Q_i} \|b_i\| = \frac{2}{\lambda} \|K\|_{\text{Hör}} \|b\|_1 \leqslant \frac{4}{\lambda} \|K\|_{\text{Hör}} \|f\|_1.$$

It remains to estimate $|\{\|Tg\| > \lambda/2\}|$. If $p_0 < \infty$, we have

$$|\{\|Tg\| > \lambda/2\}| \leqslant \frac{A_0^{p_0}}{(\lambda/2)^{p_0}} \|g\|_{p_0}^{p_0} \leqslant \frac{2^{p_0}}{\lambda^{p_0}} A_0^{p_0} \cdot (2^d \alpha \lambda)^{p_0 - 1} \|f\|_1,$$

so that altogether

$$|\{\|Tf\| > \lambda\}| \leqslant \Big(\frac{(2A_0 \cdot 2^d \alpha)^{p_0}}{2^d \alpha} + 4\|K\|_{\text{Hör}} + \frac{c_d}{\alpha} \Big) \frac{\|f\|_1}{\lambda},$$

where we are still free to choose $\alpha > 0$. Taking

$$\alpha = 2^{-d-1}/A_0 \tag{11.12}$$

leads to

$$|\{\|Tf\| > \lambda\}| \leqslant (c_d A_0 + 4\|K\|_{\text{Hör}}) \frac{\|f\|_1}{\lambda}. \tag{11.13}$$

If $p_0 = \infty$, we observe that $\|Tg\|_\infty \leqslant A_0 \|g\|_\infty \leqslant A_0 2^d \alpha \lambda$, so that the same choice of α guarantees that $|\{\|Tg\| > \lambda/2\}| = 0$. Thus, in this case, we only need to estimate the last two terms in (11.11), and these have exactly the same bounds in the case $p_0 < \infty$ that was already handled.

We have hence confirmed (11.13) for all $f \in L^{p_0}(\mathbb{R}^d; X) \cap L^1(\mathbb{R}^d; X)$ and $\lambda > 0$, and this proves the existence of a unique bounded extension $T \in \mathscr{L}(L^1(\mathbb{R}^d; X), L^{1,\infty}(\mathbb{R}^d; X))$ by the density of $L^{p_0}(\mathbb{R}^d; X) \cap L^1(\mathbb{R}^d; X)$ in $L^1(\mathbb{R}^d; X)$. This completes the proof of (1b).

(1b) *in case* (4)*:* Let then $\mathbb{R}^d$ be replaced by a cube Q_0. Note that

$$\|Tf\|_{L^{1,\infty}(Q_0;Y)} := \sup_{\lambda > 0} \lambda |Q_0 \cap \{|Tf| > \lambda\}|.$$

If $\lambda \leqslant 2A_0 \fint_{Q_0} \|f\|$, then

$$\lambda |Q_0 \cap \{|Tf| > \lambda\}| \leqslant 2A_0 \fint_{Q_0} \|f\| \times |Q_0| = 2A_0 \|f\|_1 \tag{11.14}$$

If $\lambda > 2A_0 \fint_{Q_0} \|f\|$ and α is as in (11.12), then

$$\alpha \lambda > 2^{-d} \fint_{Q_0} \|f\|$$

is in the admissible range to have Calderón–Zygmund decomposition at level $\alpha\lambda$ fully localised within the cube Q_0 (Proposition 11.2.6). Thus, the earlier argument for the full space $\mathbb{R}^d$ localises to Q_0 to produce the same conclusion (11.13), but with the integral defining $\|K\|_{\text{Hör}}$ restricted to Q_0 only. A combination with (11.14) shows that this estimate holds for all $\lambda > 0$, and hence we have the desired weak-type bound on Q_0.

The case of a quadrant S is an immediate variant of the case of $\mathbb{R}^d$, since Proposition 11.2.6 guarantees that the Calderón–Zygmund decomposition is localised to this quadrant for all values of the level parameter.

(1a)*:* A direct application of Marcinkiewicz Interpolation Theorem 2.2.3 (with 1 in place of p_0, and p_0 in place of p_1) shows that

$$\|T\|_{\mathscr{L}(L^p(\mathbb{R}^d;X),L^p(\mathbb{R}^d;Y)} \leqslant c(\theta,1,p_0)\Big(\frac{c_d(A_0+\|K\|_{\text{Hör}})}{1-\theta}\Big)^{1-\theta}\Big(\frac{A_0}{\theta}\Big)^{\theta},$$

where $\theta \in (0,1)$ is such that $1/p = (1-\theta)/1 + \theta/p_0$,

$$c(\theta,1,p_0) = \Big\{p_0^{\frac{p-1}{p_0-1}}\frac{p_0-p}{(p_0-p)(p-1)}\Big\}^{\frac{1}{p}}$$

if $p_0 \in (1,\infty)$, and $c(\theta,1,\infty) = (p-1)^{-\frac{1}{p}}$. By the arithmetic–geometric mean inequality, we have

$$\Big(\frac{1}{1-\theta}\Big)^{1-\theta}\Big(\frac{1}{\theta}\Big)^{\theta} \leqslant 1-\theta\frac{1}{1-\theta}+\theta\frac{1}{\theta} = 2, \tag{11.15}$$

and by elementary calculus one verifies that $p_0^{\frac{1}{p_0-1}} \leqslant e$ for $p_0 \in (1,\infty)$. Substituting these estimates, we obtain

$$\|T\|_{\mathscr{L}(L^p(\mathbb{R}^d;X),L^p(\mathbb{R}^d;Y)} \leqslant 2e\cdot c_d\cdot\Big\{\frac{p_0-p}{(p_0-p)(p-1)}\Big\}^{\frac{1}{p}}(A_0+\|K\|_{\text{Hör}}),$$

which coincides with the claim after redefining c_d. Since the Marcinkiewicz Interpolation Theorem 2.2.3 is valid for general measure spaces, the same argument applies equally well in the case of a cube or a quadrant as the underlying domain. □

11.2.b Local oscillations of Tf and case $p \in (p_0,\infty)$

We next turn to the study of extrapolation of the boundedness to $p > p_0$, which will involve the dual Hörmander condition. A reader familiar with the scalar-valued counterpart of the theory might expect a duality argument at this point. While this might not be strictly out of question here, either, one should note that at least some number of technicalities would have to be tackled by such an approach. To begin with, the adjoint of $T \in \mathscr{L}(L^p(\mathbb{R}^d;X), L^p(\mathbb{R}^d;Y))$ would be an operator

$$T^* \in \mathscr{L}(L^p(\mathbb{R}^d;Y)^*, L^p(\mathbb{R}^d;X)^*),$$

where each $L^p(\mathbb{R}^d;Z)^*$ is in general a larger space than $L^{p'}(\mathbb{R}^d;Z^*)$, unless additional assumptions are imposed on Z^* (see Section 1.3). Rather than dwelling into such issues, we prefer a direct approach within the original spaces of X and Y valued functions that we are interested in.

We still need to settle a technical issue about the validity of the integral representation of $Tf(x)$ for certain non-compactly supported functions f:

Lemma 11.2.7. *Let X and Y be Banach spaces and $p_0 \in [1,\infty]$. Let $T \in \mathscr{L}(L^{p_0}(\mathbb{R}^d;X), L^{p_0,\infty}(\mathbb{R}^d;Y))$ be an operator with dual Hörmander kernel K. If $B \subseteq \mathbb{R}^d$ is a ball and $f \in L^{p_0}(\mathbb{R}^d;X) \cap L^\infty(\mathbb{R}^d;X)$ is supported in $\complement B$, then for almost all $s, s' \in \frac{1}{2}B$, we have*

$$\langle Tf(s) - Tf(s'), y^* \rangle = \int_{\complement B} \langle [K(s,t) - K(s',t)]f(t), y^* \rangle \,\mathrm{d}t \qquad \forall y^* \in Y^*.$$

Proof. Consider an increasing sequence of balls $B_1 \subseteq B_2 \subseteq \dots$ such that $\bigcup_{n=1}^\infty B_n = \mathbb{R}^d$, and let $f_n := \mathbf{1}_{B_n} f$. Since $f_n = \mathbf{1}_{\complement B} f_n$ is compactly supported away from B, for almost every $s \in \frac{1}{2}B$ we have

$$\langle Tf_n(s), y^* \rangle = \int_{\complement B} \langle K(s,t) f_n(t), y^* \rangle \,\mathrm{d}t \qquad \forall y^* \in Y^*.$$

Thus, for almost every $s, s' \in \frac{1}{2}B$, the following holds for every $y^* \in Y^*$:

$$\langle Tf_n(s) - Tf_n(s'), y^* \rangle = \int_{\complement B} \langle f_n(t), [K(s,t)^* - K(s',t)^*] y^* \rangle \,\mathrm{d}t. \tag{11.16}$$

Now consider the limit $n \to \infty$. Since $f_n \to f$ in $L^{p_0}(\mathbb{R}^d;X)$ and $T \in \mathscr{L}(L^{p_0}(\mathbb{R}^d;X), L^{p_0,\infty}(\mathbb{R}^d;Y))$, we have $Tf_n \to Tf$ in $L^{p_0,\infty}(\mathbb{R}^d;Y)$. Hence a subsequence, which we keep denoting simply by f_n, also satisfies $Tf_n(s) \to Tf(s)$ at almost every $s \in \frac{1}{2}B$. This means that

$$LHS(11.16) \to \langle Tf(s) - Tf(s'), y^* \rangle.$$

It is also clear that $f_n(t) \to f(t)$ pointwise. On the other hand, the integrand in (11.16) is pointwise dominated by

$$(\|[K(s,t)^* - K(z_B,t)^*]y^*\|_{Y^*} + \|[K(s',t)^* - K(z_B,t)^*]y^*\|_{Y^*})\|f\|_\infty,$$

which is integrable over $t \in \complement B$ (thus $|t - z_B| \geqslant r_B \geqslant 2\max\{|s - z_B|, |s' - z_B|\}$) by the dual Hörmander condition. Hence

$$RHS(11.16) \to \int_{\complement B} \langle f(t), [K(s,t)^* - K(s',t)^*] y^* \rangle \,\mathrm{d}t$$

by dominated convergence. The equality of the limits is what we claimed. □

Recall the John–Strömberg maximal function and the local oscillations

$$M^{\#}_{0,\lambda} f(x) = \sup_{Q \ni x} \operatorname{osc}_\lambda(f;Q), \quad \operatorname{osc}_\lambda(f;Q) := \inf_{c\in X} \inf_{|E|\leqslant \lambda|Q|} \|(f-c)\mathbf{1}_{Q\setminus E}\|_\infty.$$

The following lemma contains the technical core of the upper extrapolation:

Lemma 11.2.8. *Under the assumptions of Theorem 11.2.5(2), for all* $f \in L^{p_0}(\mathbb{R}^d;X) \cap L^\infty(\mathbb{R}^d;X)$ *we have*

$$\|M^{\#}_{0,\lambda}(Tf)\|_\infty \leqslant (c_{d,\lambda}^{1/p_0} A_0 + 2\|K\|_{\text{Hör}^*})\|f\|_\infty.$$

If $\mathbb{R}^d$ *is replaced by a cube* $Q_0 \subseteq \mathbb{R}^d$ *or a quadrant* $S \subseteq \mathbb{R}^d$*, the conclusion remains valid with the following modifications:*

(a) *in the maximal operator* $M^{\#}_{0,\lambda}$*, the supremum is restricted to cubes* Q *contained in the initial cube* Q_0 *or the quadrant* S*;*
(b) *in the Hörmander norm* $\|K\|_{\text{Hör}^*}$*, the variables and the integrals are again restricted to* Q_0 *or* S*.*

Proof. Let $f \in L^{p_0}(\mathbb{R}^d;X) \cap L^\infty(\mathbb{R}^d;X)$ and let $Q \subseteq \mathbb{R}^d$ be a cube. Let B be a ball with the same centre and three time the diameter. We decompose

$$Tf = T(\mathbf{1}_B f) + [T(\mathbf{1}_{\complement B} f) - T(\mathbf{1}_{\complement B} f)(z)] + c,$$

where $c = T(\mathbf{1}_{\complement B} f)(z)$, and $z \in Q$ is fixed as one of the (almost all) points of Q where the conclusion of Lemma 11.2.7 is valid for the function $\mathbf{1}_{\complement B} f$. Thus

$$\|(Tf-c)\mathbf{1}_{Q\setminus E}\|_\infty \leqslant \|T(\mathbf{1}_B f)\mathbf{1}_{Q\setminus E}\|_\infty + \|[T(\mathbf{1}_{\complement B} f) - T(\mathbf{1}_{\complement B} f)(z)]\mathbf{1}_Q\|_\infty.$$

For the first term, we observe that

$$\|T(\mathbf{1}_B f)\|_{L^{p_0,\infty}} \leqslant A_0\|\mathbf{1}_B f\|_{p_0} \leqslant A_0|B|^{1/p_0}\|f\|_\infty,$$

and hence

$$|E_\Lambda| := |\{\|T(\mathbf{1}_B f)\| > \Lambda\}| \leqslant c_d\Big(\frac{A_0\|f\|_\infty}{\Lambda}\Big)^{p_0}|Q| \leqslant \lambda|Q|$$

if we choose $\Lambda := (c_d/\lambda)^{1/p_0} A_0\|f\|_\infty$. We conclude that

$$\|T(\mathbf{1}_B f)\mathbf{1}_{Q\setminus E_\Lambda}\|_\infty \leqslant (c_d/\lambda)^{1/p_0} A_0\|f\|_\infty.$$

For the other term, we estimate pointwise at almost every $s \in Q$ where the conclusion of Lemma 11.2.7 is valid. Recalling that $z \in Q$ was also chosen in this way and dualising against $y^* \in Y^*$, we get

$$|\langle T(\mathbf{1}_{\complement B} f)(s) - T(\mathbf{1}_{\complement B} f)(z), y^*\rangle| = \Big|\int_{\complement B}\langle f(t), [K(s,t)^* - K(z,t)^*]y^*\rangle\,\mathrm{d}t\Big|$$

$$\leqslant \int_{\complement B} \|[K(s,t)^* - K(z,t)^*]y^*\|_{X^*} \,\mathrm{d}t \|f\|_\infty$$
$$\leqslant 2\|K\|_{\mathrm{Hör}^*} \|y^*\|_{X^*} \|f\|_\infty.$$

Taking the supremum over y^* in the unit ball of Y^* and the essential supremum over $s \in Q$, we arrive at

$$\|\mathbf{1}_Q[T(\mathbf{1}_{\complement B} f) - T(\mathbf{1}_{\complement B} f)(z)]\|_\infty \leqslant 2\|K\|_{\mathrm{Hör}^*} \|f\|_\infty.$$

Hence altogether

$$\mathrm{osc}_\lambda(Tf;Q) \leqslant \|(Tf-c)\mathbf{1}_{Q\setminus E_\Lambda}\|_\infty \leqslant (c_d/\lambda)^{1/p_0} A_0 \|f\|_\infty + 2\|K\|_{\mathrm{Hör}^*}\|f\|_\infty,$$

and taking the supremum over all $Q \subseteq \mathbb{R}^d$ proves the lemma.

The modifications in the case of a cube Q_0 or a quadrant S in place of $\mathbb{R}^d$ are immediate by inspection. We note that the balls B featuring in the argument may extend beyond Q_0 or S; one simply thinks of $B \cap Q_0$ or $B \cap S$ in this case, while the complement $\complement B$ will be replaced by $Q_0 \setminus B$ or $S \setminus B$, respectively. □

Proof of Theorem 11.2.5(2a). Let us first consider the mapping properties of the sub-linear operator $M^{\#}_{0,\lambda} \circ T$, where $\lambda = 2^{-2-d}$.

By assumption, $T : L^{p_0}(\mathbb{R}^d;X) \to L^{p_0,\infty}(\mathbb{R}^d;Y)$ is bounded (with norm A_0), and Proposition 11.1.21 gives the boundedness of $M^{\#}_{0,\lambda} : L^{p_0,\infty}(\mathbb{R}^d;Y) \to L^{p_0,\infty}(\mathbb{R}^d)$ (with norm bounded by $c_{d,\lambda}^{1/p_0} \leqslant c_d$, since λ depends only on d, and $1/p_0 \leqslant 1$); thus the composition $M^{\#}_{0,\lambda} \circ T : L^{p_0}(\mathbb{R}^d;X) \to L^{p_0,\infty}(\mathbb{R}^d)$ is also bounded (with norm at most $c_d A_0$).

On the other hand, the previous Lemma 11.2.8 says that $M^{\#}_{0,\lambda} \circ T : L^{p_0}(\mathbb{R}^d;X) \cap L^\infty(\mathbb{R}^d;X) \to L^\infty(\mathbb{R}^d)$ is bounded (with norm at most $c_{d,\lambda}^{1/p_0} A_0 + \|K\|_{\mathrm{Hör}^*} \leqslant c_d A_0 + \|K\|_{\mathrm{Hör}^*}$), where the subspace $L^{p_0}(\mathbb{R}^d;X) \cap L^\infty(\mathbb{R}^d;X) \subseteq L^\infty(\mathbb{R}^d;X)$ is equipped with the norm of $L^\infty(\mathbb{R}^d;X)$.

This is essentially a setting to apply the Marcinkiewicz Interpolation Theorem 2.2.3: by inspection, one checks that the relaxed assumption

$$M^{\#}_{0,\lambda} \circ T : L^{p_0}(\mathbb{R}^d;X) \cap L^\infty(\mathbb{R}^d;X) \to L^\infty(\mathbb{R}^d)$$

(in place of $M^{\#}_{0,\lambda} \circ T : L^\infty(\mathbb{R}^d;X) \to L^\infty(\mathbb{R}^d)$) allows us to deduce the relaxed conclusion

$$M^{\#}_{0,\lambda} \circ T : L^{p_0}(\mathbb{R}^d;X) \cap L^p(\mathbb{R}^d;X) \to L^p(\mathbb{R}^d), \qquad p \in (p_0,\infty), \tag{11.17}$$

where $L^{p_0}(\mathbb{R}^d;X) \cap L^p(\mathbb{R}^d;X) \subseteq L^p(\mathbb{R}^d;X)$ is equipped with the norm of $L^p(\mathbb{R}^d;X)$. In fact, the proof of the Marcinkiewicz Interpolation Theorem 2.2.3 is based on decomposing a function f in the domain space into the two truncations, at varying level t,

$$\widetilde{f}^t := \big(f - t\frac{f}{\|f\|}\big) \cdot \mathbf{1}_{\{\|f\|>t\}},$$
$$\widetilde{f}_t := f \cdot \mathbf{1}_{\{\|f\|\leqslant t\}} + t\frac{f}{\|f\|} \cdot \mathbf{1}_{\{\|f\|>t\}},$$

and it is immediate to verify that, if $f \in L^{p_0}(\mathbb{R}^d;X)$, these remain in the space $L^{p_0}(\mathbb{R}^d;X)$, in addition to the other function space memberships used in the proof of Theorem 2.2.3.

If $\theta \in (0,1)$ is such that $1/p = (1-\theta)/p_0 + \theta/\infty = (1-\theta)/p_0$, the Marcinkiewicz Interpolation Theorem 2.2.3 shows that the norm of the operator in (11.17) is at most

$$\begin{aligned} c(\theta,p_0,\infty)\Big(\frac{\|M^{\#}_{0,\lambda}\circ T\|_{L^{p_0}(\mathbb{R}^d;X)\to L^{p_0,\infty}(\mathbb{R}^d)}}{1-\theta}\Big)^{1-\theta} \times \\ \times \Big(\frac{\|M^{\#}_{0,\lambda}\circ T\|_{(L^{p_0}\cap L^\infty)(\mathbb{R}^d;X)\to L^\infty(\mathbb{R}^d)}}{\theta}\Big)^{\theta} \\ \leqslant c(\theta,p_0,\infty)\Big(\frac{c_dA_0}{1-\theta}\Big)^{1-\theta}\Big(\frac{c_dA_0+\|K\|_{\text{Hör}^*}}{\theta}\Big)^{\theta} \\ \leqslant c(\theta,p_0,\infty)\cdot 2\cdot (c_dA_0+\|K\|_{\text{Hör}^*}) \end{aligned}$$

by the arithmetic–geometric mean inequality (11.15) in the last step. Moreover, still from Theorem 2.2.3 and the identity $\Gamma(x+1) = x\Gamma(x)$,

$$c(\theta,p_0,\infty) = \Big\{\frac{\Gamma(p-p_0)\Gamma(p_0+1)}{\Gamma(p)}\Big\}^{1/p} = \{p_0B(p-p_0,p_0)\}^{1/p},$$

where the beta function is

$$\begin{aligned} B(p-p_0,p_0) = \frac{\Gamma(p-p_0)\Gamma(p_0)}{\Gamma(p)} &= \int_0^1 u^{p-p_0-1}(1-u)^{p_0-1}\,\mathrm{d}u \\ &\leqslant \int_0^1 u^{p-p_0-1}\,\mathrm{d}u = \frac{1}{p-p_0}, \end{aligned}$$

since $p_0 \geqslant 1$ here. Substituting back (and redefining c_d), we find that the norm of the operator in (11.17) is at most

$$\Big(\frac{p_0}{p-p_0}\Big)^{1/p}(c_dA_0 + 2\|K\|_{\text{Hör}^*}).$$

Now Theorem 11.1.18, together with Remark 11.1.19 and the *a priori* condition that $Tf \in L^{p_0,\infty}(\mathbb{R}^d;X)$, show that

$$\begin{aligned} \|Tf\|_{L^p(\mathbb{R}^d;Y)} &\leqslant 8p\|M^{\#}_{0,\lambda}(Tf)\|_{L^p(\mathbb{R}^d)} \\ &\leqslant p\Big(\frac{p_0}{p-p_0}\Big)^{1/p}(c_dA_0+16\|K\|_{\text{Hör}^*})\|f\|_{L^p(\mathbb{R}^d;X)} \end{aligned}$$

for all $f \in L^p(\mathbb{R}^d;X) \cap L^{p_0}(\mathbb{R}^d;X)$ and $p \in (p_0,\infty)$. Since this is a dense subspace of $L^p(\mathbb{R}^d;X)$, the operator T has a unique extension to this space, with the same norm estimate above.

The case of a cube or a quadrant in place of $\mathbb{R}^d$ follows by the same argument, since all results quoted are also valid in these settings. □

It is also immediate from Lemma 11.2.8 and Proposition 11.1.24 that

$$\|Tf\|_{\mathrm{BMO}(\mathbb{R}^d;X)} \leqslant 8\|M^{\#}_{0,\lambda}(Tf)\|_{L^\infty(\mathbb{R}^d)} \leqslant (c_d A_0 + 8\|K\|_{\mathrm{H\ddot{o}r}^*})\|f\|_{L^\infty(\mathbb{R}^d;Y)}$$

for all $f \in L^{p_0}(\mathbb{R}^d;X) \cap L^\infty(\mathbb{R}^d;X)$. Since this is *not* a dense subspace of $L^\infty(\mathbb{R}^d;X)$, extending this estimate, and indeed the very meaning of "Tf", to all $f \in L^\infty(\mathbb{R}^d;X)$ requires an additional effort, to which we turn in Section 11.2.c below.

Proof of Theorem 11.2.5(3). We now assume that K satisfies both Hörmander and dual-Hörmander conditions, and hence we have access to both cases (1) and (2) that we already proved. By Theorem 11.2.5(1b), we have

$$\|T\|_{\mathscr{L}(L^1(\mathbb{R}^d;X),L^{1,\infty}(\mathbb{R}^d;Y)} \leqslant c_d(A_0 + \|K\|_{\mathrm{H\ddot{o}r}}).$$

We now use this estimate (rather than the original assumption) as input to Theorem 11.2.5(2a), i.e., we apply the latter with 1 in place of p_0 and $c_d(A_0 + \|K\|_{\mathrm{H\ddot{o}r}})$ in place of A_0. This gives, for all $p \in (1,\infty)$, the estimate

$$\begin{aligned}\|T\|_{\mathscr{L}(L^p(\mathbb{R}^d;X),L^p(\mathbb{R}^d;Y)} &\leqslant c_d p\Big(\frac{1}{p-1}\Big)^{1/p}\Big(c_d(A_0 + \|K\|_{\mathrm{H\ddot{o}r}}) + \|K\|_{\mathrm{H\ddot{o}r}^*}\Big)\\ &\leqslant c_d^2 pp'(A_0 + \|K\|_{\mathrm{H\ddot{o}r}} + \|K\|_{\mathrm{H\ddot{o}r}^*}),\end{aligned}$$

where we estimated

$$\Big(\frac{1}{p-1}\Big)^{1/p} \leqslant \Big(\frac{p}{p-1}\Big)^{1/p} = (p')^{1/p} \leqslant p'.$$

The conclusion agrees with the claim, after redefining c_d.

The case of a cube or a quadrant in place of $\mathbb{R}^d$ is immediate, since both (1) and (2) of the theorem, which we used above, were already proved in these cases as well. □

11.2.c The action of singular integrals on L^∞

The goal of this section is to establish the following theorem, in which indistinguishability of $\mathrm{BMO}(\mathbb{R}^d;X)$ functions only differing by an additive constant manifests itself.

Theorem 11.2.9. *Let X and Y be Banach spaces, $p_0 \in (1,\infty)$, and $T \in \mathscr{L}(L^{p_0}(\mathbb{R}^d;X), L^{p_0}(\mathbb{R}^d;Y))$ be an operator with a dual Hörmander kernel K. Suppose, moreover, at least one of the following:*

(1) *Y does not contain a copy of c_0, or*
(2) *K is a dual operator-Hörmander kernel.*

Then there is an operator $\widetilde{T} \in \mathscr{L}(L^\infty(\mathbb{R}^d;X), \mathrm{BMO}^{p_0}(\mathbb{R}^d;Y)/Y)$ of norm at most $(c_d A_0 + \|K\|_{\mathrm{H\ddot{o}r}^*})$ such that

(a) *for all $f \in L^{p_0}(\mathbb{R}^d;X) \cap L^\infty(\mathbb{R}^d;X)$, we have $Tf \equiv \widetilde{T}f$ modulo constants,*
(b) *for all $f \in L^\infty(\mathbb{R}^d;X)$ and $g \in L^\infty_{c,0}(\mathbb{R}^d;Y^*)$ (compactly supported with vanishing integral), we have*

$$\langle \widetilde{T}f, g\rangle = \lim_{n\to\infty} \langle T(\mathbf{1}_{E_n} f), g\rangle \tag{11.18}$$

for any bounded measurable sets $E_n \subseteq \mathbb{R}^d$ such that $\mathrm{dist}(\complement E_n, 0) \to \infty$.

Remark 11.2.10.

(1) By the John–Nirenberg inequality, the target space $\mathrm{BMO}^p(\mathbb{R}^d;Y)/Y$ of $\widetilde{T}$ is independent of the value of $p \in [1,\infty)$; however, the estimate for the operator norm need not be, and we specifically state it with $p = p_0$.
(2) The left-hand side of (11.18) could be more pedantically written as "$\langle h, g\rangle$, where $h \in [\widetilde{T}f]$ is arbitrary": the vanishing integral of g guarantees that this expression is independent of the choice of h.
(3) The boundedness requirement on T in Theorem 11.2.9 may seem stronger than in Theorem 11.2.5(2) (where it was only assumed that T maps boundedly into the larger space $L^{p_0,\infty}(\mathbb{R}^d;Y)$ and for some p_0 in the larger range $[1,\infty]$), but this is only superficial, as we can always arrange ourselves to be in the situation of Theorem 11.2.9 even under the apparently weaker boundedness hypothesis:
First, if $p_0 = \infty$, there is nothing to prove, as we can simply take $\widetilde{T} = T$, which already maps into $L^\infty(\mathbb{R}^d;Y) \subseteq \mathrm{BMO}(\mathbb{R}^d;Y)$. If, on the other hand, $p_0 \in [1,\infty)$, Theorem 11.2.5(2a) guarantees that $T \in \mathscr{L}(L^p(\mathbb{R}^d;X), L^p(\mathbb{R}^d;Y))$ for all $p \in (p_0,\infty) \subseteq (1,\infty)$, and choosing one such p as a new p_0, we are in the situation assumed in Theorem 11.2.9.

To deal with the equivalence classes modulo additive constants, it is convenient to make the following preliminary observation:

Lemma 11.2.11. *Let S be a set and X be a Banach spaces. There is a bijective linear correspondence between the following two classes of objects:*

(1) *equivalence classes $[b]$ of functions $b : S \to X$, where*

$$[b] = \{f : S \to X; s \mapsto f(s) - b(s) \text{ is constant on } S\},$$

(2) *functions $\Delta : S \times S \to X$ with the property*

$$\Delta(s,t) + \Delta(t,u) = \Delta(s,u) \qquad \forall\ s,t,u \in S. \tag{11.19}$$

This correspondence is realised by

$$[s \mapsto b(s)] \qquad \leftrightarrow \qquad (s,t) \mapsto \Delta(s,t) = b(s) - b(t).$$

Proof. To every $[b]$, we associate $\Delta(s,t) := b(s) - b(t)$, and it is clear that this is independent of the chosen representative of the equivalence class.

For the other direction, it is convenient to first record some additional algebraic relations automatically satisfied by Δ. Taking $s = t = u$, we have $2\Delta(s,s) = \Delta(s,s)$, and hence $\Delta(s,s) = 0$ for all $s \in S$. Then taking $u = s$, we have $\Delta(s,t) + \Delta(t,s) = \Delta(s,s) = 0$, and hence $\Delta(s,t) = -\Delta(t,s)$ for all $s,t \in S$. Now, to every Δ, we associate $[\Delta(\cdot,t)]$, where each $t \in S$ defines the same equivalence class. Indeed,

$$\Delta(s,t) - \Delta(s,u) = \Delta(u,s) + \Delta(s,t) = \Delta(u,t),$$

which is constant as a function of $s \in S$. It is immediate to verify that these operations sending $[b]$ to Δ, and Δ to $[b]$, are inverses of each other. $\square$

For $S \subseteq \mathbb{R}^d$ (where we are mainly interested in the case that $S = \mathbb{R}^d$ or one of its dyadic quadrants), we define

$$\widetilde{\mathrm{BMO}}^p(S;X) := \Big\{ \Delta \in L^1_{\mathrm{loc}}(S \times S; X) \text{ with property (11.19)}, \\ \|\Delta\|_{*,p} := \sup_{\substack{Q \subseteq S \\ \text{cube}}} \Big(\fint_{Q\times Q} \|\Delta(s,t)\|_X^p \,\mathrm{d}s \,\mathrm{d}t \Big)^{1/p} < \infty \Big\}$$

and $\widetilde{\mathrm{BMO}}^p_{\mathscr{D}}(S;X)$ by replacing "$Q \subseteq S$ cube" by "$Q \in \mathscr{D}(S)$".

Lemma 11.2.12. *Under the correspondence $[b] \leftrightarrow \Delta$ of functions as in Lemma 11.2.11, we have the correspondence of spaces:*

$$\mathrm{BMO}^p(\mathbb{R}^d;X)/X \simeq \widetilde{\mathrm{BMO}}^p(\mathbb{R}^d;X),$$

with the equivalence of norms

$$\|b\|_{\mathrm{BMO}^p(\mathbb{R}^d;X)} \leqslant \|\Delta\|_{*,p} \leqslant 2\|b\|_{\mathrm{BMO}^p(\mathbb{R}^d;X)}. \tag{11.20}$$

The similar correspondence is valid with any of the dyadic quadrants S in place of $\mathbb{R}^d$ and the dyadic $\mathrm{BMO}^p_{\mathscr{D}}$ (both with and without tilde) in place of BMO^p.

Proof. For each cube $Q \subseteq \mathbb{R}^d$, we have

$$\begin{aligned} \inf_{c\in X} \Big(\fint_Q \|b(s) - c\|_X^p \,\mathrm{d}s \Big)^{1/p} &\leqslant \Big(\fint_Q \Big\| b(s) - \fint_Q b(t)\,\mathrm{d}t \Big\|_X^p \,\mathrm{d}s \Big)^{1/p} \\ &\leqslant \Big(\fint_Q \fint_Q \|b(s) - b(t)\|_X^p \,\mathrm{d}s\,\mathrm{d}t \Big)^{1/p} \\ &= \Big(\fint_Q \fint_Q \Big\| (b(s) - c) - (b(t) - c) \Big\|_X^p \,\mathrm{d}s\,\mathrm{d}t \Big)^{1/p} \leqslant 2\Big(\fint_Q \|b(s) - c\|_X^p \,\mathrm{d}s \Big)^{1/p} \end{aligned}$$

and taking the infimum over $c \in X$ on the right, and then the supremum over all cubes $Q \subseteq \mathbb{R}^d$ of the whole chain, we derive (11.20). The dyadic case follows by taking the supremum over $Q \in \mathscr{D}(S)$ instead. □

In view of Lemma 11.2.12, the construction of an extension

$$\widetilde{T} \in \mathscr{L}(L^\infty(\mathbb{R}^d; X), \mathrm{BMO}(\mathbb{R}^d; Y))$$

of $T \in \mathscr{L}(L^{p_0}(\mathbb{R}^d; X), L^{p_0}(\mathbb{R}^d; Y))$ is reduced to the construction of $\Delta_T \in \mathscr{L}(L^\infty(\mathbb{R}^d; X), \widetilde{\mathrm{BMO}}(\mathbb{R}^d; X))$ such that

$$\Delta_T f(s,u) = Tf(s) - Tf(u) \qquad \forall\, f \in L^{p_0}(\mathbb{R}^d; X) \cap L^\infty(\mathbb{R}^d; X).$$

It is convenient to define this as *a priori* mapping into Y^{**}-valued functions:

Lemma 11.2.13. *For $f \in L^\infty(\mathbb{R}^d; X)$, $y^* \in Y^*$ and $s, u \in \mathbb{R}^d$, the expression*

$$\begin{aligned} \langle y^*, \Delta_T f(s,u)\rangle := \langle T(\mathbf{1}_B f)(s) - T(\mathbf{1}_B f)(u), y^*\rangle \\ + \int_{\complement B} \langle [K(s,t) - K(u,t)] f(t), y^*\rangle \,\mathrm{d}t, \end{aligned} \tag{11.21}$$

is independent of the auxiliary ball B with $s, u \in \frac{1}{2}B$.

Proof. With f, y^*, s, u fixed, let us temporarily denote the expression of interest by $\delta(B)$. If B and B' are two such balls, we can choose a third such ball B'' that contains both of them. So it is enough to prove the equality $\delta(B) = \delta(B')$ for balls $B \subseteq B'$, hence $\complement B' \subseteq \complement B$. Note that $(\complement B) \setminus (\complement B') = B' \setminus B$. Then

$$\begin{aligned} \delta(B') - \delta(B) = \langle T(\mathbf{1}_{B'\setminus B} f)(s) - T(\mathbf{1}_{B'\setminus B} f)(u), y^*\rangle \\ + \Big(\int_{\complement B'} - \int_{\complement B}\Big) \langle [K(s,t) - K(u,t)] f(t), y^*\rangle \,\mathrm{d}t, \end{aligned}$$

where the difference of the integrals is

$$\int_{B'\setminus B} \langle [K(u,t) - K(s,t)] f(t), y^*\rangle \,\mathrm{d}t = \langle T(\mathbf{1}_{B'\setminus B} f)(u) - T(\mathbf{1}_{B'\setminus B} f)(s), y^*\rangle,$$

which exactly cancels out the first term in the formula of $\delta(B') - \delta(B)$. □

Let us then check how Δ_T compares with the original T on the intersection of their domains of definition:

Lemma 11.2.14. *If $f \in L^{p_0}(\mathbb{R}^d; X) \cap L^\infty(\mathbb{R}^d; X)$, then*

$$\Delta_T f(s,u) = Tf(s) - Tf(u).$$

Proof. Under the stated assumptions, Lemma 11.2.7 guarantees that

$$\int_{\complement B} \langle [K(s,t) - K(u,t)] f(t), y^* \rangle \, \mathrm{d}t = \langle T(\mathbf{1}_{\complement B} f)(s) - T(\mathbf{1}_{\complement B} f)(u), y^* \rangle$$

for almost all $s, u \in \frac{1}{2}B$ and all $y^* \in Y^*$, and hence

$$\begin{aligned} \langle y^*, \Delta_T f(s,u) \rangle &= \langle T(\mathbf{1}_B f)(s) - T(\mathbf{1}_B f)(u), y^* \rangle \\ &\quad + \langle T(\mathbf{1}_{\complement B} f)(s) - T(\mathbf{1}_{\complement B} f)(u), y^* \rangle = \langle Tf(s) - Tf(u), y^* \rangle \end{aligned}$$

Since this is true for all $y^* \in Y^*$, the claimed identity follows. □

To justify that the *a priori* Y^{**}-valued function $\Delta_T f$ actually takes values in Y, we invoke the following corollary of the Bessaga–Pełczyński Theorem 1.2.40. This is where the condition $c_0 \not\subseteq Y$ comes to use:

Proposition 11.2.15. *Let Y be a Banach space that does not contain an isomorphic copy of c_0. If $y_j \in Y$ satisfy*

$$\sum_{j=1}^{\infty} |\langle y_j, y^* \rangle| < \infty \qquad \forall y^* \in Y^*, \tag{11.22}$$

then the series $\sum_{j=1}^{\infty} y_j$ converges in norm in Y.

Proof. Let us first note that the condition (11.22) says that $y^* \mapsto (\langle y_j, y^* \rangle)_{j=1}^{\infty}$ defines a linear operator from Y^* into ℓ^1, which is easily seen to be closed, and therefore bounded. Thus the closed graph theorem improves (11.22) to

$$\sum_{j=1}^{\infty} |\langle y_j, y^* \rangle| \leqslant C \|y^*\|_{Y^*} \qquad \forall y^* \in Y^*.$$

If $\sum_{j=1}^{\infty} y_j$ does not converge, then the partial sums $\sum_{j=1}^{n} y_j$ fail the Cauchy criterion, and hence we can find $m_1 < n_1 < m_2 < \ldots$ and $\delta > 0$ such that

$$\|v_k\|_Y \geqslant \delta > 0, \qquad v_k := \sum_{j=m_k}^{n_k} y_j. \tag{11.23}$$

On the other hand, for any $\epsilon_k = \pm 1$ and any $y^* \in Y^*$, we also have

$$\Big| \Big\langle \sum_{k=1}^{K} \epsilon_k v_k, y^* \Big\rangle \Big| \leqslant \sum_{k=1}^{K} |\langle v_k, y^* \rangle| \leqslant \sum_{j=1}^{\infty} |\langle y_j, y^* \rangle| \leqslant C \|y^*\|_{Y^*};$$

hence

$$\Big\| \sum_{k=1}^{K} \epsilon_k v_k \Big\|_Y \leqslant C. \tag{11.24}$$

But the two conditions (11.23) and (11.24) are precisely those of the Bessaga–Pełczyński Theorem 1.2.40 that guarantee the containment of an isomorphic copy of c_0 in $\overline{\operatorname{span}}(v_k)_{k=1}^{\infty} \subseteq Y$. This contradicts the assumption on Y. □

After this interlude, we return to the main topic of this section:

Lemma 11.2.16. *Under the assumptions of Theorem 11.2.9, for every $f \in L^\infty(\mathbb{R}^d; X)$, the function $\Delta_T f$ in (11.21) is well defined, takes values in $Y \subseteq Y^{**}$, is strongly measurable, and satisfies*

$$\|\Delta_T f\|_{L^{p_0}(Q\times Q;Y)} \leqslant (c_d A_0 + \|K\|_{\text{Hör}^*})\|f\|_\infty |Q|^{2/p_0}$$

for every cube $Q \subseteq \mathbb{R}^d$.

Proof. Let B be the ball concentric with Q and with twice the diameter of Q; hence $Q \subseteq \frac{1}{2}B$. From the assumption that $T \in \mathscr{L}(L^{p_0}(\mathbb{R}^d; X), L^{p_0}(\mathbb{R}^d; Y))$ and $f \in L^\infty(\mathbb{R}^d; X)$, it is immediate that $T(\mathbf{1}_B f) \in L^{p_0}(\mathbb{R}^d; Y)$ and

$$\|T(\mathbf{1}_B f)\|_{p_0} \leqslant A_0 \|\mathbf{1}_B f\|_{p_0} \leqslant A_0 |B|^{1/p_0} \|f\|_\infty,$$

so that

$$\begin{aligned}&\|(s,u) \mapsto T(\mathbf{1}_B f)(s) - T(\mathbf{1}_B f)(u)\|_{L^{p_0}(Q\times Q;Y)} \\ &\leqslant 2|Q|^{1/p_0} \|T(\mathbf{1}_B f)\|_{L^{p_0}(\mathbb{R}^d;Y)} \leqslant c_d A_0 |Q|^{2/p_0} \|f\|_\infty,\end{aligned}$$

The more delicate matter is the integral in (11.21). Certainly this integral exists, since the dual Hörmander condition guarantees that $[K(s,t)^* - K(u,t)^*]y^*$ is jointly measurable and belongs to $L^1(\complement B, \mathrm{d}t; Y^*)$ uniformly in $(s,u) \in Q$, while $f \in L^\infty(\mathbb{R}^d; Y)$ by assumption. An immediate estimate with the dual Hörmander condition shows that this integral is bounded by $\|K\|_{\text{Hör}^*} \|f\|_\infty \|y^*\|_{Y^*}$, uniformly in $x \in Q$, and hence defines a Y^{**}-valued function $h(s,u)$ with the pointwise bound

$$\|h(s,u)\|_{Y^{**}} \leqslant \|K\|_{\text{Hör}^*} \|f\|_\infty. \tag{11.25}$$

What remains is to justify the Y-valuedness and the strong measurability of this weakly defined function. To this end, we write $f_n = \mathbf{1}_{2^n B \setminus 2^{n-1}B} f$, so that $\mathbf{1}_{\complement B} f = \sum_{n\geqslant 1} f_n$, say pointwise. Since each $f_n \in L^{p_0}(\mathbb{R}^d; X) \cap L^\infty(\mathbb{R}^d; X)$, we can apply Lemma 11.2.7 to see that

$$\begin{aligned}&\int_{\complement B} \langle [K(s,t) - K(u,t)] f_n(t), y^* \rangle \,\mathrm{d}t \\ &= \langle T f_n(s) - T f_n(u), y^* \rangle =: \langle h_n(s,u), y^* \rangle\end{aligned}$$

is the pairing of y^* with a Y-valued, strongly measurable function $h_n(s,u)$.

If we denote by h the *a priori* Y^{**}-valued function defined by

$$\langle y^*, h(s,u) \rangle := \int_{\complement B} \langle [K(s,t) - K(u,t)] f(t), y^* \rangle \,\mathrm{d}t,$$

then

$$\langle y^*, h(s,u)\rangle = \sum_{n=1}^{\infty} \langle h_n(s,u), y^*\rangle \qquad \forall y^* \in Y^*. \tag{11.26}$$

If K satisfies the dual operator-Hörmander condition, then

$$\sum_{n=1}^{\infty} \|h_n(s,u)\| \leqslant \int_{\complement B} \|K(s,t) - K(u,t)\| \|f\|_\infty \,\mathrm{d}t \leqslant 2\|K\|_{\text{Hör}^*_{op}} \|f\|_\infty,$$

so the series $\sum_{n=1}^{\infty} h_n(s,u)$ converges absolutely, and hence in norm. Under the mere dual Hörmander condition, but with the assumption that Y does not contain an isomorphic copy of c_0, the needed norm convergence of $\sum_{n=1}^{\infty} h_n(s,u)$ follows by Proposition 11.2.15 and the bound

$$\begin{aligned}\sum_{n=1}^{\infty} |\langle h_n(s,u), y^*\rangle| &\leqslant \int_{\complement B} |\langle f(y), [K(s,t)^* - K(u,t)^*]y^*\rangle| \,\mathrm{d}t \\ &\leqslant 2\|K\|_{\text{Hör}^*} \|y^*\|_{Y^*} \|f\|_\infty < \infty \qquad \forall y^* \in Y^*.\end{aligned}$$

In both cases, by (11.26), the limit of $\sum_{n=1}^{\infty} h_n(s,u)$ must be $h(s,u)$. Thus, as a pointwise limit of Y-valued strongly measurable functions, h itself must be both Y-valued and strongly measurable. Once these qualitative properties are verified, the quantitative $L^{p_0}(Q \times Q; Y)$ estimate is immediate by integrating over $Q \times Q$ the already observed pointwise bound (11.25). ☐

Now we are prepared to complete:

Proof of Theorem 11.2.9. The operator $\varDelta_T : L^\infty(\mathbb{R}^d; X) \to L^{p_0}_{\text{loc}}(\mathbb{R}^{d+d}; Y)$ is well defined by Lemma 11.2.16 and satisfies

$$\|\varDelta_T f\|_{*,p_0} \leqslant (c_d A_0 + \|K\|_{\text{Hör}^*})\|f\|_\infty$$

for the norm defined in Lemma 11.2.12. By Lemma 11.2.12, we obtain a bounded linear operator $\widetilde{T} \in \mathscr{L}(L^\infty(\mathbb{R}^d; X), \text{BMO}^{p_0}(\mathbb{R}^d; Y)/Y)$, with the same norm bound, by setting

$$\widetilde{T}f := [\varDelta_T f(\cdot, u)] \quad \text{(the equivalence class modulo constants)}, \tag{11.27}$$

where the choice of $u \in \mathbb{R}^d$ is irrelevant. By Lemma 11.2.14, we have $\varDelta_T f(s,u) = Tf(s) - Tf(u)$ for $f \in L^{p_0}(\mathbb{R}^d; X) \cap L^\infty(\mathbb{R}^d; X)$, and hence $\widetilde{T}f = [Tf]$ in this case. This completes the proof of Claim (a) of the theorem.

As for Claim (b), we note that pairing a $g \in L^\infty_{c,0}(\mathbb{R}^d; Y^*)$ with an element of $\text{BMO}^{p_0}(\mathbb{R}^d; Y)/Y$ is well defined, and independent of the representative of the equivalence class, since the integral of g against any constant $c \in Y$ will vanish. By the assumptions on E_n, we can choose balls $B_n := B(0, r_n) := B(0, \text{dist}(\complement E_n, 0)) \subseteq E_n$ with $r_n \to \infty$. Let n be so large that $\operatorname{supp} g \subseteq \frac{1}{2}B_n$. Since $\widetilde{T}$ is linear, we have

$$\langle \widetilde{T}f, g\rangle = \langle \widetilde{T}(\mathbf{1}_{E_n} f), g\rangle + \langle \widetilde{T}(\mathbf{1}_{\complement E_n} f), g\rangle =: I_n + II_n.$$

By Claim (a), which we already proved, we have

$$I_n = \langle T(\mathbf{1}_{E_n} f), g\rangle.$$

For II_n, recalling the construction of $\widetilde{T}$ from (11.27) with $u = 0$, and then the definition of $\Delta_T f(s,u)$ from (11.21) with $B = B_n$, we have

$$\begin{aligned} II_n &= \langle \Delta_T(\mathbf{1}_{\complement E_n} f)(\cdot, 0), g\rangle \\ &= \int_{\mathbb{R}^d} \langle \Delta_T(\mathbf{1}_{\complement E_n} f)(s,0), g(s)\rangle \,\mathrm{d}s \\ &= \int_{\mathbb{R}^d} \langle T(\mathbf{1}_{B_n}\mathbf{1}_{\complement E_n} f)(s) - T(\mathbf{1}_{B_n}\mathbf{1}_{\complement E_n} f)(0), g(s)\rangle \,\mathrm{d}s \\ &\qquad + \int_{\mathbb{R}^d}\int_{\complement B_n} \langle [K(s,t) - K(0,t)](\mathbf{1}_{\complement E_n} f)(t), g(s)\rangle \,\mathrm{d}t\,\mathrm{d}s \\ &=: III_n + IV_n = 0 + IV_n, \end{aligned}$$

since $B_n \subseteq E_n$. Finally,

$$|IV_n| \leqslant \|f\|_{L^\infty(\mathbb{R}^d;X)} \int_{\mathbb{R}^d}\int_{\complement B_n} \|[K(s,t) - K(0,t)]^* g(s)\|_{X^*} \,\mathrm{d}t\,\mathrm{d}s.$$

For every fixed $s \in \operatorname{supp} g \subseteq \frac{1}{2}B_n$, the inner integral is bounded by $\|K\|_{\text{Hör}^*}\|g(s)\|_{Y^*}$, and, as $n \to \infty$, it converges to 0 by dominated convergence; the same is also true for $s \notin \operatorname{supp} g$, since both the integral and the upper bound vanish in this case. Thus also the double integral converges to 0 by another application of dominated convergence.

Altogether, we have seen that

$$\langle \widetilde{T} f, g\rangle - \langle T(\mathbf{1}_{E_n} f), g\rangle = II_n = IV_n \to 0 \quad \text{as} \quad n \to \infty,$$

which concludes the proof of the remaining Claim (b) of Theorem 11.2.9. □

11.3 Calderón–Zygmund operators and sparse bounds

The goal of this section is to derive a powerful pointwise domination of Calderón–Zygmund operators by simple averaging operators over sparse families of dyadic cubes; from this domination, norm estimates for Calderón–Zygmund operators in various different spaces follow almost instantly.

The assumptions that we have to make on the kernel of the operator in order to carry out this programme are somewhat stronger than those needed for the L^p extrapolation of the previous section:

Definition 11.3.1 (Calderón–Zygmund kernel). *Let Z be a Banach space, and $K : \dot{\mathbb{R}}^{2d} \to Z$. We define the quantities*

$$c_K := \sup\{|s-t|^d \cdot \|K(s,t)\| : (s,t) \in \dot{\mathbb{R}}^{2d}\},$$

and, for $u \in [0, \frac{1}{2}]$,

$$\begin{aligned}
\omega_K^1(u) &:= \sup\Big\{|s-t|^d \|K(s,t) - K(s',t)\| : |s-s'| \leqslant u|s-t|\Big\},\\
\omega_K^2(u) &:= \sup\Big\{|s-t|^d \|K(s,t) - K(s,t')\| : |t-t'| \leqslant u|s-t|\Big\},\\
\omega_K(u) &:= \max_{i=1,2} \omega_K^i(u).
\end{aligned}$$

For $K \in C^1(\dot{\mathbb{R}}^{2d}; Z)$, *let further*

$$\begin{aligned}
c_K^1 &:= \sup\{|s-t|^{d+1} \|\nabla_s K(s,t)\| : s \neq t\},\\
c_K^2 &:= \sup\{|s-t|^{d+1} \|\nabla_t K(s,t)\| : s \neq t\}.
\end{aligned}$$

We say that a kernel K *with* $c_K < \infty$ *is*

(i) *a* standard kernel *if* $\omega_K(u) \leqslant c_\delta u^\delta$ *for some* $\delta \in (0,1]$,
(ii) *a* Dini kernel *if* ω_K *satisfies the* Dini condition

$$\|\omega_K\|_{\text{Dini}} := \int_0^{1/2} \omega_K(u) \frac{\mathrm{d}u}{u} < \infty,$$

(iii) *a* C^1-Calderón–Zygmund kernel *if* $K \in C^1(\dot{\mathbb{R}}^{2d}; Z)$ *and* $c_K^i < \infty$, $i = 1, 2$,
(iv) *an* ω-Calderón–Zygmund kernel *if* $\omega_K \leqslant \omega$,
(v) *an* (ω_1, ω_2)-Calderón–Zygmund kernel *if* $\omega_K^i \leqslant \omega_i$, $i = 1, 2$.

We also apply these notions to kernels K *defined on* $\{(s,t) : s, t \in S, s \neq t\}$, *where* S *is either a cube or a quadrant of* $\mathbb{R}^d$*; in this case, each supremum above is taken only over the respective domain of definition.*

It is immediate that a standard kernel is a Dini kernel with $\|\omega\|_{\text{Dini}} \leqslant \delta^{-1} c_\delta$.

Remark 11.3.2. For a convolution kernel $K(x,y) = k(x-y)$, we have

$$\begin{aligned}
c_K &= \sup\{|s|^d \|k(s)\| : s \neq 0\},\\
c_K^i &= \sup\{|s|^{d+1} \|\nabla k(s)\| : s \neq 0\}, \qquad i = 1, 2,\\
\omega_K(u) = \omega_K^i(u) &= \sup\{|s|^d \|k(s) - k(s-t)\| : |t| \leqslant u|s|\}, \qquad i = 1, 2,
\end{aligned}$$

with no difference between $i = 1$ and $i = 2$ in the last two formulas.

Lemma 11.3.3.

$$\omega_K^i(\frac{1}{2}) \leqslant (1 + 2^d) c_K, \qquad \sum_{k=2}^{\infty} \omega_K^i(2^{-k}) \leqslant \frac{1}{\log 2} \|\omega_K^i\|_{\text{Dini}}.$$

Proof. If $|t-t'| \leqslant \frac{1}{2}|s-t|$, then $|s-t| \leqslant |s-t'|+|t-t'| \leqslant |s-t'|+\frac{1}{2}|s-t|$, and hence $|s-t| \leqslant 2|s-t'|$. Thus

$$|s-t|^d\|K(s,t)-K(s,t')\| \leqslant c_K + 2^d|s-t'|^d\|K(s,t')\| \leqslant (1+2^d)c_K,$$

and hence $\omega_K^2(\frac{1}{2}) \leqslant (1+2^d)c_K$. The proof for ω_K^1 is entirely similar.

If ω is increasing, which is obviously the case with $\omega = \omega_K^i$, it follows that

$$\omega(2^{-k-1}) \leqslant \omega(u), \quad u \in (2^{-k-1}, 2^{-k}),$$

hence

$$\omega(2^{-k-1}) \log 2 \leqslant \int_{2^{-k-1}}^{2^{-k}} \omega(u) \frac{\mathrm{d}u}{u},$$

and thus

$$\sum_{k=2}^{\infty} \omega(2^{-k}) = \sum_{k=1}^{\infty} \omega(2^{-1-k}) \leqslant \frac{1}{\log 2} \int_0^{1/2} \omega(u) \frac{\mathrm{d}u}{u}.$$

□

Lemma 11.3.4. *For $K : \dot{\mathbb{R}}^{2d} \to Z = \mathscr{L}(X,Y)$, we have:*

(1) *If $\|\omega_K^1\|_{\mathrm{Dini}} < \infty$, then K is a dual operator-Hörmander kernel, and*

$$\|K\|_{\mathrm{Hör}_{\mathrm{op}}^*} \leqslant \sigma_{d-1}\|\omega_K^1\|_{\mathrm{Dini}}.$$

(2) *If $\|\omega_K^2\|_{\mathrm{Dini}} < \infty$, then K is an operator-Hörmander kernel, and*

$$\|K\|_{\mathrm{Hör}_{\mathrm{op}}} \leqslant \sigma_{d-1}\|\omega_K^2\|_{\mathrm{Dini}}.$$

(3) *Every standard kernel is a Dini kernel with*

$$\|\omega\|_{\mathrm{Dini}} \leqslant 2^{d+1}\frac{c_K}{\delta}\Big(1+\log_+ \frac{c_\delta}{2^{d+1}c_K}\Big).$$

(4) *Every C^1-Calderón–Zygmund kernel is a standard kernel with*

$$\omega_K^i(u) \leqslant 2^{d+1}c_K^i \cdot u$$

and a Dini kernel with

$$\|\omega_K^i\|_{\mathrm{Dini}} \leqslant 2^{d+1}c_K\Big(1+\log_+ \frac{c_K^i}{c_K}\Big).$$

Here σ_{d-1} is the $(d-1)$-dimensional measure of the unit sphere in $\mathbb{R}^d$. The same conclusions hold with $\dot{\mathbb{R}}^{2d}$ replaced by $\dot{S}^2 := \{(s,t) : s,t \in S, s \neq t\}$, where S is either a cube or a quadrant of $\mathbb{R}^d$, and both the Dini and the Hörmander conditions are modified by restricting the variables to the respective domain of definition.

Note that, in concrete situations, the constants c_δ or c_K^i are often much larger than c_K. The point of the bounds in parts (3) and (4) is that these larger constants contribute to the Dini bounds only logarithmically.

Proof. We will first prove (2); the proof of (1) is analogous.

$$\begin{aligned}
&\int_{|x-y|>2|y-y'|} \|K(x,y)-K(x,y')\| \,\mathrm{d}x \\
&\leqslant \int_{|x-y|>2|y-y'|} \omega_K^2\Big(\frac{|y-y'|}{|x-y|}\Big)\frac{1}{|x-y|^d}\,\mathrm{d}x \\
&= \sigma_{d-1}\int_{2|y-y'|}^{\infty} \omega_K^2\Big(\frac{|y-y'|}{r}\Big)\frac{\mathrm{d}r}{r} = \sigma_{d-1}\int_0^{\frac12}\omega_K^2(t)d\frac{\mathrm{d}t}{t} = \sigma_{d-1}\|\omega_K^2\|_{\mathrm{Dini}}
\end{aligned}$$

and this is the required bound.

For the remaining claims, we begin with the following observation. For $|x-x'|\leqslant u|x-y|$ and $v\in[0,1]$, we have

$$|x+v(x'-x)-y| \geqslant |x-y|-|x'-x| \geqslant (1-u)|x-y| \geqslant \frac12|x-y|.$$

This implies the crude bound

$$\|K(x',y)-K(x,y)\| \leqslant \frac{c_K}{|x'-y|^d}+\frac{c_K}{|x-y|^d} \leqslant (2^d+1)\frac{c_K}{|x-y|^d} \leqslant \frac{2^{d+1}c_K}{|x-y|^d}.$$

This shows that $\omega_K^i(u)\leqslant 2^{d+1}c_K$ for all $u\in[0,\frac12]$ and $i=1$, and the proof for $i=2$ is similar.

(3): By the previous observation, denoting $c_0 := 2^{d+1}c_K$, the standard estimate $\omega(u)\leqslant c_\delta u^\delta$ bootstraps to $\omega(u)\leqslant\min\{c_0,c_\delta u^\delta\}$. If $c_0\leqslant c_\delta$, then

$$\begin{aligned}
\|\omega\|_{\mathrm{Dini}} &\leqslant \int_0^{(c_0/c_\delta)^{1/\delta}} c_\delta u^\delta\frac{\mathrm{d}u}{u} + \int_{(c_0/c_\delta)^{1/\delta}}^1 c_0\frac{\mathrm{d}u}{u} \\
&= \frac{c_\delta}{\delta}\frac{c_0}{c_\delta} + c_0\log\Big(\frac{c_\delta}{c_0}\Big)^{1/\delta} = \frac{c_0}{\delta}\Big(1+\log\frac{c_\delta}{c_0}\Big).
\end{aligned}$$

If $c_0>c_\delta$, we simply estimate $\|\omega\|_{\mathrm{Dini}}\leqslant\int_0^1 c_\delta u^{\delta-1}\,\mathrm{d}u = c_\delta/\delta\leqslant c_0/\delta$. Hence, in each case, we have

$$\|\omega\|_{\mathrm{Dini}} \leqslant \frac{c_0}{\delta}\Big(1+\log_+\frac{c_\delta}{c_0}\Big).$$

We will prove (4) in the case $i=1$, the case of $i=2$ is analogous. Hence

$$
\begin{aligned}
\|K(x',y)-K(x,y)\| &= \Big\| \Big/_{v=0}^{1} K(x+u(x'-x),y)\Big\| \\
&= \Big\| \int_0^1 (x'-x)\cdot \nabla_x K(x+v(x'-x),y)\,\mathrm{d}v\Big\| \\
&\leqslant |x'-x| \int_0^1 \frac{c_K^1}{|x+v(x'-x)-y|^{d+1}}\,\mathrm{d}v \\
&\leqslant u|x-y| \int_0^1 \frac{c_K^1}{(\frac{1}{2}|x-y|)^{d+1}}\,\mathrm{d}v = u\frac{2^{d+1}c_K^1}{|x-y|^d}.
\end{aligned}
$$

This is the claimed standard estimate, and the Dini estimate follows from part (3) with $\delta = 1$ and $c_\delta = 2^{d+1}c_K^i$.

The version with a cube or a quadrant follows with the same argument by simply restricting all the variables and the integrals to the relevant domain of definition. □

In particular, Dini kernels satisfy both Hörmander and dual Hörmander conditions, and hence all the results of the previous section apply to them:

Corollary 11.3.5 (Calderón–Zygmund). *Let X and Y be Banach spaces and $p_0 \in [1,\infty]$. Let $T \in \mathscr{L}(L^{p_0}(\mathbb{R}^d;X), L^{p_0,\infty}(\mathbb{R}^d;Y))$ be an operator with a Calderón–Zygmund kernel K. Then all conclusions of Theorem 11.2.5 hold with $\|K\|_{\text{Hör}}$ replaced by $\|\omega_K^2\|_{\text{Dini}}$ and $\|K\|_{\text{Hör}^*}$ by $\|\omega_K^1\|_{\text{Dini}}$ in the estimates.*

Proof. This follows at once from Theorem 11.2.5, where the same conclusions are deduced for Hörmander and/or dual Hörmander kernels K, and Lemma 11.3.4, where these assumptions are verified for under the Dini conditions. □

11.3.a An abstract domination theorem

We will first present an abstract form of the domination theorem, i.e., we postulate the relevant properties of the operator needed to carry out the proof, and only then return to the question of checking these properties in the concrete case of Calderón–Zygmund operators.

We will formulate the theorem for *positive sub-linear* operators mapping a linear space of X-valued functions into $L^0(\mathbb{R}^d;\mathbb{R}_+)$. By this we mean that for all functions f and g we have that $Tf \geqslant 0$ is a non-negative function, $T(\alpha f) = |\alpha| Tf$ for constants α, and $T(f+g) \leqslant Tf + Tg$ for all f, g in the domain of T. Note that if T is a *linear* operator mapping into $L^0(\mathbb{R}^d;Y)$, then the operator $f \mapsto \|Tf(\cdot)\|_Y$ is a positive sub-linear one, and this is the way that such operators will be naturally covered by the theory.

Theorem 11.3.6 (Abstract sparse domination). *Let X be a Banach space, let T be a positive sub-linear operator from $L^1(\mathbb{R}^d;X)$ into $L^0(\mathbb{R}^d;\mathbb{R}_+)$, and consider the associated maximal operator*

$$
M_T^{\#} f(x) := \sup_{Q \ni x}\sup_{y,z\in Q} |T(\mathbf{1}_{\complement(5Q)}f)(y) - T(\mathbf{1}_{\complement(5Q)}f)(z)|. \tag{11.28}
$$

Suppose that both T and $M_T^\#$ are bounded from $L^1(\mathbb{R}^d;X)$ to $L^{1,\infty}(\mathbb{R}^d)$. Then for every boundedly supported $f \in L^1(\mathbb{R}^d)$ and $\varepsilon \in (0,1)$, there is a $(1-\varepsilon)$-sparse family $\mathscr{S}$ of dyadic cubes such that, almost everywhere,

$$Tf \leqslant \frac{8 \cdot 10^d \cdot c_T}{\varepsilon} \sum_{S \in \mathscr{S}} \mathbf{1}_S \fint_{5S} \|f\|,$$

where

$$c_T := \|T\|_{1 \to 1,\infty} + \|M_T^\#\|_{1 \to 1,\infty}. \tag{11.29}$$

The heart of Theorem 11.3.6 is contained in the following lemma:

Lemma 11.3.7. *Under the assumptions of Theorem 11.3.6, for any $f \in L^1_{\text{loc}}(\mathbb{R}^d;X)$, any cube Q_0 and $\varepsilon \in (0,1)$, there are disjoint subcubes $Q_j \in \mathscr{D}(Q_0)$ such that*

$$\sum_j |Q_j| \leqslant \varepsilon |Q_0| \tag{11.30}$$

and, almost everywhere,

$$\mathbf{1}_{Q_0} T(\mathbf{1}_{5Q_0} f) \leqslant \frac{4 \cdot 10^d c_T}{\varepsilon} \Big(\mathbf{1}_{Q_0} \fint_{5Q_0} \|f\| + \sum_j \mathbf{1}_{Q_j} \fint_{5Q_j} \|f\| \Big) + \sum_j \mathbf{1}_{Q_j} T(\mathbf{1}_{5Q_j} f),$$

where c_T was defined in (11.29).

Proof. Given a cube Q_0, consider any disjoint family of its subcubes $Q_j \in \mathscr{D}(Q_0)$. Then we have

$$\begin{aligned} \mathbf{1}_{Q_0} T(\mathbf{1}_{5Q_0} f) &= \mathbf{1}_{Q_0 \setminus \bigcup_j Q_j} T(\mathbf{1}_{5Q_0} f) + \sum_j \mathbf{1}_{Q_j} T(\mathbf{1}_{5Q_0} f) \\ &\leqslant \mathbf{1}_{Q_0 \setminus \bigcup_j Q_j} T(\mathbf{1}_{5Q_0} f) + \sum_j \mathbf{1}_{Q_j} T(\mathbf{1}_{5Q_0 \setminus 5Q_j} f) + \sum_j \mathbf{1}_{Q_j} T(\mathbf{1}_{5Q_j} f), \end{aligned} \tag{11.31}$$

and

$$\begin{aligned} \mathbf{1}_{Q_j} T(\mathbf{1}_{5Q_0 \setminus 5Q_j} f) &\leqslant \mathbf{1}_{Q_j} [\inf_{Q_j} M_T^\#(\mathbf{1}_{5Q_0} f) + \inf_{Q_j} T(\mathbf{1}_{5Q_0 \setminus 5Q_j} f)] \\ &\leqslant \mathbf{1}_{Q_j} [\inf_{Q_j} M_T^\#(\mathbf{1}_{5Q_0} f) + \inf_{Q_j} \{ T(\mathbf{1}_{5Q_0} f) + T(\mathbf{1}_{5Q_j} f) \}] \end{aligned} \tag{11.32}$$

where we used sublinearity and the definition of $M_T^\#$ to get the estimates. Note that no convergence issues arise when viewing the above lines in the pointwise sense.

The last term in (11.31) already has the correct form, and it remains to choose the cubes Q_j in such a way that we have (11.30) as well as

$$\mathbf{1}_{Q_0 \setminus \bigcup_j Q_j} T(\mathbf{1}_{5Q_0} f) + \sum_j \mathbf{1}_{Q_j} T(\mathbf{1}_{5Q_0 \setminus 5Q_j} f) \leqslant \mathbf{1}_{Q_0} \frac{c_d c_T}{\varepsilon} \fint_{5Q_0} \|f\|.$$

For a $\lambda > 0$ to be chosen and every $Q \in \mathscr{D}(Q_0)$, we define $F(Q) \subseteq Q$ by

$$F(Q) := Q \cap [\{T(\mathbf{1}_{5Q} f) > \lambda \langle \|f\| \rangle_{5Q}\} \cup \{M_T^{\#}(\mathbf{1}_{5Q} f) > \lambda \langle \|f\| \rangle_{5Q}\}].$$

Thus, by the assumed $L^1(\mathbb{R}^d; X)$ to $L^{1,\infty}(\mathbb{R})$ bounds,

$$\begin{aligned} |F(Q)| &\leqslant |\{T(\mathbf{1}_{5Q} f) > \lambda \langle \|f\| \rangle_{5Q}\}| + |\{M_T^{\#}(\mathbf{1}_{5Q} f) > \lambda \langle \|f\| \rangle_{5Q}\}| \\ &\leqslant (\|T\|_{1\to 1,\infty} + \|M_T\|_{1\to 1,\infty}) \frac{\|\mathbf{1}_{5Q} f\|_1}{\lambda \langle \|f\| \rangle_{5Q}} = \frac{5^d}{\lambda} c_T \cdot |Q|. \end{aligned} \tag{11.33}$$

Let then $Q_j \in \mathscr{D}(Q_0)$ be the maximal dyadic subcubes such that

$$\frac{|Q_j \cap F(Q_0)|}{|Q_j|} > 2^{-d-1}.$$

The cubes Q_j are disjoint, so that

$$\sum_j |Q_j| \leqslant \sum_j \frac{|Q_j \cap F(Q_0)|}{2^{-d-1}} \leqslant 2^{d+1} |F(Q_0)| \leqslant \frac{2 \cdot 10^d}{\lambda} c_T \cdot |Q_0| = \varepsilon |Q_0|,$$

which is (11.30), if we choose

$$\lambda := \frac{2 \cdot 10^d}{\varepsilon} c_T.$$

Substituting back to (11.33), this choice gives in particular that

$$|F(Q)| \leqslant 2^{-d-1} |Q|.$$

Since $\mathbf{1}_{F(Q_0)} \leqslant M_{\mathscr{D}}(\mathbf{1}_{F(Q_0)})$ almost everywhere, we see that $F(Q_0)$ is contained in $\bigcup_j Q_j = \{M_{\mathscr{D}}(\mathbf{1}_{F(Q_0)}) > 2^{-d-1}\}$, except perhaps for a subset of measure zero. In particular, we have (a.e.)

$$\mathbf{1}_{Q_0 \setminus \bigcup_j Q_j} T(\mathbf{1}_{5Q_0} f) \leqslant \mathbf{1}_{Q_0 \setminus \bigcup_j Q_j} \lambda \langle \|f\| \rangle_{5Q_0}. \tag{11.34}$$

On the other hand, the maximality of Q_j implies that its dyadic parent $\widehat{Q}_j$ satisfies the opposite inequality, and hence

$$\frac{|Q_j \cap F(Q_0)|}{|Q_j|} \leqslant \frac{|\widehat{Q}_j \cap F(Q_0)|}{2^{-d} |\widehat{Q}_j|} \leqslant \frac{2^{-d-1}}{2^{-d}} = \frac{1}{2}.$$

But also $|F(Q_j)| \leqslant 2^{-d-1} |Q_j| \leqslant \frac{1}{4} |Q_j|$, and hence

$$|Q_j \setminus [F(Q_0) \cup F(Q_j)]| \geqslant (1 - \frac{1}{2} - \frac{1}{4}) |Q_j| > 0.$$

With any z_j in the non-empty set $Q_j \setminus [F(Q_0) \cup F(Q_j)]$, we can now complete the estimation of (11.32) as follows:

$$\begin{aligned}\mathbf{1}_{Q_j} T(\mathbf{1}_{5Q_0 \setminus 5Q_j} f) &\leqslant \mathbf{1}_{Q_j}[M_T^{\#}(\mathbf{1}_{5Q_0} f)(z_j) + T(\mathbf{1}_{5Q_0} f)(z_j) + T(\mathbf{1}_{5Q_j} f)(z_j)] \\ &\leqslant \mathbf{1}_{Q_j}[\lambda\langle\|f\|\rangle_{5Q_0} + \lambda\langle\|f\|\rangle_{5Q_0} + \lambda\langle\|f\|\rangle_{5Q_j}],\end{aligned}$$

where we used the bounds for $M_T^{\#}(\mathbf{1}_{5Q_0} f)$ and $T(\mathbf{1}_{5Q_0} f)$ on $\complement F(Q_0)$ that follow directly from the definition of these sets, and the analogous bound for $T(\mathbf{1}_{5Q_j} f)$ on $\complement F(Q_j)$. Hence

$$\sum_j \mathbf{1}_{Q_j} T(\mathbf{1}_{5Q_0 \setminus 5Q_j} f) \leqslant \mathbf{1}_{\bigcup_j Q_j} 2\lambda\langle\|f\|\rangle_{5Q_0} + \sum_j \mathbf{1}_{Q_j} \lambda\langle\|f\|\rangle_{5Q_j},$$

and together with (11.31), (11.34) and the choice of λ, this completes the proof of the lemma. □

Iterating the previous lemma, we obtain:

Lemma 11.3.8. *Under the assumptions of Theorem 11.3.6, for any cube Q_0 and $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d; X)$ and $\varepsilon \in (0,1)$, there is a $(1-\varepsilon)$-sparse subcollection $\mathscr{S}(Q_0) \subseteq \mathscr{D}(Q_0)$ such that, almost everywhere,*

$$\mathbf{1}_{Q_0} T(\mathbf{1}_{5Q_0} f) \leqslant \frac{8 \cdot 10^d c_T}{\varepsilon} \sum_{S \in \mathscr{S}(Q_0)} \mathbf{1}_S \fint_{5S} \|f\|.$$

Proof. By Lemma 11.3.7, almost everywhere we have

$$\mathbf{1}_{Q_0} T(\mathbf{1}_{5Q_0} f) \leqslant \frac{c_d c_T}{\varepsilon}\Big(\mathbf{1}_{Q_0} \fint_{5Q_0} \|f\| + \sum_j \mathbf{1}_{Q_j} \fint_{5Q_j} \|f\|\Big) + \sum_j \mathbf{1}_{Q_j^1} T(\mathbf{1}_{5Q_j^1} f)$$

for disjoint subcubes $Q_j^1 \in \mathscr{D}(Q_0)$ such that

$$\sum_j |Q_j^1| \leqslant \varepsilon |Q_0|,$$

and $c_d = 4 \cdot 10^d$ Applying the same estimate to each Q_j^1 in place of Q_0, and continuing by induction, almost everywhere we obtain

$$\begin{aligned}\mathbf{1}_{Q_0} T(\mathbf{1}_{5Q_0} f) \leqslant \frac{c_d c_T}{\varepsilon}\Big(\mathbf{1}_{Q_0} \fint_{5Q_0} \|f\| + 2\sum_{n=1}^{N-1} \sum_j \mathbf{1}_{Q_j^n} \fint_{5Q_j^n} \|f\| \\ + \sum_k \mathbf{1}_{Q_k^N} \fint_{5Q_k^n} \|f\|\Big) + \sum_k \mathbf{1}_{Q_k^N} T(\mathbf{1}_{5Q_k^N} f),\end{aligned} \tag{11.35}$$

where the Q_j^n are dyadic subcubes of some Q_i^{n-1} in such that

$$\sum_{j: Q_j^n \subseteq Q_i^{n-1}} |Q_j^n| \leqslant \varepsilon |Q_i^{n-1}|.$$

In particular,

$$\sum_j |Q_j^n| \leqslant \varepsilon \sum_i |Q_i^{n-1}| \leqslant \ldots \leqslant \varepsilon^n |Q_0|,$$

so that the support of the last term in (11.35) becomes negligible in the limit $N \to \infty$. Thus, almost everywhere, we have

$$\mathbf{1}_{Q_0} T(\mathbf{1}_{5Q_0} f) \leqslant 2\frac{c_d c_T}{\varepsilon} \sum_{n=0}^{\infty} \sum_j \mathbf{1}_{Q_j^n} \fint_{5Q_j^n} \|f\|, \tag{11.36}$$

where the pairwise disjoint subsets

$$E_j^n := Q_j^n \setminus \bigcup_k Q_k^{n+1}$$

have measure $|E_j^n| \geqslant (1-\varepsilon)|Q_j^n|$. In other words, the cubes Q_j^n form a $(1-\varepsilon)$-sparse subcollection $\mathscr{S}(Q_0) \subseteq \mathscr{D}(Q_0)$, and (11.36) is precisely the estimate asserted in the lemma. □

In order to pass from the local Lemma 11.3.8 to the global Theorem 11.3.6, we use:

Lemma 11.3.9. *Let $E \subseteq \mathbb{R}^d$ satisfy $0 < diam(E) < \infty$. Then there is a partition $\mathscr{Q}$ of $\mathbb{R}^d$ by dyadic cubes Q such that $E \subseteq 5Q$ for every $Q \in \mathscr{Q}$.*

Proof. Consider all dyadic cubes $Q \in \mathscr{D}$ with the property that $E \nsubseteq 2Q$. Clearly all cubes with $\operatorname{diam}(Q) < \frac{1}{2}\operatorname{diam}(E)$ will satisfy this condition. On the other hand, every cube $Q \in \mathscr{D}$ is contained in some $\widetilde{Q} \in \mathscr{D}$ such that $E \subseteq 2\widetilde{Q}$: if we fix some $x \in Q$ and then $r > 0$ large enough so that $E \subseteq B(x, r)$, then it suffices to take $\widetilde{Q} \supseteq Q$ with $\ell(\widetilde{Q}) > 2r$, since then $2\widetilde{Q} \supseteq B(x, \frac{1}{2}\ell(\widetilde{Q})) \supseteq E$.

Let $\mathscr{Q}$ be the collection of *maximal* dyadic cubes with the property that $E \nsubseteq 2Q$. Maximality implies disjointness, and from what we just checked, it follows that every $x \in \mathbb{R}^d$ is contained in some $Q \in \mathscr{Q}$, so these cubes form a partition of $\mathbb{R}^d$.

Since Q is maximal, its dyadic parent $\widehat{Q}$ satisfies $E \subseteq 2\widehat{Q}$. It remains to observe that $2\widehat{Q} \subseteq 5Q$ to complete the proof. □

We now return to:

Proof of Theorem 11.3.6. If $f \equiv 0$, there is nothing to prove, so fix a non-zero, compactly supported $f \in L^1_c(\mathbb{R}^d; X)$. Thus $E = \operatorname{supp} f$ satisfies $0 < \operatorname{diam}(E) < \infty$ as required to apply Lemma 11.3.9. This lemma produces a partition $\mathscr{Q} \subseteq \mathscr{D}$ of $\mathbb{R}^d$ such that $\operatorname{supp} f \subseteq 5Q$, and thus $\mathbf{1}_{5Q} f = f$, for every $Q \in \mathscr{Q}$. This means that

$$Tf = \sum_{Q \in \mathscr{Q}} \mathbf{1}_Q Tf = \sum_{Q \in \mathscr{Q}} \mathbf{1}_Q T(\mathbf{1}_{5Q} f).$$

Now Lemma 11.3.8 applies to each term on the right, producing $(1-\varepsilon)$-sparse subcollections $\mathscr{S}(Q) \subseteq \mathscr{D}(Q)$ for each $Q \in \mathscr{Q}$, and

$$\sum_{Q\in\mathscr{Q}} \mathbf{1}_Q T(\mathbf{1}_{5Q} f) \leqslant \sum_{Q\in\mathscr{Q}} \frac{c_d c_T}{\varepsilon} \sum_{S\in\mathscr{S}(Q)} \mathbf{1}_S \fint_{5S} \|f\| = \frac{c_d c_T}{\varepsilon} \sum_{S\in\mathscr{S}} \mathbf{1}_S \fint_{5S} \|f\|,$$

where $\mathscr{S} := \bigcup_{Q\in\mathscr{Q}} \mathscr{S}(Q)$ and $c_d = 8\cdot 10^d$. It is immediate that this union of disjointly supported sparse collections remains sparse, as the same pairwise disjoint subsets $E(S) \subseteq S$ remain pairwise disjoint also among all $S \in \mathscr{S}$. □

11.3.b Sparse operators and domination

With Theorem 11.3.6 at our disposal, the following notion should not appear too alien to the reader:

Definition 11.3.10 (Sparse operator). *Given a sparse collection of sets* $\mathscr{S} \subseteq \mathscr{D}$*, the associated* sparse operator *is*

$$A_{\mathscr{S}} f := \sum_{S\in\mathscr{S}} \mathbf{1}_S \fint_S f.$$

More generally, with a dilation factor $\varrho \geqslant 1$*, we define*

$$A^{\varrho}_{\mathscr{S}} f := \sum_{S\in\mathscr{S}} \mathbf{1}_S \fint_{\varrho S} f.$$

In contrast to most other operators that we encounter, the boundedness properties of the sparse operators tend to be extremely easy. As a first illustration, we check the L^p boundedness of $A_{\mathscr{S}}$ by dualising against $g \in L^{p'}$:

$$\begin{aligned}\int A_{\mathscr{S}} f \cdot g &= \sum_{S\in\mathscr{S}} \fint_S f \cdot \fint_S g \cdot |S| \leqslant \sum_{S\in\mathscr{S}} \inf_S M_{\mathscr{D}} f \cdot \inf_S M_{\mathscr{D}} g \cdot \frac{|E(S)|}{\gamma} \\ &\leqslant \frac{1}{\gamma}\int M_{\mathscr{D}} f \cdot M_{\mathscr{D}} g \leqslant \frac{1}{\gamma}\|M_{\mathscr{D}} f\|_p \cdot \|M_{\mathscr{D}} g\|_{p'} \leqslant \frac{1}{\gamma} p'\|f\|_p \cdot p\|g\|_{p'}.\end{aligned}$$

This shows that $\|A_{\mathscr{S}}\|_{p\to p} \leqslant \gamma^{-1} pp'$, where γ is the sparseness parameter; since $A_{\mathscr{S}}$ is manifestly positive, it suffices to consider positive functions above, and the same bound persists for vector-valued functions.

Looking back at the statement of Theorem 11.3.6, it *almost* says that $Tf \leqslant c \cdot A_{\mathscr{S}}\|f\|$ under the assumptions of the theorem, but the presence of the expanded cubes $5S$ prevents this from being strictly true in the stated form. While the variant of a sparse operator implicitly appearing in Theorem 11.3.6 would be almost as good as $A_{\mathscr{S}}$ for many purposes, the use of the more symmetric (indeed, self-dual) operators $A_{\mathscr{S}}$ as in Definition 11.3.10 is often preferred.

A trivial way to achieve this in Theorem 11.3.6 is to dominate $\mathbf{1}_S \leqslant \mathbf{1}_{5S}$, after which the same cube $5S$ will appear in both the indicator and the integral. These cubes will still be sparse, if only with a smaller parameter $\gamma = 5^{-d}(1-\varepsilon)$, since the disjoint major subsets $E(S) \subseteq S \subseteq 5S$ satisfy $|E(S)| \geqslant (1-\varepsilon)|S| = (1-\varepsilon)5^{-d}|5S|$ and hence also qualify for the disjoint major subsets of the expanded cubes $5S$. An apparent loss in this construction is the fact that these $5S$ are no longer *dyadic* cubes. Even this problem, however, can be fixed, by a variant of the shifted dyadic cubes that we introduced in Definition 3.2.25. Recall that the standard dyadic system is

$$\mathscr{D}^0 := \bigcup_{j\in\mathbb{Z}} \mathscr{D}_j^0, \qquad \mathscr{D}_j^0 := \{2^{-j}([0,1)^d + m) : \; m \in \mathbb{Z}^d\}.$$

We will need the case $N = 5$ of the following statement, but we record the general formulation for convenience of reference, as the case $N = 3$ also features in various applications.

Proposition 11.3.11 (Dilated dyadic cubes). *Let $N \in \mathbb{Z}_+$ be odd. Then the collection of N-fold concentric dilations $\{NQ : Q \in \mathscr{D}(\mathbb{R}^d)\}$ can be partitioned into N^d subcollections $\mathscr{D}^{n;N}$, $n \in \mathbb{Z}_N^d$, each of which has the same covering and nestedness properties as $\mathscr{D}$, namely,*

$$\mathscr{D}^{n;N} = \bigcup_{j\in\mathbb{Z}} \mathscr{D}_j^{n;N},$$

where for each $j \in \mathbb{Z}$:

(1) *$\mathscr{D}_j^{n;N}$ is a partition of $\mathbb{R}^d$ consisting cubes of side-length $N \cdot 2^{-j}$, and*
(2) *$\mathscr{D}_{j+1}^{n;N}$ is a refinement of $\mathscr{D}_j^{n;N}$.*

Proof. Since $\mathscr{D}_j(\mathbb{R}^d) = \{I_1 \times \cdots \times I_d : I_i \in \mathscr{D}_j(\mathbb{R})\}$ and $N(I_1 \times \cdots \times I_d) = NI_1 \times \cdots \times NI_d$, it suffices to verify the case $d = 1$. In the calculation that follows, we will need to dilate an interval $I = [c-r, c+r)$ both by the algebraic multiplication $a \cdot I = \{a \cdot t : t \in I\} = [ac - ar, ac + ar)$ and by the concentric dilation, for which we temporarily adopt the heavier notation $a \odot I = [c - ar, c + ar)$ for the sake of distinction.

With these notations fixed, we have

$$\begin{aligned}
\{N \odot I : I \in \mathscr{D}_j\} &= \{N \odot 2^{-j}([0,1) + m) : m \in \mathbb{Z}\} \\
&= \{2^{-j}([-N', N'+1) + m) : m \in \mathbb{Z}\} \quad \big(N := 2N'+1\big) \\
&= \{2^{-j}([0,N) + m - N') : m \in \mathbb{Z}\} \\
&= \{2^{-j}([0,N) + m) : m \in \mathbb{Z}\} \\
&= \Big\{N2^{-j}\Big([0,1) + \frac{m}{N}\Big) : m \in \mathbb{Z}\Big\}.
\end{aligned}$$

The sought-after partition of this collection is now achieved as follows: For each $n \in \mathbb{Z}_N = \{0, 1, \ldots, N-1\}$ and $j \in \mathbb{Z}$, we define

$$\mathscr{D}_j^{n;N} := \Big\{N2^{-j}\Big([0,1)+k+\frac{\alpha(n,j)}{N}\Big) : k\in\mathbb{Z}\Big\} \tag{11.37}$$

for appropriate $\alpha(n,j)\in\mathbb{Z}_N$ to be shortly determined. It is clear that each $\mathscr{D}_j^{n;N}$ satisfies (1) from the statement of the Proposition, no matter how we choose $\alpha(n,j)$. To ensure (2), it suffices to check that the left (or equivalently right) half of any $I\in\mathscr{D}_j^{n;N}$ belongs to $\mathscr{D}_{j+1}^{n;N}$. For a generic I as written above, the left half will be

$$N2^{-j}\Big([0,\tfrac12)+k+\frac{\alpha(n,j)}{N}\Big) = N2^{-j-1}\Big([0,1)+2k+\frac{2\alpha(n,j)}{N}\Big).$$

For this to be in $\mathscr{D}_{j+1}^{n;N}$, it is necessary and sufficient that

$$2\alpha(n,j)\equiv\alpha(n,j+1)\mod N \tag{11.38}$$

If we specify $\alpha(n,0):=n$, all other $\alpha(n,j)$, $j\in\mathbb{Z}\setminus\{0\}$ will be uniquely determined by (11.38), since 2 has a multiplicative inverse in $\mathbb{Z}_N$ for odd N. Indeed, the solution is given by

$$\alpha(n,j)\equiv 2^j n\mod N, \tag{11.39}$$

where the negative powers are interpreted in the sense of the multiplicative inverse mod N.

For each $j\in\mathbb{Z}$, the map $n\mapsto 2^j n\mod N$ is a bijection on $\mathbb{Z}_N$, and thus

$$\begin{aligned}\bigcup_{n=0}^{N}\mathscr{D}_j^{n;N} &= \Big\{N2^{-j}\Big([0,1)+k+\frac{a}{N}\Big) : k\in\mathbb{Z}, a\in\mathbb{Z}_N\Big\}\\ &= \Big\{N2^{-j}\Big([0,1)+\frac{m}{N}\Big) : m\in\mathbb{Z}\Big\} = \{N\odot I : I\in\mathscr{D}_j\},\end{aligned}$$

so indeed $\{N\odot I : I\in\mathscr{D}\}$ is a disjoint union of the collections $\mathscr{D}^{n;N}$, $n\in\mathbb{Z}_N$, and we already checked that each $\mathscr{D}^{n;N}$ has the properties (1) and (2). □

Remark 11.3.12 (Shifted dyadic cubes). The cube families $\mathscr{D}^{n;N}$ constructed above are close relatives of the *shifted dyadic cubes* of Definition 3.2.25, and they satisfy a variant of the Covering Lemma 3.2.26:

> Given an odd $N\in\mathbb{Z}_+$, for every cube $Q\subseteq\mathbb{R}^d$, there exist a vector $n\in\mathbb{Z}_N^d$ and a cube $D\in\mathscr{D}^{n;N}$ such that
>
> $$\frac{N}{N-1}\ell(Q)<\ell(D)\leqslant\frac{2N}{N-1}\ell(Q)\quad\text{and}\quad Q\subseteq D. \tag{11.40}$$

In fact, let $R\in\mathscr{D}$ be a cube of side-length $\ell(R)\in(\ell(Q)/(2N'),\ell(Q)/N']$ that contains the centre z_Q of Q, where $N=2N'+1$ as before. Then $D=NR\in\mathscr{D}^{n;N}$ for some $n\in\mathbb{Z}_N^d$, and D contains the cube of side-length $2N'\ell(R)>\ell(Q)$ centred at z_Q; thus $D\supseteq Q$, and $\ell(D)=N\ell(R)$ lies exactly in the range asserted in (11.40).

Also note that both the partition and refinement properties (1) and (2) of Proposition 11.3.11 of each $\mathscr{D}^{n;N}$, as well as the covering property of every cube $Q \subseteq \mathbb{R}^d$ by a cube in some $\mathscr{D}^{n;N}$, remain invariant if we drop the algebraic dilation factor N in (11.37), so as to be back to cubes of side-length 2^{-j}. When $N = 3$, this reproduces precisely the shifted dyadic cubes of Definition 3.2.25; since $2 \equiv -1 \mod 3$, (11.39) reduces in this case to the simpler form $\alpha(n,j) = (-1)^j n$, where reference to modular arithmetic can be avoided.

It is now easy to show that the sparse operators with a dilation, $A^{\varrho}_{\mathscr{S}}$, may always be dominated by a finite number of the simple sparse operators $A_{\mathscr{S}^n}$. It is technically convenient to take an odd integer N for the dilation factor. This causes little loss of generality since, choosing $N \geqslant \varrho$, we can always dominate

$$\fint_{\varrho Q} f \leqslant \Big(\frac{N}{\varrho}\Big)^d \fint_{NQ} f$$

and hence $A^{\varrho}_{\mathscr{S}} f \leqslant (N/\varrho)^d A^N_{\mathscr{S}} f$ for $f \geqslant 0$.

Lemma 11.3.13. *Let $\mathscr{S} \subseteq \mathscr{D}$ be ε-sparse for some $\varepsilon \in (0,1)$, and $N \in \mathbb{Z}_+$ be odd. Then there are $N^{-d}\varepsilon$-sparse collections $\mathscr{S}^n \subseteq \mathscr{D}^{n;N}$ for each $n \in \mathbb{Z}^d_N$ such that, for every non-negative $f \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$,*

$$A^N_{\mathscr{S}} f \leqslant \sum_{n \in \mathbb{Z}^d_N} A_{\mathscr{S}^n} f$$

Proof. We note that the collection $\{5Q : Q \in \mathscr{S}\}$ is $N^{-d}\varepsilon$-sparse, with the same disjoint subsets $E(Q) \subseteq Q \subseteq NQ$ that satisfy $|E(Q)| \geqslant \varepsilon|Q| = \varepsilon N^{-d}|NQ|$. By Proposition 11.3.11, we have a decomposition $\{NQ : Q \in \mathscr{D}\} = \bigcup_{n \in \mathbb{Z}^d_N} \mathscr{D}^{n;N}$ into dyadic systems $\mathscr{D}^{n;N}$. We then define $\mathscr{S}^n := \{NQ : Q \in \mathscr{S}\} \cap \mathscr{D}^{n;5}$. Thus

$$\begin{aligned} A^N_{\mathscr{S}} f &= \sum_{Q \in \mathscr{S}} \mathbf{1}_Q \fint_{NQ} f \leqslant \sum_{Q \in \mathscr{S}} \mathbf{1}_{NQ} \fint_{NQ} f = \sum_{n \in \mathbb{Z}^d_N} \sum_{\substack{Q \in \mathscr{S} \\ NQ \in \mathscr{D}^{n;N}}} \mathbf{1}_{NQ} \fint_{NQ} f \\ &= \sum_{n \in \mathbb{Z}^d_N} \sum_{Q' \in \mathscr{S}^n} \mathbf{1}_{Q'} \fint_{Q'} f = \sum_{n \in \mathbb{Z}^d_N} A_{\mathscr{S}^n} f. \end{aligned}$$

□

We can now reformulate Theorem 11.3.6 in terms of sparse operators:

Theorem 11.3.14 (Abstract sparse domination II). *Let X be a Banach space, and let T be a positive sub-linear operator from $L^1(\mathbb{R}^d; X)$ into $L^0(\mathbb{R}^d; \mathbb{R}_+)$, and consider the associated Lerner's maximal operator*

$$M^{\#}_T f(x) := \sup_{Q \ni x} \sup_{y,z \in Q} |T(\mathbf{1}_{\complement(5Q)} f)(y) - T(\mathbf{1}_{\complement(5Q)} f)(z)|.$$

Suppose that both T and M_T are bounded from $L^1(\mathbb{R}^d;X)$ to $L^{1,\infty}(\mathbb{R}^d)$. Then for every boundedly supported $f\in L^1(\mathbb{R}^d)$, there is a 5^{-1}-sparse collection $\mathscr{S}\subseteq\mathscr{D}$ and, for every $n\in\mathbb{Z}_5^d$, a 5^{-d-1}-sparse collection $\mathscr{S}^n\subseteq\mathscr{D}^{n;5}$ of the dyadic systems as in Proposition 11.3.11, such that almost everywhere

$$Tf\leqslant 10^{d+1}c_T A^5_{\mathscr{S}}\|f\|\leqslant 10^{d+1}c_T\sum_{n\in\mathbb{Z}_5^d}A_{\mathscr{S}^n}\|f\|,$$

where $c_T:=\|T\|_{1\to 1,\infty}+\|M_T^{\#}\|_{1\to 1,\infty}$.

Proof. Choosing $\varepsilon=4/5$ in Theorem 11.3.6, we find a $\frac{1}{5}$-sparse collection $\mathscr{S}\subseteq\mathscr{D}$ such that

$$Tf\leqslant\frac{8\cdot 10^d\cdot c_T}{4/5}\sum_{S\in\mathscr{S}}\mathbf{1}_S\fint_{5S}\|f\|=10^{d+1}c_T A^5_{\mathscr{S}}\|f\|.$$

This is the first claim, and the second one follows from Lemma 11.3.13. □

11.3.c Sparse domination of Calderón–Zygmund operators

The goal of this section is to specialise the abstract Theorem 11.3.14 to the case of Calderón–Zygmund operators in the following form:

Theorem 11.3.15 (Sparse domination of singular integrals). *Let X and Y be Banach spaces, $p_0\in[1,\infty]$, and let*

$$T\in\mathscr{L}(L^{p_0}(\mathbb{R}^d;X),L^{p_0,\infty}(\mathbb{R}^d;Y))$$

be an operator with a Dini kernel K. Then for every boundedly supported $f\in L^1(\mathbb{R}^d)$, there is a 5^{-1}-sparse collection $\mathscr{S}\subseteq\mathscr{D}$ and, for every $n\in\mathbb{Z}_5^d$, a 5^{-d-1}-sparse collection $\mathscr{S}^n\subseteq\mathscr{D}^{n;5}$ of the dyadic systems as in Proposition 11.3.11, such that almost everywhere

$$\|Tf\|_Y\leqslant c_{d,T}A^5_{\mathscr{S}}\|f\|_X\leqslant c_{d,T}\sum_{n\in\mathbb{Z}_5^d}A_{\mathscr{S}^n}\|f\|,$$

where

$$c_{d,T}\leqslant c_d\big(\|T\|_{L^{p_0}(\mathbb{R}^d;X)\to L^{p_0,\infty}(\mathbb{R}^d;Y)}+c_K+\|\omega\|_{\mathrm{Dini}}\big)$$

with c_K and ω as in Definition 11.3.1.

The result remains true if $\mathbb{R}^d$ is systematically replaced by a cube or a quadrant of $\mathbb{R}^d$, both in the function spaces where the boundedness is considered, and in the definition of the kernel bounds c_K and $\|\omega_K\|_{\mathrm{Dini}}$.

Proof. By Theorem 11.3.14, applied to the positive sub-linear operator $U: f\mapsto\|Tf(\cdot)\|_Y$, the result follows if we can estimate $\|U\|_{L^1\to L^{1,\infty}}$ and $\|M_U\|_{L^1\to L^{1,\infty}}$ by the bound for $c_{d,T}$ given above. For the former, this is immediate from the Calderón–Zygmund Theorem 11.2.5 and Lemma 11.3.4, which show that

$$\begin{aligned}\|U\|_{L^1(\mathbb{R}^d;X)\to L^{1,\infty}(\mathbb{R}^d)} &= \|T\|_{L^1(\mathbb{R}^d;X)\to L^{1,\infty}(\mathbb{R}^d;Y)} \\ &\leqslant c_d\big(\|T\|_{L^{p_0}(\mathbb{R}^d;X)\to L^{p_0,\infty}(\mathbb{R}^d;Y)} + \|K\|_{\text{Hör}}\big) \\ &\leqslant c_d\big(\|T\|_{L^{p_0}(\mathbb{R}^d;X)\to L^{p_0,\infty}(\mathbb{R}^d;Y)} + \|K\|_{\text{Dini}}\big).\end{aligned}$$

For M_U, we first observe that, by the triangle inequality,

$$\begin{aligned}|U(\mathbf{1}_{\complement 5Q}f)(y) - U(\mathbf{1}_{\complement 5Q}f)(z)| &= |\|T(\mathbf{1}_{\complement 5Q}f)(y)\|_Y - \|T(\mathbf{1}_{\complement 5Q}f)(z)\|_Y| \\ &\leqslant \|T(\mathbf{1}_{\complement 5Q}f)(y) - T(\mathbf{1}_{\complement 5Q}f)(z)\|_Y.\end{aligned}$$

Hence, taking the supremum over $y, z \in Q$ and then over cubes $Q \ni x$, it follows that

$$M_U f(x) \leqslant M_T^{\#} f(x) := \sup_{Q\ni x}\sup_{y,z\in Q} \|T(\mathbf{1}_{\complement(5Q)}f)(y) - T(\mathbf{1}_{\complement(5Q)}f)(z)\|_Y.$$

The norm estimate of the latter is the content of the following lemma. □

Lemma 11.3.16. *Let X and Y be Banach spaces, $p_0 \in [1,\infty]$, and let T be an operator with a Dini kernel $K : \dot{\mathbb{R}}^{2d} \to \mathscr{L}(X,Y)$. Then the maximal operator*

$$M_T^{\#} f(x) := \sup_{Q\ni x}\sup_{y,z\in Q} \|T(\mathbf{1}_{\complement(5Q)}f)(y) - T(\mathbf{1}_{\complement(5Q)}f)(z)\|_Y$$

satisfies

$$M_T^{\#} f(x) \leqslant c_d(c_K + \|\omega_K\|_{\text{Dini}}) M f(x)$$

and

$$\|M_T^{\#}\|_{L^1(\mathbb{R}^d;X)\to L^{1,\infty}(\mathbb{R}^d)} \leqslant c_d(c_K + \|\omega_K\|_{\text{Dini}}).$$

The result remains true if $\mathbb{R}^d$ is systematically replaced by a cube $Q_0 \subseteq \mathbb{R}^d$ or a quadrant $S \subseteq \mathbb{R}^d$, both in the function spaces where the boundedness is considered, and in the definition of the kernel bounds c_K and $\|\omega_K\|_{\text{Dini}}$.

Proof. For $x, x_0, x_1 \in Q$, we have

$$\begin{aligned}&T(\mathbf{1}_{\complement(5Q)}f)(x_0) - T(\mathbf{1}_{\complement(5Q)}f)(x_1) \\ &\qquad = \sum_{j=0}^{1}(-1)^j [T(\mathbf{1}_{\complement(5Q)}f)(x_j) - T(\mathbf{1}_{\complement(5Q)}f)(x)] \\ &\qquad = \sum_{j=0}^{1}(-1)^j \int_{\complement(5Q)} [K(x_j,y) - K(x,y)] f(y)\,\mathrm{d}y,\end{aligned}$$

where

$$\begin{aligned}
&\Big\| \int_{\complement(5Q)} [K(x_j,y)-K(x,y)]f(y)\,\mathrm{d}y \Big\| \\
&\leqslant \int_{\complement B(x,4\sqrt{d}\ell(Q))} \|[K(x_j,y)-K(x,y)]f(y)\|\,\mathrm{d}y \\
&\qquad + \int_{B(x,4\sqrt{d}\ell(Q))\setminus(5Q)} \|[K(x_j,y)-K(x,y)]f(y)\|\,\mathrm{d}y =: I + II
\end{aligned}$$

where, observing that $|x_j - x| < \sqrt{d}\ell(Q) \leqslant \frac{1}{4}|x-y|$ for $x, x_j \in Q$ and $y \in \complement B(x, 4\sqrt{d}\ell(Q))$,

$$\begin{aligned}
I &\leqslant \int_{\complement B(x,4\sqrt{d}\ell(Q))} \omega_K^1\Big(\frac{|x_j - x|}{|x-y|}\Big)\frac{1}{|x-y|^d}\|f(y)\|\,\mathrm{d}y \\
&\leqslant \sum_{k=2}^{\infty} \int_{2^k\sqrt{d}\ell(Q)\leqslant|y-x|<2^{k+1}\sqrt{d}\ell(Q)} \omega_K^1\Big(\frac{\sqrt{d}\ell(Q)}{2^k\sqrt{d}\ell(Q)}\Big)\frac{\|f(y)\|\,\mathrm{d}y}{(2^k\sqrt{d}\ell(Q))^d} \\
&\leqslant \sum_{k=2}^{\infty} \omega_K^1(2^{-k})c_d \fint_{B(x,2^{k+1}\sqrt{d}\ell(Q))} \|f(y)\|\,\mathrm{d}y \\
&\leqslant c_d Mf(x)\sum_{k=2}^{\infty}\omega_K^1(2^{-k}) \leqslant c_d Mf(x)\|\omega_K^1\|_{\mathrm{Dini}},
\end{aligned}$$

by Lemma 11.3.3 in the last step. On the other hand, since $|x_j - y|, |x-y| \geqslant 2\ell(Q)$ for $x, x_j \in Q$ and $y \notin 5Q$, we obtain

$$\begin{aligned}
II &\leqslant \int_{B(x,2\sqrt{d}\ell(Q))\setminus(5Q)} c_K\frac{2}{(2\ell(Q))^d}\|f(y)\|\,\mathrm{d}y \\
&\leqslant c_K c_d \fint_{B(x,2\sqrt{d}\ell(Q))} \|f(y)\|\,\mathrm{d}y \leqslant c_K c_d Mf(x).
\end{aligned}$$

These bounds give the pointwise estimate for $M_T^{\#}f(x)$, and the norm estimate is then immediate from the corresponding bound for the Hardy–Littlewood maximal operator M.

The case of a cube or a quadrant in place of $\mathbb{R}^d$ follows by inspection of the same argument: if all variables under consideration are restricted like this, it is evident that only the corresponding restrictions of the kernel conditions will be needed to make the estimates. □

11.3.d Weighted norm inequalities and the A_2 theorem

We are now ready to provide the main application of the sparse domination of Calderón–Zygmund operators: their weighted norm inequalities with an optimal dependence of the weight. A function $w \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ is called a *weight* if $w(x) \in (0,\infty)$ almost everywhere. We recall from Appendix J the following definition, which we now extend to the local situation as well:

Definition 11.3.17. *For $p \in (1,\infty)$ the Muckenhoupt A_p characteristic of a weight w is defined by*

$$[w]_{A_p} := \sup_Q \Big(\fint_Q w(x)\,\mathrm{d}x \Big) \Big(\fint_Q w^{1-p'}(x)\,\mathrm{d}x \Big)^{p-1},$$

where the supremum is over all (axes-parallel) cubes $Q \subseteq \mathbb{R}^d$. We say that w is an A_p weight if $[w]_{A_p} < \infty$.

For a cube or quadrant $Q_0 \subseteq \mathbb{R}^d$, we define the local weight characteristic $[w]_{A_p(Q_0)}$ and the weight class $A_p(Q_0)$ in a similar way, but restricting the supremum to cubes $Q \subseteq Q_0$ only.

For the treatment of weighted norm inequalities, it is useful to introduce the following simple but far-reaching idea:

Remark 11.3.18 (Dual weight trick). Given an operator T, a weight w and an exponent $p \in (1,\infty)$, consider an inequality of the form

$$\|T(h)\|_{L^p(w)} \leqslant C\|h\|_{L^p(w)} \qquad \forall h \in L^p(w). \tag{11.41}$$

If σ is another weight, we observe that $h = f\sigma$ is in $L^p(w)$ if and only if $f \in L^p(\sigma^p w)$. With this substitution, the previous estimate becomes

$$\|T(f\sigma)\|_{L^p(w)} \leqslant C\|f\sigma\|_{L^p(w)} = C\|f\|_{L^p(\sigma^p w)} \qquad \forall f \in L^p(\sigma^p w).$$

Equating the weights inside the operator and on the right hand side, we want to arrange that $\sigma = \sigma^p w$, i.e., that $\sigma = w^{-1/(p-1)}$; this is called the (L^p-)*dual weight* of w. With this choice, the previous display reduces to

$$\|T(f\sigma)\|_{L^p(w)} \leqslant C\|f\|_{L^p(\sigma)} \quad \forall f \in L^p(\sigma), \quad \sigma := w^{-1/(p-1)}. \tag{11.42}$$

Applying duality in $L^p(w)$, yet another equivalent condition is given by the conveniently symmetric formulation

$$\int T(f\sigma)\cdot gw \leqslant C\|f\|_{L^p(\sigma)}\|g\|_{L^{p'}(w)} \quad \forall f \in L^p(\sigma), g \in L^{p'}(w). \tag{11.43}$$

Thus all three formulations (11.41), (11.42) and (11.43) are equivalent.

We now give the A_2 theorem for the sparse operators $A_{\mathscr{S}}$. The simplicity of this argument is a manifestation of the usefulness of dominating other operators by the sparse ones.

Theorem 11.3.19 (Cruz-Uribe–Martell–Pérez). *Let $\varepsilon \in (0,1)$ and $\mathscr{S} \subseteq \mathscr{D}$ be ε-sparse. Let $N \in \mathbb{Z}_+$ be odd. If $w \in A_2$, then the sparse operator $A^N_{\mathscr{S}}$ is bounded on $L^2(w)$, and*

$$\|A^N_{\mathscr{S}}\|_{\mathscr{L}(L^2(w))} \leqslant \frac{4}{\varepsilon} N^{2d} [w]_{A_2}.$$

Proof. By the dual weight trick (Remark 11.3.18), with $\sigma := w^{-1}$ we need to prove that

$$\int A^N_{\mathscr{S}}(f\sigma)\cdot gw \leqslant \frac{4}{\varepsilon}N^{2d}[w]_{A_2}\|f\|_{L^2(\sigma)}\|g\|_{L^2(w)} \qquad \forall f\in L^2(\sigma), g\in L^2(w).$$

Since $A_{\mathscr{S}}$ is a positive operator, both g and h may be taken to be positive, and there are no subtle convergence issues in the computation that follows. We first observe that

$$\langle f\sigma\rangle_Q = \frac{1}{|Q|}\int_Q f\sigma = \frac{\sigma(Q)}{|Q|}\frac{1}{\sigma(Q)}\int_Q f\sigma = \langle\sigma\rangle_Q\langle f\rangle^\sigma_Q,$$

where $\sigma(Q) = \int_Q \sigma$ and $\langle f\rangle^\sigma_Q$ is the average of f with respect to the measure induced by the weight σ. We denote the corresponding dyadic maximal operator by $M^\sigma_{\mathscr{D}} f := \sup_{Q\in\mathscr{D}} \mathbf{1}_Q\langle f\rangle^\sigma_Q$; this operator is bounded on $L^2(\sigma)$ with norm 2 according to Doob's maximal inequality (Theorem 3.2.2, cf. explanations preceding Theorem 3.2.27) with $p = p' = 2$.

We can then estimate, using that $[w]_{A_2} = \sup_Q\langle w\rangle_Q\langle\sigma\rangle_Q$ by definition,

$$\begin{aligned}\int A^N_{\mathscr{S}}(f\sigma)\cdot gw &= \sum_{Q\in\mathscr{S}}\langle f\sigma\rangle_{NQ}\int_{\mathbb{R}^d}\mathbf{1}_Q\cdot gw\\ &= \sum_{Q\in\mathscr{S}}\langle f\sigma\rangle_{NQ}\langle gw\rangle_Q|Q|\\ &= \sum_{Q\in\mathscr{S}}\langle\sigma\rangle_{NQ}\langle w\rangle_Q\langle f\rangle^\sigma_{NQ}\langle g\rangle^w_Q|Q|,\end{aligned}$$

where

$$\langle\sigma\rangle_{NQ}\langle w\rangle_Q \leqslant \langle\sigma\rangle_{NQ}\langle w\rangle_{NQ}N^d \leqslant [w]_{A_2}N^d.$$

Hence

$$\int A^N_{\mathscr{S}}(f\sigma)\cdot gw \leqslant N^d[w]_{A_2}\sum_{Q\in\mathscr{S}}\langle f\rangle^\sigma_{NQ}\langle g\rangle^w_Q\frac{|E(Q)|}{\varepsilon},$$

where

$$\langle g\rangle^w_Q \leqslant \inf_{z\in Q} M^w_{\mathscr{D}}g(z)$$

by definition of the dyadic maximal operator. As for $\langle f\rangle^\sigma_{NQ}$, we observe by Proposition 11.3.11 that the dilated cube NQ belongs to one of the N^d dyadic system $\mathscr{D}^{n;N}$, where $n\in\mathbb{Z}^d_N$, and the average over NQ is then something that appears in the corresponding maximal operator $M_{\mathscr{D}^{n;N}}$. Hence

$$\begin{aligned}&\sum_{Q\in\mathscr{S}}\langle f\rangle^\sigma_{NQ}\langle g\rangle^w_Q\frac{|E(Q)|}{\varepsilon} \qquad (11.44)\\ &\quad = \sum_{n\in\mathbb{Z}^d_N}\sum_{\substack{Q\in\mathscr{S}\\ NQ\in\mathscr{D}^{n;N}}}\langle f\rangle^\sigma_{NQ}\langle g\rangle^w_Q\frac{|E(Q)|}{\varepsilon}\end{aligned}$$

$$
\begin{aligned}
&\leqslant \sum_{n\in\mathbb{Z}_N^d} \sum_{\substack{Q\in\mathscr{S} \\ NQ\in\mathscr{D}^{n;N}}} \inf_Q M^\sigma_{\mathscr{D}^{n;N}} f \cdot \inf_Q M^w_{\mathscr{D}} g \cdot \frac{|E(Q)|}{\varepsilon} \\
&\leqslant \frac{1}{\varepsilon} \sum_{n\in\mathbb{Z}_N^d} \sum_{\substack{Q\in\mathscr{S} \\ NQ\in\mathscr{D}^{n;N}}} \int_{E(Q)} M^\sigma_{\mathscr{D}^{n;N}} f \cdot M^w_{\mathscr{D}} g \\
&\leqslant \frac{1}{\varepsilon} \sum_{n\in\mathbb{Z}_N^d} \int_{\mathbb{R}^d} M^\sigma_{\mathscr{D}^{n;N}} f \cdot M^w_{\mathscr{D}} g \cdot \sigma^{1/2} w^{1/2} \\
&\leqslant \frac{1}{\varepsilon} \sum_{n\in\mathbb{Z}_N^d} \|M^\sigma_{\mathscr{D}^{n;N}} f\|_{L^2(\sigma)} \|M^w_{\mathscr{D}} g\|_{L^2(w)} \\
&\leqslant \frac{1}{\varepsilon} \sum_{n\in\mathbb{Z}_N^d} 2\|f\|_{L^2(\sigma)} \cdot 2\|g\|_{L^2(w)} \\
&= \frac{4}{\varepsilon} N^d \|f\|_{L^2(\sigma)} \|g\|_{L^2(w)}.
\end{aligned}
$$

Substituting back, this gives the claimed bound for $\|A_{\mathscr{S}}\|_{\mathscr{L}(L^2(w))}$. □

Corollary 11.3.20. *Let $\varepsilon \in (0,1)$ and $\mathscr{S} \subseteq \mathscr{D}$ be ε-sparse, and let $Q_0 \in \mathscr{D}$. If $N \in \mathbb{Z}_+$ is odd, $p \in (1,\infty)$, and $w \in A_p$, then the sparse operator $A^N_{\mathscr{S}}$ is bounded on $L^p(w)$, and*

$$
\|A^N_{\mathscr{S}}\|_{\mathscr{L}(L^p(w))} \leqslant c_{d,p} \frac{N^{2d}}{\varepsilon} [w]_{A_p}^{\max(1,\frac{1}{p-1})}.
$$

Proof. This is an immediate consequence of Theorem 11.3.19 and Rubio de Francia's Extrapolation Theorem J.2.1. (In the latter, ϕ_{pr} and c_{pr} should be replaced by ϕ_{dpr} and c_{dpr}; the omission of dependence on d is a systematic typo in Theorem J.2.1 and its proof. This explains a need a constant $c_{d,p}$ rather than just c_p in the statement of the corollary.) □

It is also useful to record the following localised version:

Proposition 11.3.21. *Let $\varepsilon \in (0,1)$ and $\mathscr{S} \subseteq \mathscr{D}$ be ε-sparse, and let $Q_0 \in \mathscr{D}$. If $N \in \mathbb{Z}_+$ is odd and $w \in A_2(Q_0)$, then the sparse operator $A^N_{\mathscr{S}}$ is bounded on $L^2(Q_0, w)$, and*

$$
\|A^N_{\mathscr{S}}\|_{\mathscr{L}(L^2(Q_0,w))} \leqslant \Big(\frac{4}{\varepsilon} N^{2d} + 1\Big) [w]_{A_2(Q_0)}.
$$

The same result is true if the cube Q_0 is replaced by a quadrant of $\mathbb{R}^d$.

We start with a simple:

Lemma 11.3.22. *For every $Q \in \mathscr{D}(Q_0)$, there exists a cube $\widetilde{Q}$ such that $NQ \cap Q_0 \subseteq \widetilde{Q} \subseteq Q_0$ and $\ell(\widetilde{Q}) = \min\{N\ell(Q), \ell(Q_0)\}$.*

Proof. If $NQ \subseteq Q_0$, we take $\widetilde{Q} := NQ$, and if $N\ell(Q) \geqslant \ell(Q_0)$, we define $\widetilde{Q} := Q_0$.

Let us finally consider $Q \in \mathscr{D}(Q_0)$ such that $NQ \not\subseteq Q_0$ but $N\ell(Q) < \ell(Q_0)$. Let first $d = 1$, so that both $Q_0 = [a, b)$ and Q are intervals. If NQ extends to the left of a, then $\widetilde{Q} := [a, a + N\ell(Q))$ satisfies the desired properties. If NQ extends to the right of b, then $\widetilde{Q} := [b - N\ell(Q), b)$ works. For general $d \geqslant 1$ with $Q = I_1 \times \cdots \times I_d$ and $Q_0 = J_1 \times \cdots \times J_d$, we take $\widetilde{Q} := \widetilde{I}_1 \times \cdots \times \widetilde{I}_d$, where each $\widetilde{I}_i$ is built relative to the respective interval J_i as in the one-dimensional construction just given. This completes the proof. □

Proof of Proposition 11.3.21. The norm on the left is the $L^2(w)$-norm of the operator $f \mapsto \mathbf{1}_{Q_0} A^N_{\mathscr{S}}(\mathbf{1}_{Q_0} f)$, i.e., both the domain and the range of the operator is restricted to functions supported on Q_0. Since $Q_0 \in \mathscr{D}$, each $Q \in \mathscr{D} \subseteq \mathscr{S}$ that contributes to $\mathbf{1}_{Q_0} A_{\mathscr{S}}(\mathbf{1}_{Q_0} f)$ satisfies either $Q \subseteq Q_0$ or $Q \supsetneq Q_0$. Letting $\mathscr{S}' := \{Q \in \mathscr{S} : Q \subseteq Q_0\}$, we hence have

$$\begin{aligned}
&\int_{Q_0} A^N_{\mathscr{S}}(\mathbf{1}_{Q_0} f\sigma) \cdot gw \\
&\qquad \leqslant \int_{Q_0} A^N_{\mathscr{S}'}(\mathbf{1}_{Q_0} f\sigma) \cdot gw + \int_{Q_0} \sum_{Q \supsetneq Q_0} \langle \mathbf{1}_{Q_0} f \rangle_{NQ} \cdot gw =: I + II.
\end{aligned}$$

By the dual weight trick with $\sigma = w^{-1}$, estimating the left-hand side uniformly over $f \in L^2(Q_0, \sigma)$ and $g \in L^2(Q_0, w)$ of unit norm is equivalent to bounding $\|A^N_{\mathscr{S}}\|_{\mathscr{L}(L^2(Q_0,w))}$.

Term II is dominated by

$$\sum_{Q \supsetneq Q_0} \fint_{NQ} (\mathbf{1}_{Q_0} f) = \sum_{Q \supsetneq Q_0} \frac{|Q_0|}{|Q|} \fint_{NQ_0} (\mathbf{1}_{Q_0} f) = \sum_{k=1}^{\infty} 2^{-kd} \fint_{NQ_0} (\mathbf{1}_{Q_0} f),$$

where $\sum_{k=1}^{\infty} 2^{-kd} \leqslant \sum_{k=1}^{\infty} 2^{-k} = 1$. Thus

$$II \leqslant \left\| \mathbf{1}_{Q_0} \langle \mathbf{1}_{Q_0} f \rangle_{NQ_0} \right\|_{L^2(w)} \|g\|_{L^2(w)},$$

where

$$\begin{aligned}
\left\| \mathbf{1}_{Q_0} \langle \mathbf{1}_{Q_0} f \rangle_{NQ_0} \right\|_{L^2(w)} &= \frac{w(Q_0)^{1/2}}{|NQ_0|} \int_{Q_0} f w^{1/2} \sigma^{1/2} \\
&\leqslant \frac{w(Q_0)^{1/2}}{|Q_0|} \|f\|_{L^2(Q_0,w)} \sigma(Q_0)^{1/2} \\
&\leqslant [w]^{1/2}_{A_2(Q_0)} \|f\|_{L^2(Q_0,w)} \leqslant [w]_{A_2(Q_0)} \|f\|_{L^2(Q_0,w)}.
\end{aligned}$$

We then turn to the main part I involving $\mathscr{S}' := \{Q \in \mathscr{S} : Q \subseteq Q_0\}$. We can largely follow the proof of Theorem 11.3.19, but some care is needed to ensure that we only apply the A_2 condition to cubes contained in Q_0, which need not be the case with the dilated cubes NQ. We start with

$$\begin{aligned} I &= \sum_{Q\in\mathscr{S}'} \langle \mathbf{1}_{Q_0} f\sigma\rangle_{NQ} \int_{Q_0} \mathbf{1}_Q \cdot gw \\ &= \sum_{Q\in\mathscr{S}'} \frac{1}{|NQ|} \int_{NQ\cap Q_0} f\sigma \cdot \int_Q gw \\ &= \sum_{Q\in\mathscr{S}'} \frac{\sigma(NQ\cap Q_0)}{|NQ|} \frac{w(Q)}{|Q|} \langle f\rangle^{\sigma}_{NQ\cap Q_0} \cdot \langle g\rangle^w_Q |Q| \end{aligned}$$

By Lemma 11.3.22, for every $Q \in S' \subseteq \mathscr{D}(Q_0)$, there is a cube $\widetilde{Q}$ such that $Q \subseteq NQ \cap Q_0 \subseteq \widetilde{Q} \subseteq Q_0$ and $\ell(\widetilde{Q}) \leqslant N\ell(Q)$. Thus

$$\sigma(NQ\cap Q_0) \leqslant \sigma(\widetilde{Q}), \quad w(Q) \leqslant w(\widetilde{Q}), \quad |\widetilde{Q}| \leqslant |NQ| = N^d|Q|.$$

Hence

$$\frac{\sigma(NQ\cap Q_0)}{|NQ|}\frac{w(Q)}{|Q|} \leqslant \frac{\sigma(\widetilde{Q})}{|\widetilde{Q}|}\frac{w(\widetilde{Q})}{|\widetilde{Q}|} N^d \leqslant [w]_{A_2(Q_0)} N^d,$$

since $\widetilde{Q}$ is a cube contained in Q_0. Substituting back, and using sparseness, it follows that

$$\int_{Q_0} A^N_{\mathscr{S}'}(\mathbf{1}_{Q_0} f\sigma)\cdot gw \leqslant N^d [w]_{A_2(Q_0)} \sum_{Q\in\mathscr{S}'} \langle f\rangle^{\sigma}_{NQ\cap Q_0} \cdot \langle g\rangle^w_Q \frac{|E(Q)|}{\varepsilon}.$$

As in the proof of Theorem 11.3.19, we have $\langle g\rangle^w_Q \leqslant \inf_{z\in Q} M^w_{\mathscr{D}(Q_0)} g$. Also, using Proposition 11.3.11, each NQ belongs to one of the dilated dyadic systems $\mathscr{D}^{n;N}$, where $n \in \mathbb{Z}^d_N$. A key observation is that then also

$$\mathscr{C}^{n;N} := \{NQ\cap Q_0 : Q\in\mathscr{D}, NQ \in \mathscr{D}^{n;N}\}$$

is a nested family with *set-theoretic* (if not geometric) properties matching those of $\mathscr{D}(Q_0)$: Each of the subfamilies

$$\mathscr{C}^{n;N}_k := \{NQ\cap Q_0 \in \mathscr{C}^{n;N} : \ell(Q) = 2^{-k}\ell(Q_0)\}$$

is a partition of Q_0, and each $\mathscr{C}^{n;N}_{k+1}$ refines the previous $\mathscr{C}^{n;N}_k$. Thus, the corresponding maximal operators

$$M^{\sigma}_{\mathscr{C}^{n;N}} f := \sup_{R\in\mathscr{C}^{n;N}} \mathbf{1}_R \langle f\rangle^{\sigma}_R$$

are still instances of the Doob maximal operator with respect on abstract filtered spaces. Repeating the computation (11.44) *mutatis mutandis*, we then obtain

$$
\begin{aligned}
\sum_{Q\in\mathscr{S}'} \langle f\rangle^{\sigma}_{NQ\cap Q_0}\langle g\rangle^{w}_{Q}\frac{|E(Q)|}{\varepsilon}
&\leqslant \frac{1}{\varepsilon}\sum_{n\in\mathbb{Z}^d_N}\|M^{\sigma}_{\mathscr{C}^{n;N}}f\|_{L^2(\sigma)}\|M^{w}_{\mathscr{D}}g\|_{L^2(w)}\\
&\leqslant \frac{1}{\varepsilon}\sum_{n\in\mathbb{Z}^d_N}2\|f\|_{L^2(\sigma)}\cdot 2\|g\|_{L^2(w)}=\frac{4}{\varepsilon}N^d\|f\|_{L^2(\sigma)}\|g\|_{L^2(w)}.
\end{aligned}
$$

Hence

$$
I\leqslant N^d[w]_{A_2(Q_0)}\cdot\frac{4}{\varepsilon}N^d\|f\|_{L^2(\sigma)}\|g\|_{L^2(w)}.
$$

In combination with the bound

$$
II\leqslant [w]_{A_2(Q_0)}\|f\|_{L^2(\sigma)}\|g\|_{L^2(w)}.
$$

Recalling that

$$
\int_{Q_0}A^N_{\mathscr{S}}(\mathbf{1}_{Q_0}f\sigma)\cdot gw\leqslant\int_{Q_0}A^N_{\mathscr{S}'}(\mathbf{1}_{Q_0}f\sigma)\cdot gw+\int_{Q_0}\sum_{Q\supsetneq Q_0}\langle\mathbf{1}_{Q_0}f\rangle_{NQ}\cdot gw
$$

and the dual weight trick, we conclude the proof in the case of a cube.

If Q_0 is replaced by a quadrant S, we note by density that it suffices to consider the integrals above compactly supported f and g. But then, if Q_0 is a sufficiently large cube contained in the quadrant and having one corner at the corner of the quadrant, then such f and g will be supported in Q_0. Thus the previous considerations apply and give a bound in terms of $[w]_{A_2(Q_0)}$, which is clearly dominated by $[w]_{A_2(S)}$. □

An extension of Proposition 11.3.21 to $p\neq 2$ follows, in principle, by Rubio de Francia's Extrapolation Theorem J.2.1 just like Corollary 11.3.20 from Theorem 11.3.19. Since Theorem J.2.1 was formulated for global $A_p(\mathbb{R}^d)$ weights only, we include some remarks about its local version. As a rule, all dyadic considerations carry over without any change. However, one needs to play a little attention to the interplay of dyadic and non-dyadic cubes in the local setting. The following is a local variant of the Covering Lemma 3.2.26:

Lemma 11.3.23. *For cubes $Q\subseteq Q_0\subseteq\mathbb{R}^d$, there exist a vector $\alpha\in\{0,\frac13,\frac23\}^d$ and a dyadic cube*

$$
D\in\mathscr{D}^{\alpha}(Q_0):=\{P+\alpha(-1)^{\log_2\frac{\ell(P)}{\ell(Q_0)}}\ell(P):P\in\mathscr{D}(Q_0)\}\tag{11.45}
$$

(the shifted dyadic cubes from Definition 3.2.25) such that

$$
\ell(D)\leqslant 3\ell(Q)\quad\textit{and}\quad Q\subseteq D\subseteq Q_0.
$$

In (11.45), the point of the factor $(-1)^{\log_2\frac{\ell(P)}{\ell(Q_0)}}$ is simply to alternate between ±1 with each consecutive generation of the dyadic cubes. We refer the reader to the discussion preceding Lemma 3.2.26 for why such a factor is needed.

Proof. If $3\ell(Q) \geqslant \ell(Q_0)$, then clearly $D := Q_0 \in \mathscr{D}(Q_0) = \mathscr{D}^0(Q_0)$ satisfies the required properties.

Let then $3\ell(Q) < \ell(Q_0)$. By Lemma 3.2.26 (a global version of the lemma that we are proving), there exists a cube D as asserted, expect that we do not know whether $D \subseteq Q_0$ or not. If yes, then we are done, so suppose that $D \nsubseteq Q_0$. We will check that an appropriate shift of D will be a cube that we are looking for.

Let first $d = 1$ so that $Q_0 = [a, b)$ as well as Q and D are just intervals. If D extends to the left of a, then we can take $D' := [a, a + \ell(D)) \in \mathscr{D}(Q_0)$, and if D extends to the right of b, then we can take $D' := [b - \ell(D), b) \in \mathscr{D}(Q_0)$.

Let then $d \geqslant 1$ be arbitrary, $Q = I_1 \times \cdots \times I_d \subseteq D = J_1 \times \cdots \times J_d \in \mathscr{D}^\alpha(\mathbb{R}^d)$, and $Q_0 := K_1 \times \cdots \times K_d$. For each $i \in \{1, \ldots, d\}$, we run the previous construction: If $J_i \subseteq K_i$, we let $J_i' := J_i \in \mathscr{D}^{\alpha_i}(\mathbb{R})$. If $J_i \nsubseteq K_i$, we let J_i' be the interval of lengths $\ell(J_i)$ that meets the same end-point of K_i as J_i. Then $J_i' \in \mathscr{D}(K_i)$. Defining $D' := J_1' \times \cdots \times J_d'$, we have $D' \in \mathscr{D}^{\alpha'}$, where $\alpha_i' = \alpha_i$ if $J_i \subseteq K_i$ and $\alpha_i' = 0$ otherwise. This D' in place of D satisfies the claimed properties, and the proof of the lemma is complete. □

As in (3.36), we can now easily dominate the local maximal operator

$$M_{Q_0} f(x) := \sup_{\substack{Q \subseteq Q_0 \\ \text{cube}}} \mathbf{1}_Q(x) \fint_Q \|f(y)\| \, \mathrm{d}y$$

by the local dyadic maximal operators

$$M_{Q_0}^\alpha f(x) := \sup_{\substack{P \in \mathscr{D}^\alpha(Q_0) \\ P \subseteq Q_0}} \mathbf{1}_P(x) \fint_P \|f(y)\| \, \mathrm{d}y, \quad \alpha \in \{0, \tfrac{1}{3}, \tfrac{2}{3}\}^d$$

with

$$M_{Q_0} f \leqslant 3^d \max_{\alpha \in \{\{0, \frac{1}{3}, \frac{2}{3}\}^d} M_{Q_0}^\alpha f \leqslant 3^d \sum_{\alpha \in \{\{0, \frac{1}{3}, \frac{2}{3}\}^d} M_{Q_0}^\alpha f. \tag{11.46}$$

Proposition 11.3.24. *Let $p, r \in (1, \infty)$ and cube $Q_0 \subseteq \mathbb{R}^d$ be a cube. Then*

(1) $\|M_{Q_0} f\|_{L^p(Q_0, w)} \leqslant c_d p' [w]_{A_p}^{1/(p-1)} \|f\|_{L^p(Q_0, w)}$;

(2) *if a pair of functions (f, h) satisfies*

$$\|h\|_{L^r(Q_0, w)} \leqslant \phi_r([w]_{A_r(Q_0)}) \|f\|_{L^r(Q_0, w)}$$

for all $w \in A_r(Q_0)$, where ϕ_r is a non-negative increasing function, then

$$\|h\|_{L^p(Q_0, w)} \leqslant \phi_{dpr}([w]_{A_p}) \|f\|_{L^p(Q_0, w)}$$

for all $w \in A_p(Q_0)$, where each ϕ_{dpr} is a non-negative increasing function. In particular, if $\phi_r(t) = c_r t^\tau$, then $\phi_{dpr}(t) \leqslant c_{dpr} t^{\tau \max\{\frac{r-1}{p-1}, 1\}}$.

Proof. (1) follows by repeating the proof of Theorem J.1.1: the dyadic considerations are unchanged, and in the last step of the proof, one replaces an application of (3.36) by its localised version (11.46).

The proof of (2) is the same as the proof of Theorem J.2.1, except that the all references to the maximal operator M are replaced by the local version M_{Q_0} and, accordingly, all applications of Theorem J.1.1 by case (1) of the proposition that we already proved. (We note that the ϕ_{pr} and c_{pr} should be replaced by ϕ_{dpr} and c_{dpr} already in Theorem J.2.1; the omission of the dependence on d is a systematic typo in Theorem J.2.1 and its proof.) □

Corollary 11.3.25. *Let $\varepsilon \in (0,1)$ and $\mathscr{S} \subseteq \mathscr{D}$ be ε-sparse, and let $Q_0 \in \mathscr{D}$. If $N \in \mathbb{Z}_+$ is odd, $p \in (1,\infty)$, and $w \in A_p(Q_0)$, then the sparse operator $A^N_{\mathscr{S}}$ is bounded on $L^p(Q_0, w)$, and*

$$\|A^N_{\mathscr{S}}\|_{\mathscr{L}(L^p(Q_0,w))} \leqslant c_{d,p}\Big(\frac{4}{\varepsilon}N^{2d} + 1\Big)[w]_{A_p(Q_0)}^{\max(1,\frac{1}{p-1})}.$$

The same result is true if Q_0 is replaced by a quadrant of $\mathbb{R}^d$.

Proof. The case of a cube is immediate from case $p = 2$ established in Proposition 11.3.21 and extrapolation established in Proposition 11.3.24(2). The case of a quadrant follows from this by the same considerations as in the last paragraph of the proof of Proposition 11.3.21. □

Thanks to sparse domination, we also obtain the corresponding results for Calderón–Zygmund operators:

Theorem 11.3.26 (A_2 theorem). *Let X and Y be Banach spaces, $p_0 \in [1,\infty]$, and let*

$$T \in \mathscr{L}(L^{p_0}(\mathbb{R}^d;X), L^{p_0,\infty}(\mathbb{R}^d;Y))$$

with norm N_0 be an operator with a Dini kernel K. Then for every $p \in (1,\infty)$ and every $w \in A_p$, the operator T extends uniquely to

$$T \in \mathscr{L}(L^p(w;X), L^p(w;Y))$$

with norm estimate

$$\|T\|_{\mathscr{L}(L^p(w;X),L^p(w;Y))} \leqslant c_{d,p}\Big(N_0 + c_K + \|\omega_K\|_{\mathrm{Dini}}\Big)[w]_{A_p}^{\max(1,\frac{1}{p-1})}$$

where c_K, ω_K are as in Definition 11.3.1.

The result remain true if $\mathbb{R}^d$ is systematically replaced by a cube $Q_0 \subseteq \mathbb{R}^d$ or a quadrant $S \subseteq \mathbb{R}^d$, as the domain of the function spaces, in the definition of the Calderón–Zygmund constants c_K and $\|\omega_K\|_{\mathrm{Dini}}$, as well as in the definition of the weight class A_p.

Proof. Let us first consider the global case. Let $f \in L^p_c(w;X)$ be supported on a compact set F. Denoting by $\sigma = w^{-1/(p-1)}$ the dual weight, we have

$$\int \|f\| = \int_K \|f\| w^{1/p}\sigma^{1/p'} \leqslant \|f\|_{L^p(w)}\sigma(K)^{1/p'} < \infty,$$

so that $f \in L^1_c(\mathbb{R}^d;X)$ as well, and Tf is well defined by the Calderón–Zygmund theorem 11.2.5. Then Theorem 11.3.15 guarantees the existence of a $\frac{1}{5}$-sparse collection $\mathscr{S} \subseteq \mathscr{D}$ such that, pointwise almost everywhere,

$$\|Tf(x)\|_Y \leqslant c_d c_T (A^5_{\mathscr{S}}\|f\|_X)(x), \qquad c_T = N_0 + c_K + \|\omega\|_{\mathrm{Dini}}.$$

Thus, by Corollary 11.3.20, we have

$$\begin{aligned}\|Tf(x)\|_{L^p(w;Y)} &\leqslant c_d c_T \|A^5_{\mathscr{S}}(\|f\|_X)\|_{L^p(w)} \\ &\leqslant c_d c_T c_{d,p}[w]_{A_p}^{\max(1,\frac{1}{p-1})}\Big\|\|f\|_X\Big\|_{L^p(w)} \\ &= c_{d,p} c_T [w]_{A_p}^{\max(1,\frac{1}{p-1})}\|f\|_{L^p(w;X)}.\end{aligned} \tag{11.47}$$

Recalling the definition of c_T, this is the required norm estimate for T restricted to $L^p_c(w;X)$; since this subspace is dense in $L^p(w;X)$, it allows to uniquely extend T to the whole space with the same norm.

The proof in the case of a cube or a quadrant in place of $\mathbb{R}^d$ remains the same, just using the local Corollary 11.3.25 in place of Corollary 11.3.20 to replace (11.47) by

$$\begin{aligned}\|Tf(x)\|_{L^p(Q_0,w;Y)} &\leqslant c_d c_T \|A^5_{\mathscr{S}}(\|f\|_X)\|_{L^p(Q_0,w)} \\ &\leqslant c_d c_T c_{d,p}[w]_{A_p(Q_0)}^{\max(1,\frac{1}{p-1})}\Big\|\|f\|_X\Big\|_{L^p(Q_0,w)} \\ &= c_{d,p} c_T [w]_{A_p(Q_0)}^{\max(1,\frac{1}{p-1})}\|f\|_{L^p(Q_0,w;X)}.\end{aligned}$$

□

Corollary 11.3.27 (A_2 theorem for the Hilbert transform). *Let X be a UMD space, $p \in (1,\infty)$ and $w \in A_p(\mathbb{R})$. Then the Hilbert transform*

$$Hf(s) = \lim_{\varepsilon\downarrow 0}\frac{1}{\pi}\int_{|s-t|>\varepsilon}\frac{f(t)}{s-t}\,\mathrm{d}t$$

extends uniquely to $H \in \mathscr{L}(L^p(w;X))$ with

$$\|H\|_{\mathscr{L}(L^p(w;X))} \leqslant c_p [w]_{A_p}^{\max(1,\frac{1}{p-1})}\hbar_{2,X}, \quad \hbar_{2,X} := \|H\|_{\mathscr{L}(L^2(\mathbb{R};X))}.$$

Proof. Recall that the Hilbert transform is bounded on $L^2(\mathbb{R};X)$ when X is a UMD space (Theorem 5.1.13). In particular, taking $T = H$ and $p_0 = 2$

in Theorem 11.3.26, we have $N_0 \leqslant \hbar_{2,X} < \infty$, using the notation from the statement of that theorem. The kernel of the Hilbert transform is $K(s,t) = \frac{1}{s-t}$, so that $c_K = 1$ qualifies for the constant in Definition 11.3.1. Moreover,

$$\Big|\frac{1}{s-t} - \frac{1}{s'-t}\Big| = \Big|\frac{s'-s}{(s-t)(s'-t)}\Big| \leqslant 2\frac{|s'-t|}{|s-t|^2} \quad \forall |s-s'| \leqslant \frac{1}{2}|s-t|,$$

so that we can take the modulus of continuity $\omega_1(u) = 2u$ in Definition 11.3.1. Checking that $\omega_2(u) = 2u$ also works in entirely similar. Thus $\|\omega\|_{\mathrm{Dini}} = \int_0^1 2u \frac{\mathrm{d}u}{u} = 2$. Finally, it is easy to check that the norm $\hbar_{2,X} = \|H\|_{\mathscr{L}(L^2(\mathbb{R};X))}$ is *at least* 1, say by Proposition 5.2.2, which says that H acts as multiplication by $-i$ on functions with Fourier transform supported on $\mathbb{R}_+$. Thus $N_0 + c_K + \|\omega\|_{\mathrm{Dini}} \leqslant \hbar_{2,X} + 1 + 2 \leqslant 4\hbar_{2,X}$. Substituting this into the result of Theorem 11.3.26 gives the claimed bound for $\|H\|_{\mathscr{L}(L^p(w;X))}$. □

11.3.e Sharpness of the A_2 theorem

Already in the scalar-valued case $X = \mathbb{K}$, Corollary 11.3.27, and hence Theorem 11.3.26, is sharp in its dependence on the weight characteristic $[w]_{A_p}$. In order to see this, we need to know about the behaviour of $[w]_{A_p}$ for some concrete examples of weights, for which we can also estimate the weighted norm of the Hilbert transform. The following important power weights will serve this purpose:

Example 11.3.28 (Power weights). Let $\alpha \in \mathbb{R}$, $p \in (1,\infty)$, $w(x) = |x|^\alpha$ for $x \in \mathbb{R}^d$, and $\sigma(x) = w(x)^{-1/(p-1)} = |x|^{-\alpha/(p-1)}$. Then

$$w \in A_p(\mathbb{R}^d) \quad \Leftrightarrow \quad w, \sigma \in L^1_{\mathrm{loc}}(\mathbb{R}^d) \quad \Leftrightarrow \quad -d < \alpha < d(p-1),$$

and if these equivalent conditions holds, then

$$c_{d,p}[w]_{A_p} \leqslant \frac{1}{1+\alpha}\Big(\frac{1}{p-1-\alpha}\Big)^{p-1} \leqslant C_{d,p}[w]_{A_p}.$$

To verify the claims of this example, we make use of the following:

Lemma 11.3.29. *If $Q \subseteq \mathbb{R}^d$ is any cube, and $\tilde{Q}$ is a cube of the same size centred at the origin, then*

$$\fint_Q |x|^{-\gamma}\,\mathrm{d}x \leqslant \fint_{\tilde{Q}} |x|^{-\gamma}\,\mathrm{d}x \eqsim_d \frac{\ell(Q)^{-\gamma}}{d-\gamma}, \qquad \gamma \in [0,d),$$
$$\fint_Q |x|^{\gamma}\,\mathrm{d}x \geqslant \fint_{\tilde{Q}} |x|^{\gamma}\,\mathrm{d}x \eqsim_{d,\Gamma} \ell(Q)^{\gamma}, \qquad \gamma \in [0,\Gamma], \qquad \Gamma > 0.$$

Proof. Let $Q = \prod_{i=1}^d I_i$ and $\tilde{Q} = \prod_{i=1}^d \tilde{I}_d$. Then Q is the disjoint union of the sets $Q_{\mathscr{I}} := \prod_{i\in\mathscr{I}}(I_i \cap \tilde{I}_i) \times \prod_{i\in\complement\mathscr{I}}(I_j \setminus \tilde{I}_j)$, where $\mathscr{I}$ ranges over all subsets of $\{1, \ldots, d\}$, and $\complement\mathscr{I} := \{1, \ldots, d\} \setminus \mathscr{I}$. Of course $\tilde{Q}$ is a similar union over $\tilde{Q}_{\mathscr{I}}$, defined by interchanging the roles of I_j and $\tilde{I}_j$ in $Q_{\mathscr{I}}$.

Since $\ell(I_j) = \ell(\tilde{I}_j)$ is the common side-length of Q and $\tilde{Q}$, it follows that also $|I_j \setminus \tilde{I}_j| = |\tilde{I}_j \setminus I_j|$. Since $\tilde{I}_j$ is centred at the origin, if $x_j \in I_j \setminus \tilde{I}_j$ and $\tilde{x}_j \in \tilde{I}_j \setminus I_j$, then $|\tilde{x}_j| \leqslant |x_j|$.

Now all $x = (x_i)_{i=1}^d \in Q_{\mathscr{I}}$ are in measure-preserving correspondence with $\tilde{x} = (\tilde{x}_i)_{i=1}^d \in \tilde{Q}_{\mathscr{I}}$, such that $|x_i| = |\tilde{x}_i|$ for all $i \in \mathscr{I}$, and $|x_j| \geqslant |\tilde{x}_j|$ for all $j \in \complement\mathscr{I}$; hence altogether $|x| \geqslant |\tilde{x}|$.

This implies inequalities like the first ones on each line of the lemma, for $Q_{\mathscr{I}}$ and $\tilde{Q}_{\mathscr{I}}$ in place of Q and $\tilde{Q}$, and thus also these inequalities as claimed, by summing over all $\mathscr{I} \subseteq \{1, \ldots, d\}$.

To estimate the integrals over $\tilde{Q}$, we note that $B(0, \frac{1}{2}\ell(Q)) \subseteq \tilde{Q} \subseteq B(0, \frac{1}{2}\sqrt{d}\ell(Q))$, where, for $\alpha > -d$,

$$\int_{B(0,c_d\ell(Q))} |x|^\alpha \,\mathrm{d}x = \int_0^{c_d\ell(Q)} r^\alpha r^{d-1}\sigma_{d-1}\,\mathrm{d}r = \frac{(c_d\ell(Q))^{d+\alpha}}{d+\alpha}\sigma_{d-1},$$

thus

$$2^{-d-\alpha}\sigma_{d-1}\frac{\ell(Q)^\alpha}{d+\alpha} \leqslant \fint_{\tilde{Q}} |x|^\alpha \,\mathrm{d}x \leqslant (2^{-1}\sqrt{d})^{d+\alpha}\sigma_{d-1}\frac{\ell(Q)^\alpha}{d+\alpha}.$$

For $\alpha = -\gamma \in (-d, 0]$, the quantities multiplying $\ell(Q)^\alpha/(d+\alpha) = \ell(Q)^{-\gamma}/(d-\gamma)$ are clearly uniformly bounded from above and away from zero, with bounds depending on d only. Similarly, for $\alpha = \gamma \in [0, \Gamma]$, the quantities multiplying $\ell(Q)^\alpha = \ell(Q)^\gamma$ have this property, with bounds depending on d and Γ only. □

Proof of Example 11.3.28. The second $\Leftrightarrow$ in the claim is immediate.

Note that at least one of w and σ is $|x|$ to a non-negative exponent, and therefore locally integrable with a strictly positive integral over every cube Q. Thus, in order that $[w]_{A_p}$ is finite, it is necessary that the other of the two functions is locally integrable as well, showing the first $\Rightarrow$ in the claim.

It remains to check that $-d < \alpha < d(p-1)$ implies that $w \in A_p(\mathbb{R}^d)$, together with the claimed estimate for $[w]_{A_p}$.

Let first $\alpha \geqslant 0$, and denote $\delta_Q := \operatorname{dist}(Q, 0)/\ell(Q)$. For $x \in Q$, we have $|x| \leqslant (\delta_Q + \sqrt{d})\ell(Q)$, and thus $\fint_Q w \leqslant (\delta_Q + \sqrt{d})^\alpha \ell(Q)^\alpha$. If $\delta_Q > 0$, we also have $|x|^{-1} \leqslant \delta_Q^{-1}\ell(Q)^{-1}$, and hence $(\fint_Q \sigma)^{p-1} \leqslant \delta_Q^{-\alpha}\ell(Q)^{-1}$. Thus

$$\sup_{Q:\delta_Q \geqslant \delta} \fint_Q w\Big(\fint_Q \sigma\Big)^{p-1} \leqslant \sup_{Q:\delta_Q \geqslant \delta} (\delta_Q + \sqrt{d})^\alpha \delta_Q^{-\alpha} = \Big(1 + \frac{\sqrt{d}}{\delta}\Big)^\alpha$$

On the other hand, for any cube Q, it follows from Lemma 11.3.29 that

$$\begin{aligned}\sup_{Q:\delta_Q \leqslant \delta} \int_Q w\Big(\fint_Q \sigma\Big)^{p-1} &\leqslant \sup_{Q:\delta_Q \leqslant \delta} (\delta_Q + \sqrt{d})^\alpha \ell(Q)^\alpha \Big(\frac{2^d d}{d - \frac{\alpha}{p-1}}\ell(Q)^{-\frac{\alpha}{p-1}}\Big)^{p-1} \\ &= (\delta + \sqrt{d})^\alpha \Big(\frac{2^d d}{d - \frac{\alpha}{p-1}}\Big)^{p-1}\end{aligned}$$

Fixing some $\delta = \delta_{d,p}$, it is then immediate that

$$[w]_{A_p} \leqslant \frac{c_{d,p}}{(d - \frac{\alpha}{p-1})^{p-1}} = \frac{c'_{d,p}}{[d(p-1) - \alpha]^{p-1}}$$

For a matching lower bound, it is enough to consider just the unit cube Q, in which case the estimates of Lemma 11.3.29 apply with $\Gamma = d(p-1)$ to give that

$$[w]_{A_p} \geqslant \fint_Q w\Big(\fint_Q \sigma\Big)^{p-1} \eqsim_{d,p} 1 \cdot \Big(\frac{1}{d - \frac{\alpha}{p-1}}\Big)^{p-1} \eqsim_{d,p} \Big(\frac{1}{d(p-1) - \alpha}\Big)^{p-1}.$$

This completes the proof for $\alpha \in [0, d(p-1))$, noting that $\frac{1}{1+\alpha} \eqsim_{d,p} 1$ in this case.

For $\alpha = -\gamma < 0$, we note that

$$\begin{aligned} [|x|^{-\gamma}]_{A_p} = [|x|^{\frac{\gamma}{p-1}}]_{A_{p'}}^{p-1} &\eqsim_{d,p} \Big\{\Big(\frac{1}{d(p'-1) - \frac{\gamma}{p-1}}\Big)^{p'-1}\Big\}^{p-1} \\ &= \frac{p-1}{d-\gamma} \eqsim_{d,p} \frac{1}{d+\alpha} \end{aligned}$$

by applying the previous case to $\frac{\gamma}{p-1} \geqslant 0$ and p' in place of α and p, and noting that $(p-1)(p'-1) = 1$. □

We are now fully equipped to confirm the sharpness of Corollary 11.3.27.

Proposition 11.3.30 (Buckley). *Fix $p \in (1, \infty)$, and suppose that $\phi : [1, \infty) \to [1, \infty)$ is an increasing function such that*

$$\|H\|_{\mathscr{L}(L^p(w))} \leqslant \phi([w]_{A_p}) \qquad \forall w \in A_p,$$

or even just for all power weights in A_p. Then

$$\phi(t) \geqslant c_p \cdot t^{\max(1, \frac{1}{p-1})} \qquad \forall t \geqslant 1.$$

Proof. Let $\sigma = w^{-1/(p-1)}$ denote the dual weight. Using the dualised formulation (11.43) of the $L^p(w)$-boundedness of $T = H$, and choosing f and g with positively separated compact supports, so that the kernel representation is available, we have

$$\frac{1}{\pi} \iint \frac{f(y)\sigma(y)g(x)w(x)}{x-y} \,\mathrm{d}x \,\mathrm{d}y \leqslant \phi([w]_{A_p}) \|f\|_{L^p(\sigma)} \|g\|_{L^{p'}(w)} \tag{11.48}$$

for all such f and g. If these functions are non-negative with $\operatorname{supp} f \subseteq \mathbb{R}_-$ and $\operatorname{supp} g \subseteq \mathbb{R}_+$, then the integrand is non-negative, and by monotone convergence (11.48) persists even if the supports of f and g meet at the origin.

The crucial point in bounding the Hilbert transform form below is the following observation: if $h(y) = |y|^{-\alpha}\mathbf{1}_{(-1,0)}(y)$, then for $x \in (0, 1)$,

$$Hh(x) = \frac{1}{\pi}\int_0^1 \frac{y^{-\alpha}}{x+y}\,\mathrm{d}y \geqslant \frac{1}{\pi}\int_0^x \frac{y^{-\alpha}}{2x}\,\mathrm{d}y = \frac{1}{2\pi}\frac{x^{-\alpha}}{1-\alpha}, \tag{11.49}$$

which is essentially h again, but with a factor $\frac{1}{1-\alpha}$ that blows up as $\alpha \to 1-$.

We now "test" (11.48) with two choices of (f, g, σ, w), so that $(f\sigma, gw)$ is either $(|y|^{-\alpha}\mathbf{1}_{(-1,0)}, \mathbf{1}_{(0,1)})$ or $(\mathbf{1}_{(-1,0)}, |y|^{-\alpha}\mathbf{1}_{(0,1)})$, with $\alpha \in [0, 1)$. In either case (11.49) shows that

$$LHS(11.48) \geqslant \frac{1}{2\pi}\int_0^1 \frac{x^{-\alpha}}{1-\alpha}\,\mathrm{d}x = \frac{1}{2\pi}\frac{1}{(1-\alpha)^2},$$

where we have accumulated a quadratic blow-up.

To estimate the right hand side of (11.48), we need to specify the individual functions, not just the products $f\sigma$ and gw. In the first case, let $f = \mathbf{1}_{(-1,0)}$ and $\sigma(y) = w(y)^{-1/(p-1)} = |y|^{-\alpha}$; thus $w(y) = |y|^{\alpha(p-1)}$ and $g(y) = \mathbf{1}_{(0,1)}(y)w(y)^{-1} = \mathbf{1}_{(0,1)}(y)|y|^{-\alpha(p-1)}$. Then

$$\begin{aligned}\|f\|_{L^p(\sigma)}\|g\|_{L^{p'}(w)} &= \Big(\int_0^1 x^{-\alpha}\,\mathrm{d}x\Big)^{1/p}\Big(\int_0^1 x^{\alpha(p-1)(1-p')}\,\mathrm{d}x\Big)^{1/p'} \\ &= 1/(1-\alpha).\end{aligned} \tag{11.50}$$

noting that $(p-1)(p'-1) = 1$, and Example 11.3.28 shows that $[w]_{A_p} \leqslant c_p/(1-\alpha)^{p-1}$. Thus, altogether, we have

$$\frac{1}{2\pi}\frac{1}{(1-\alpha)^2} \leqslant (11.48) \leqslant \phi\Big(\frac{c_p}{(1-\alpha)^{p-1}}\Big)\frac{1}{1-\alpha}. \tag{11.51}$$

Denoting $t = c_p/(1-\alpha)^{p-1}$, this reduces to

$$\phi(t) \geqslant \tilde{c}_p t^{1/(p-1)} \qquad \forall t \geqslant c_p. \tag{11.52}$$

Since $H^2 = -I$, it is clear that $\|H\|_{\mathscr{L}(L^p(w))} \geqslant 1$, and hence $\phi(t) \geqslant 1 \geqslant c_p' t^{1/(p-1)}$ for $t \in [1, c_p)$ as well.

In the second case, we take $g = \mathbf{1}_{(0,1)}$ and $w(x) = \sigma(x)^{1-p} = |x|^{-\alpha}$; thus $\sigma(x) = |x|^{\alpha/(p-1)} = |x|^{\alpha(p'-1)}$ and $f(x) = \mathbf{1}_{(-1,0)}(x)|x|^{-\alpha(p'-1)}$. A computation like (11.50) gives exactly the same final result, only with a slightly different intermediate step, and Example 11.3.28 shows that $[w]_{A_p} \leqslant c_p/(1-\alpha)$. With this quantity inside ϕ in (11.51), the substitution $t = c_p/(1-\alpha)$ then gives

$$\phi(t) \geqslant \tilde{c}_p t \qquad \forall t \geqslant c_p, \tag{11.53}$$

and the same bound for $t \in [1, c_p)$ follows from $H^2 = -I$ as before. The two lower bounds (11.52) and (11.53) together prove the proposition. □

11.4 Notes

Given the emphasis of these volumes in analysis of functions having their *range* in a Banach space, we have chosen to keep the consideration related to the *domain* of the functions relatively simple, concentrating on the canonical case of the Euclidean space $\mathbb{R}^d$ and, with specific applications in the later chapters in mind, its rather special subdomains—cubes and quadrants—only. However, much of this theory could be developed on far more general domains, notably on *spaces of homogeneous type* (espaces de nature homogène) introduced by Coifman and Weiss [1971] and extensively studied ever since. Since our treatment is heavily based on the dyadic cubes on $\mathbb{R}^d$, we recall that analogous constructions are also available in the mentioned generality. The construction of a fixed family of sets, sharing the essential properties of the standard dyadic cubes of $\mathbb{R}^d$, is due to Christ [1990]. We also make use of "adjacent" and "random" families of dyadic cubes; a reasonably comprehensive account of their analogues in spaces of homogeneous type is provided by Hytönen and Kairema [2012] with several variants and elaborations due to Auscher and Hytönen [2013], Hytönen and Martikainen [2012], Hytönen and Tapiola [2014], and Nazarov, Reznikov, and Volberg [2013].

Section 11.1

This section deals with relatively classical topics but with some modern flavour. In particular, the local oscillation decomposition of Theorem 11.1.12 dates essentially back to Lerner [2010] in the scalar-valued case. The vector-valued generalisation, introducing the notion of λ-pseudomedian, was first found by Hänninen and Hytönen [2014]. Our present proof streamlines the original one.

Proposition 11.1.14 was proved by Katz and Pereyra [1999] in the scalar-valued case via a multilinear estimate, and by Hänninen and Hytönen [2016] as stated.

Theorem 11.1.30 on the vector-valued H^1–BMO duality is essentially from Bourgain [1986], although the present proof is different. In this circle of ideas, we have only covered the relatively elementary part of the theory that does not require any assumptions on the underlying Banach space. Note that Theorem 11.1.30 says that $\mathrm{BMO}_{\mathscr{D}}(\mathbb{R}^d; X^*)$ can be identified with an isometric subspace of $(H^1_{\mathscr{D},\mathrm{at}}(\mathbb{R}^d; X))^*$. The same proof works in the non-dyadic case, where arbitrary cubes are allowed both in the definition of BMO and of the Hardy space atoms. To describe the full dual $(H^1_{\mathrm{at}}(\mathbb{R}^d; X))^*$, Blasco [1988] defines a class of Banach space Y-valued measures $\mathscr{BMO}(\mathbb{R}^d; Y)$. Among other things, he shows that $(H^1_{\mathrm{at}}(\mathbb{R}^d; X))^* = \mathscr{BMO}(\mathbb{R}^d; X^*)$ for every Banach space X, whereas $\mathscr{BMO}(\mathbb{R}^d; Y) = \mathrm{BMO}(\mathbb{R}^d; Y)$, if and only if Y has the Radon–Nikodým property. A recent account with more information on the Banach space valued H^1 and BMO can be found in Chapter 7 of Pisier [2016].

Section 11.2

The material of this section is predominantly classical, and most of the results would have been available in essentially the present form by the 1980's, if not earlier, even in the Banach space valued setting. The scalar-valued origins, of course, date much further back.

The essence of Theorem 11.2.5 comes from Calderón and Zygmund [1952], who consider the scalar-valued case ($X = Y = \mathscr{L}(X, Y) = \mathbb{C}$) and Dini kernels of the special form $K(x, y) = K(x - y) = |x - y|^{-d} \Omega\big(\frac{x-y}{|x-y|}\big)$, where moreover $\int_{S^{n-1}} \Omega \, \mathrm{d}\sigma = 0$. In contrast to Theorem 11.2.5, which extrapolates other L^p-bounds from an assumed *a priori* L^{p_0}-bound, Calderón and Zygmund [1952] obtained their L^p-boundedness conclusions unconditionally, i.e., they also deduce the initial L^{p_0}-bound for $p_0 = 2$ from their special assumptions on the kernel. Once this is achieved, the extrapolation to other L^p-bounds is carried out in much the same way as in the present treatment, particularly in the case $p < p_0$. The fact that the extrapolation part of Calderón and Zygmund [1952] argument remains valid under more general assumptions on the kernel was observed by Hörmander [1960], who introduced the conditions, now bearing his name, in Definition 11.2.1 in the case of scalar-valued convolution kernels $K(x, y) = \mathfrak{K}(x - y)$. What we have called the (operator-)Hörmander class Hör was designated as K^1 by Hörmander [1960], who also defines a family of related conditions K^a with a parameter $a \in [1, \infty]$. Just like Hör $= K^1$ is relevant for the extrapolation of L^p-boundedness, the condition K^a permits the extrapolation of L^p-to-L^q boundedness from one pair (p, q) with $\frac{1}{p} - \frac{1}{q} = 1 - \frac{1}{a}$ to other such pairs.

The first Banach space-valued generalisations, which used the operator-Hörmander conditions, were found by Schwartz [1961] and, apparently independently, by Benedek, Calderón, and Panzone [1962]. According to García-Cuerva and Rubio de Francia [1985], the fact that the mere Hörmander condition (involving integrals of $\|K(s, t)x - K(s', t)x\|_Y$ rather than $\|K(s, t) - K(s', t)\|_{\mathscr{L}(X,Y)}$) is sufficient for results like Theorem 11.2.5 "should have been observed by anyone trying to adapt the proof of [the Calderón–Zygmund theorem] to the vector valued case", yet they "do not emphasize very much the interest of this weaker condition since, in most of the applications of vector valued singular integrals, [the operator Hörmander condition] does hold." Rubio de Francia, Ruiz, and Torrea [1986] provided, in their own words, an "updated review" of Benedek et al. [1962], incorporating several new developments in singular integrals into the vector-valued theory, and in particular explicitly dealing with two-variable kernels $K(s, t)$, as we have done here. Our considerations related to c_0 in Theorem 11.2.9 were inspired by Girardi and Weis [2004].

A version of Theorem 11.2.5 for convolution kernels $K(s, t) = \mathfrak{K}(s - t)$ is also presented by Grafakos [2008], where (in contrast to our approach) the upper extrapolation is achieved by a duality argument, and the interested reader is referred to this work for details of that approach. Grafakos [2008] is also explicit about the norm estimate in Theorem 11.2.5(3); this is certainly well known, but often not spelled out in many references.

Section 11.3

The main body of this section consists of results from the 2010's. Since the discovery of the original forms of many of these results, there has been significant activity in generalising and streamlining their proofs, as well as developing entirely new approaches. As a result, our order of presentation deviates from the historical timeline in favour of a smoother mathematical story. *A* main result of this section is certainly the A_2 Theorem 11.3.26, but the various Sparse Domination Theorems 11.3.6, 11.3.14, and 11.3.15, originally developed as tools for proving the A_2 Theorem 11.3.26, have by now established themselves as results of intrinsic value and models for desirable type of domination to search for in other situations.

Prehistory of the A_2 theorem

In its scalar-valued and qualitative form (i.e., saying that T is bounded on $L^p(w)$, but without tracking the estimate for the operator norm), the result goes back to Hunt, Muckenhoupt, and Wheeden [1973] in the special case that T is the Hilbert transform (as in Corollary 11.3.27) and to Coifman and Fefferman [1974] for all standard Calderón–Zygmund operators of convolution type. The question of sharp dependence of the weighted operator norms $\|T\|_{\mathscr{L}(L^p(w))}$ on the weights constant $[w]_{A_p}$ was raised by Buckley [1993], who settled the case of the Hardy–Littlewood maximal operator (Theorem J.1.1) and obtained non-matching upper and lower bounds for Calderón–Zygmund operators. In particular, Proposition 11.3.30 saying that an estimate for $\|T\|_{\mathscr{L}(L^p(w))}$ can be no better than $[w]_{A_p}^{\max(1,\frac{1}{p-1})}$, is essentially from Buckley [1993]. In many papers, results of this type a stated in a slightly weaker form along the lines that "the power of $[w]_{A_p}$ can be no better than $\max(1,\frac{1}{p-1})$". However, in some related questions, the sharp estimate is known to exhibit behaviour different from a pure power law.

The question of Buckley [1993] gained new interest through the work of Astala, Iwaniec, and Saksman [2001], who considered the following problem: Let $\mathcal{O} \subseteq \mathbb{C}$ be a domain and $k \in (0,1)$. What is the minimal q such that all functions $f \in W^{1,q}_{\mathrm{loc}}(\mathcal{O})$ with $|\bar\partial f| \leqslant k|\partial f|$ (referred to as *weakly quasiregular*) must in fact belong to $f \in W^{1,2}_{\mathrm{loc}}(\mathcal{O})$ (and then be called simply *quasiregular*)? By results of Astala [1994], $q > 1 + k$ suffices; by examples due to Iwaniec and Martin [1993], $q < 1 + k$ does not, leaving $q = 1 + k$ as the critical case. Astala, Iwaniec, and Saksman [2001] proved that $q = 1+k$ is still sufficient for the said self-improvement, under their *conjecture* that the Beurling–Ahlfors transform

$$Bf(z) := -\frac{1}{\pi}\lim_{\varepsilon\to 0}\int_{\mathbb{C}\setminus D(z,\varepsilon)} \frac{f(y)\,\mathrm{d}A(y)}{(z-y)^2}, \quad D(z,\varepsilon) := \{y \in \mathbb{C} : |y-z| < \varepsilon\}$$

satisfies the upper bound

$$\|B\|_{\mathscr{L}(L^p(w))} \leqslant c_p[w]_{A_p}, \qquad p \in [2, \infty). \tag{11.54}$$

Special cases of the A_2 theorem

Shortly after being posed, the conjecture of Astala et al. [2001] was verified by Petermichl and Volberg [2002], and another proof was found by Dragičević and Volberg [2003]. Already Petermichl and Volberg [2002] observed that (11.54) as stated may be derived from its special case $p = 2$ by keeping track of the constants in the proof of Rubio de Francia's extrapolation theorem as presented, e.g., by García-Cuerva and Rubio de Francia [1985]. This idea was systematised by Dragičević, Grafakos, Pereyra, and Petermichl [2005], whose results were treated in Appendix J and applied in the section under discussion.

The positive results for the Beurling–Ahlfors transform inspired the question of sharp weighted bounds for other operators, and the special role of the exponent $p = 2$ as the critical case for extrapolation gave rise to the name "A_2 conjecture", several further cases of which were settled over the next few years. In particular, the Hilbert transform (the scalar-valued case of Corollary 11.3.27) and the Riesz transforms were handled by Petermichl [2007, 2008], a general class of sufficiently smooth odd kernels on $\mathbb{R}$ by Vagharshakyan [2010], and powers of the Beurling–Ahlfors operator by Dragičević [2011]. All these results relied on

(A) *ad hoc* representation formulas of special singular integrals in terms of simple "dyadic shifts" as in the representation of Petermichl [2000] for the Hilbert transform (see Theorem 5.1.13 and (5.20)), and
(B) Bellman function techniques for sharp weighted bounds of these shifts.

The component (B) behind these results was first challenged by Lacey, Petermichl, and Reguera [2010], who replaced it with

(C) "corona decompositions" to verify the "testing conditions" in a
(D) dyadic two-weight $T(1)$ theorem of Nazarov, Treil, and Volberg [2008].

Shortly after, a much simpler alternative to either (B) or (C)–(D) was found by Cruz-Uribe, Martell, and Pérez [2010], who in turn replaced it by methods largely similar to the ones that we have used here:

(E) domination of dyadic shifts from (A) (not yet of singular integrals directly) by the sparse operators $A_{\mathscr{S}}$, and
(F) estimating $\|A_{\mathscr{S}}\|_{\mathscr{L}(L^2(w))}$ as in Theorem 11.3.19, whose proof follows closely the original one from Cruz-Uribe et al. [2010],

However, component (A) of the original proofs remained unchallenged and, being somewhat *ad hoc* for the specific singular integrals considered thus far, restricted their extension to wider classes of operators.

The general A_2 theorem

These limitations of (A) were overcome by Hytönen [2012], who found

(G) a general *dyadic representation formula* (a variant of which will be presented in Theorem 12.4.27) of all standard Calderón–Zygmund operators in terms of a series of dyadic shifts of increasing complexity.

Moreover, (C) and (D) had to be replaced by

(C′) refinements of (C) to control the general shifts produced by (G), and
(D′) a difficult two-weight $T(1)$ theorem of Pérez, Treil, and Volberg [2010] about the singular integral itself, rather than the dyadic shifts as in (D).

A combination of (G), (C′), and (D′) gave the first proof of the A_2 Theorem 11.3.26 for all standard Calderón–Zygmund operators in the scalar case.

In a matter of months since the announcement of Hytönen [2012] in 7/2010, several variants and extensions were found. Streamlined versions and certain improvements of the original approach were obtained in Hytönen, Pérez, Treil, and Volberg [2014], Hytönen and Pérez [2013], and Hytönen [2017], which appeared in arXiv in 10/2010, 3/2011, and 8/2011, respectively. At the same time, alternatives to (C′) and (D′) by

(B′) elaborations of (B) with good control on the shift complexity

were obtained by Nazarov and Volberg [2013] (arXiv 4/2011) and Treil [2013] (arXiv 5/2011), and these were used by Nazarov, Reznikov, and Volberg [2013] (arXiv 6/2011) to give an extension of the A_2 theorem to doubling metric space domains in place of $\mathbb{R}^d$. (Thus, the versions with a cube or a quadrant that we have stated in Theorem 11.3.26 are but very particular instances of the general domains in which the result may be formulated.)

Still over the same hectic months, Hytönen, Lacey, Martikainen, Orponen, Reguera, Sawyer, and Uriarte-Tuero [2012] (arXiv 3/2011) combined the approach of Hytönen [2012] with input from the time–frequency techniques of Lacey and Thiele [2000] to extend the A_2 theorem to *maximally truncated Calderón–Zygmund operators*

$$T_{\#}f(x) = \sup_{\varepsilon>0} \|T_\varepsilon f(x)\|, \qquad T_\varepsilon f(x) = \int_{|x-y|>\varepsilon} K(x,y)f(y)\,\mathrm{d}y. \tag{11.55}$$

However, these results were shortly superseded by Hytönen and Lacey [2012] (arXiv 6/2011) by a new approach combining (G) with elaborations of (E) and (F) from the approach of Cruz-Uribe et al. [2010]:

(E′) domination of the general dyadic shifts from (G) by operators (essentially like) $A^N_{\mathscr{S}}$, where arbitrarily large N appear, and
(F′) estimating $\|A^N_{\mathscr{S}}\|_{\mathscr{L}(L^2(w))}$ with bounds polynomial in $\log N$ (which requires much more delicate analysis than Theorem 11.3.19).

As a curiosity, the term "sparse" in its present usage seems to have been introduced by Hytönen and Lacey [2012] (line below $(*)$ on page 2042). This was pointed out by Andrei Lerner in his survey talk at the "AIM Workshop on sparse domination of singular integrals" in San José, California, in 10/2017.

Simpler proofs

The difficulties with arbitrarily high shift complexity N, which seemed unavoidable in the general A_2 theorem until this point, were finally eliminated by Lerner [2013a,b] (arXiv 2/2012). These papers provide two different proofs of the same main result, stating that

$$\|T_{\#} f\|_F \leqslant c_{d,T} \sup_{\mathscr{D},\mathscr{S}} \|A_{\mathscr{S}} f\|_F, \tag{11.56}$$

where $T_{\#}$ is the maximal truncation (11.55) of a standard Calderón–Zygmund operator, F is any Banach function space of $\mathbb{R}^d$, and the supremum is taken over all dyadic systems $\mathscr{D}$ and their sparse subcollections $\mathscr{S}$. With T in place of $T_{\#}$, this is slightly weaker than the pointwise estimate of Theorem 11.3.15 but, taking $F = L^p(w)$, quite sufficient for bounding T (or $T_{\#}$) on $L^p(w)$.

The first proof of (11.56) by Lerner [2013a] still started with (G) and (E′), but then proceeded with the key new idea of

(H) domination of the adjoints $(A^N_{\mathscr{S}})^*$ by the simple operators $A_{\mathscr{S}} = A^*_{\mathscr{S}}$.

(The fact that the argument passes through the adjoint is where the Banach function space F is needed, while everything else can be estimated pointwise.) The A_2 estimate can then be completed by the simple step (F).

At the same time, Hytönen, Lacey, and Pérez [2013] found a way of replacing the initial steps (G) and (E′) by

(I) direct domination of the singular integral by an infinite series of operators (essentially like) $A^N_{\mathscr{S}}$ with arbitrarily large N.

Thus, a self-contained proof of the A_2 theorem is obtained by concatenating the steps (I), (H), and (F), and these constitute the *simple proof of the A_2 conjecture* presented by Lerner [2013b]. As soon as things started falling into the right place, the progress was very fast, and the preprints of the just discussed papers appeared in the arXiv essentially over a weekend in February 2012: Lerner [2013a] on Thursday 9th, Hytönen et al. [2013] on Friday 10th, and Lerner [2013b] on Monday 13th.

The simple proof of Lerner [2013b] also admitted the first extension of the A_2 theorem to the weighted Bochner space $L^p(w; X)$ by Hänninen and Hytönen [2014]. At the time, the main difficulty with this Banach space valued extension was the dependence of the sparse domination (I), via its use of Lerner's local oscillation formula (Theorem 11.1.12), on the notion of median. Thus, a workable vector-valued version of this concept had to be developed; it is reproduced in Section 11.1.

Pointwise sparse domination

Although not a necessity for proving the A_2 theorem, the possibility of replacing (11.56) by pointwise domination presented itself as a natural question, which attracted some interest. This was independently achieved by Conde-Alonso and Rey [2016] (arXiv 9/2014) and Lerner and Nazarov [2019] (also announced and circulated around the same time in 2014, although in arXiv only in 8/2015). These results still slightly deviated from Theorem 11.3.15 by requiring a stronger form of the Dini condition,

$$\int_0^{1/2} \omega(t) \log_2 \Big(\frac{1}{t}\Big) \frac{\mathrm{d}t}{t} < \infty.$$

All Dini kernels were first covered by the "elementary" (but not so easy) proof of Lacey [2017] (arXiv 1/2015), which was further quantified (in terms of dependence on $\|\omega\|_{\mathrm{Dini}}$) by Hytönen, Roncal, and Tapiola [2017] (arXiv 10/2015) and remarkably simplified again by Lerner [2016] (arXiv 12/2015). In proving Theorem 11.3.15, we have followed the further simplification due to Lerner and Ombrosi [2020]. One advantage of their approach is a reduction of the prerequisites from classical Calderón–Zygmund theory necessary to run their argument. On the technical level, this is achieved by replacing the maximal operator

$$M_T f(x) = \sup_{Q \ni x} \sup_{y \in Q} T(\mathbf{1}_{\complement 5Q} f)(y)$$

of Lerner [2016] by its "sharp" version $M_T^{\#}$ defined in (11.28). While $M_T^{\#}$ can be estimated relatively directly, bounding the larger $M_T f$ originally required non-trivial classical results about the maximal truncations (11.55). However, it was later observed by Almeida, Betancor, Fariña, and Rodríguez-Mesa [2022] that the bounds for the two operators are actually equivalent under general assumptions only involving the bounds for T that are used in the theory anyway. Although not explicitly discussed by Lerner and Ombrosi [2020], the present vector-valued extensions of their results, leading to Theorems 11.3.15 and 11.3.26, involved little additional effort; this is in contrast to the first vector-valued A_2 theorem by Hänninen and Hytönen [2014]. Further abstractions are due to Lorist [2021] and Lerner, Lorist, and Ombrosi [2022]; the latter work also explicitly addresses the vector-valued case.

Routes to sharpness in weighted estimates

There are some alternative routes to see the sharpness result of Proposition 11.3.30, which goes back to Buckley [1993] well before the matching upper bounds were known. Luque, Pérez, and Rela [2015] made the curious observation that this can also be achieved without exhibiting any explicit examples in the weighted situation, but studying instead the asymptotics of the *unweighted* norms $\|T\|_{L^p \to L^p}$ as $p \to 1$ and $p \to \infty$. This depends on a variant of Rubio de Francia's Extrapolation Theorem J.2.1, where one keeps track of the p-dependence in the estimates for $\|T\|_{L^p\to L^p}$ given by extrapolating a bound of the type

$$\|T\|_{L^{p_0}(w)\to L^{p_0}(w)} \leqslant \phi([w]_{A_{q_0}}),$$

where q_0 can also be different from p_0. Via contraposition, a lower bound for $\|T\|_{L^p\to L^p}$ imposes a lower bound for ϕ. This quantitative weighted-to-unweighted extrapolation was already used earlier by Fefferman and Pipher [1997] in the "positive" direction to obtain sharp unweighted L^p-norm asymptotics for some operators by studying their weighted behaviour. They also obtained a certain predecessor of the A_2 Theorem 11.3.26 with $\|T\|_{\mathscr{L}(L^2(w))} \leqslant c_{d,T}[w]_{A_1}$, where

$$\begin{aligned}[w]_{A_1} := \|Mw/w\|_\infty &= \sup_Q \fint_Q w\Big(\operatorname*{ess\,sup}_Q w^{-1}\Big) \\ &\geqslant \sup_Q \fint_Q w\Big(\fint_Q w^{-1/(p-1)}\Big)^{p-1} = [w]_{A_p} \quad \forall p\in(1,\infty).\end{aligned}$$

Further results

For a while, it might have seemed that the new sharp weighted technology was essentially restricted to the class of Calderón–Zygmund operators. A certain discouragement against further extensions came from an observation of Orponen [2013] that *if* an operator T has a dyadic representation (G) in the sense of Hytönen [2012], *then* T must necessarily be a Calderón–Zygmund operator. However, as soon as the role of (G) in the A_2 theorem was challenged by other methods, the door was also open for extensions beyond the standard Calderón–Zygmund realm. Nevertheless, few could probably have expected how far this theory could indeed be extended.

As an application of the sharp weighted estimates for Dini kernels discussed above, Hytönen, Roncal, and Tapiola [2017] (arXiv 10/2015) showed that rough homogeneous singular integrals

$$T_\Omega f(s) := \mathrm{p.\,v.}\int_{\mathbb{R}^d} \frac{\Omega(t/|t|)}{|t|^d} f(s-t)\,\mathrm{d}t, \quad \Omega\in L_0^\infty(S^{d-1}).$$

satisfy the weighted norm inequality

$$\|T_\Omega\|_{\mathscr{L}(L^2(w))} \leqslant c_d\|\Omega\|_\infty \phi([w]_{A_2})$$

with $\phi(u)\leqslant u^2$. Although dealing with a class of operators outside the direct scope of the sparse domination technology of the time, this result may nevertheless be seen as stretching those methods, rather than introducing genuinely new ones, in that the operator T_Ω was decomposed into a series of pieces in the scope of the previously available tools by following a classical approach to qualitative versions of similar results by Duoandikoetxea and Rubio de Francia

[1986], and Watson [1990]. A more intrinsic approach has been subsequently developed by Conde-Alonso, Culiuc, Di Plinio, and Ou [2017], but $\phi(u) \leqslant u^2$ seems to remain the best available bound at the time of writing. In the other direction, Honzík [2023] constructed examples of symbols Ω and weights w to show that $\phi(u) \geqslant u^{3/2}$; hence the quantitative behaviour of T_Ω is definitely different from the linear A_2 theorem for standard Calderón–Zygmund operators, but their precise bounds remain open.

Already a few weeks before Hytönen, Roncal, and Tapiola [2017] (late 10/2015 in arXiv), a far-reaching approach to sparse domination of a wide class of operators had been revealed by Bernicot, Frey, and Petermichl [2016] (early 10/2015 in arXiv). They observed that several operators that act boundedly in L^p only in some range $(p_0, q_0) \subsetneq (1,\infty)$ (and thus are definitely outside the Calderón–Zygmund class by Theorem 11.2.5) can be proved to possess *sparse form domination* of the type

$$|\langle Tf, g\rangle| \leqslant C \sum_{Q\in\mathscr{S}} |Q| \Big(\fint_{5Q} |f|^{p_0}\Big)^{1/p_0} \Big(\fint_{5Q} |g|^{q_0'}\Big)^{1/q_0'}.$$

This in turn implies weighted norm inequalities of the form

$$\|Tf\|_{L^p(w)} \leqslant C\big([w]_{A_{p/p_0}} [w]_{\mathrm{RH}_{(q_0/p)'}}\big)^{\alpha} \|f\|_{L^p(q)}, \quad p \in (p_0, q_0),$$

where $[w]_{\mathrm{RH}_t}$ is the best constant in the *reverse Hölder inequality*

$$\Big(\fint_Q w^t\Big)^{1/t} \leqslant C \fint_Q w,$$

and $\alpha = \alpha(p_0, q_0, p)$ is a certain explicit exponent depending on the indicated quantities only.

Typical examples in the scope of the theory of Bernicot, Frey, and Petermichl [2016] are various "singular non-integral operators" arising in harmonic analysis adapted to operators other than the classical Laplacian, e.g., generalised Riesz transforms $\nabla L^{-1/2}$, where L could be a second-order divergence-form operator $L = -\operatorname{div}(A\nabla)$ with bounded coefficient matrix A, or a Schrödinger operator $L = -\Delta + V$ with some potential V.

After the key observation that it is possible to go beyond Calderón–Zygmund theory at all, sparse domination results and weighted norm inequalities, as a corollary, for several different types of operators have been obtained:

- rough singular integrals (Conde-Alonso, Culiuc, Di Plinio, and Ou [2017], Di Plinio, Hytönen, and Li [2020a]);
- Bochner–Riesz multipliers (Benea, Bernicot, and Luque [2017], Conde-Alonso et al. [2017], Lacey, Mena, and Reguera [2019]);
- oscillatory integrals (Lacey and Spencer [2017], Krause, Lacey, and Wierdl [2019]);
- bilinear Hilbert transforms and related phase-space objects (Culiuc, Di Plinio, and Ou [2018a], Di Plinio, Do, and Uraltsev [2018]);

- singular integrals along curves, Radon transforms (Cladek and Ou [2018], Culiuc, Kesler, and Lacey [2019], Oberlin [2019], Anderson, Hu, and Roos [2021]);
- spherical maximal operators both on $\mathbb{R}^d$ (Lacey [2019], Beltran, Oberlin, Roncal, Seeger, and Stovall [2022a], Borges, Foster, Ou, Pipher, and Zhou [2023]) and on the Heisenberg group (Bagchi, Hait, Roncal, and Thangavelu [2021], Ganguly and Thangavelu [2021]);
- pseudo-differential operators (Beltran and Cladek [2020]).

A relatively general theory has been developed by Beltran, Roos, and Seeger [2022b], who also explicitly discuss Banach space valued operators.

Product space theory

A related direction, in which a weighted theory of singular integrals is well developed since the works of Fefferman and Stein [1982] and Fefferman [1987, 1988], yet the sparse domination technology has met obstacles, consists of the theory of product space or multi-parameter singular integrals modelled after the product Hilbert transform $H_1 \otimes H_2$ (where H_i denotes the Hilbert transform in the ith variable of $\mathbb{R}^2$). Natural maximal operators in this theory are the *strong maximal operator*

$$M_* f(s) := \sup_{R \text{ rectangle}} \mathbf{1}_R(s) \fint_R \|f(t)\| \, \mathrm{d}t.$$

and its dyadic version, where the rectangles are restricted to be dyadic (i.e., products of dyadic intervals). Barron, Conde-Alonso, Ou, and Rey [2019] have shown that it is impossible to dominate the strong dyadic maximal operator by sparse forms based on rectangles with sides parallel to the axes, which presents an obstacle to sparse techniques in this setting. While the most obvious extension of sparse domination is thus excluded, it was shown by Barron and Pipher [2017] that one can still obtain a workable substitute by replacing the dominating averages $\fint_R |f|$ of f with the averages $\fint_R Sf$ of its dyadic square function Sf on the right-hand side.

On the other hand, the original dyadic representation (G), while largely superseded by sparse technology in applications to standard Calderón–Zygmund operators, remains available, after natural modifications, in the product space theory, as first proved by Martikainen [2012b] in the two-parameter case and extended to arbitrarily many parameters by Ou [2017]. A vector-valued approach to this theory has been developed by Hytönen, Martikainen, and Vuorinen [2019a].

Sparse domination versus causality

While the current mainstream in sparse domination, evidenced by the previous list, consists of proving and applying domination for ever wider classes of operators, one may also pose a somewhat opposite question: Suppose that a given (say, standard Calderón–Zygmund) operator T possesses some additional properties. Can this be reflected in the dominating sparse operator as well? A concrete instance of such an additional property is causality. Suppose for simplicity that $d = 1$, and that $K(s,t)$ is non-zero only if $s > r$; thus $Tf(s)$ depends only on the "past" values $f(t)$ with $t < s$. If T is a Calderón–Zygmund operator, then it satisfies the sparse domination $Tf(s) \leqslant c_T A^5_{\mathscr{S}} f(s)$ by the general theory. However, the dominating sparse operator $A^5_{\mathscr{S}}$ is no longer causal. Is it possible to exploit the causality of T to obtain a sharper form of sparse domination, where this causality is preserved also in the right-hand side? Some partial (but far from complete) results in this direction have been obtained by Hytönen and Rosén [2023].

Aimar, Forzani, and Martín-Reyes [1997] have shown that causal Calderón–Zygmund operators remain bounded on the weighted space $L^p(w)$ for the larger class of *one-sided A_p weights*, defined by the finiteness of

$$[w]_{A_p^-} := \sup_{-\infty<a<b<c<\infty} \frac{1}{(c-a)^p} \Big(\int_b^c w \Big) \Big(\int_a^b w^{-\frac{1}{p-1}} \Big)^{p-1},$$

but the optimal bound for the operator norm $\|T\|_{\mathscr{L}(L^p(w))}$ in terms of $[w]_{A_p^-}$ remains open. In analogy with the A_2 Theorem 11.3.26, it is natural to make:

Conjecture 11.4.1 (One-sided A_2 conjecture of Chen, Han, and Lacey [2020]). For all causal Calderón–Zygmund operators,

$$\|T\|_{\mathscr{L}(L^p(w))} \leqslant c_T ([w]_{A_p^-})^{\max(1,\frac{1}{p-1})}.$$

Partial results for Haar multipliers (see Section 12.1.a) in place of Calderón–Zygmund operators are obtained by Chen et al. [2020], but beyond that the conjecture remains open.

Causal operators appear very naturally; e.g., the operator-valued kernel

$$K(s,t) = \mathbf{1}_{\mathbb{R}_+}(s-t) A e^{-(s-t)A},$$

of relevance to the maximal regularity problem studied in Chapter 17, has this form. A theory of one-sided singular integrals applicable to this operator-valued situation has been developed by Chill and Król [2018].

Matrix weighted spaces and convex body domination

Let $W : \mathbb{R}^d \to \mathbb{R}^{N\times N}$ be a *matrix weight*, i.e., measurable and positive definite almost everywhere, and $f : \mathbb{R}^d \to \mathbb{R}^N$ be measurable. The norm

$$\begin{aligned} \|f\|^2_{L^2(W)} &:= \int_{\mathbb{R}^d} \langle W(t) f(t), f(t)\rangle \,\mathrm{d}t \\ &= \int_{\mathbb{R}^d} |W(t)^{\frac12} f(t)|^2 \,\mathrm{d}t = \|W^{\frac12} f\|^2_{L^2(\mathbb{R}^d;\mathbb{R}^N)}, \end{aligned}$$

appears naturally from the prediction theory for multivariate stationary stochastic processes $n \in \mathbb{Z} \mapsto \xi_n \in L^2(\Omega; \mathbb{R}^N)$ developed by Wiener and Masani [1958], where stationarity means that $\Gamma_{n-k} := \mathbb{E}\xi_n \xi_k^T \in \mathbb{R}^{N \times N}$ depends only on the difference of the discrete times $n, k \in \mathbb{Z}$. If W is the density of the spectral measure of the process, i.e., $\Gamma_k = \widehat{W}(k)$ are the Fourier coefficients of $W \in L^1(\mathbb{T}; \mathbb{R}^{N \times N})$, the boundedness of the Hilbert transform on $L^2(W)$ is equivalent to a positive angle between the past and the future of the process. Even for $N = 1$, this problem was only solved 15 years later by Hunt, Muckenhoupt, and Wheeden [1973], who characterised this boundedness in terms of the A_2 condition. For $N > 1$, it took over 20 more years before the solution was obtained by Treil and Volberg [1997], who identified the correct analogue of the A_2 condition in the matrix-valued case:

$$[W]_{A_2} := \sup_Q |\langle W \rangle_Q^{1/2} \langle W^{-1} \rangle_Q^{1/2}|^2,$$

where $|\ |$ is (say) the operator norm on $\mathscr{L}(\mathbb{R}^N)$ (but the choice of the norm on $\mathbb{R}^{N \times N}$ is irrelevant, as they are all equivalent).

With the natural definition

$$\|f\|_{L^p(W)} := \|W^{\frac{1}{p}} f\|_{L^p(\mathbb{R}^d; \mathbb{R}^N)},$$

one is led to inquire about the boundedness of the Hilbert transform on $L^p(W)$. The characterising matrix-A_p condition, identified via different approaches by Nazarov and Treil [1996] and Volberg [1997], is less intuitive for $p \neq 2$. It is perhaps most easily formulated with the help of the classical theorem of John [1948], which guarantees that every norm on $\mathbb{R}^N$ is equivalent (with constants depending only on N) to a Euclidean norm, whose unit ball is a linear transformation of the standard unit ball. If W is a matrix weight and $V := W^{\frac{1}{p}}$, it is easy to see that

$$e \in \mathbb{R}^n \mapsto \Big(\fint_Q |V(t)e|^p \, \mathrm{d}t \Big)^{1/p}$$

is a norm, and hence, by the theorem of John [1948], there is a positive definite *reducing operator* $[V]_{Q,p} \in \mathbb{R}^{N \times N}$, such that

$$|[V]_{Q,p} e| \leqslant \Big(\fint_Q |V(t)e|^p \, \mathrm{d}t \Big)^{1/p} \leqslant \sqrt{N} \cdot |[V]_{Q,p} e|.$$

The matrix-A_p condition may then be defined by the finiteness of the constant

$$[W]_{A_p} := \sup_Q |[V]_{Q,p} [V^{-1}]_{Q,p'}|^p, \quad V := W^{\frac{1}{p}}.$$

The reader is invited to check that $[V]_{Q,p} = \langle V^p \rangle_Q^{\frac{1}{p}}$ if $N = 1$ or $p = 2$ (but not in general otherwise), so that the different definitions of A_p are consistent. It is possible to give an equivalent definition of the matrix A_p condition without reference to reducing operators, but one would still need them to prove anything interesting, which is why we prefer to state the definition as above.

While the qualitative boundedness of the Hilbert transform, and in fact of more general Calderón–Zygmund operators, on $L^p(W)$ was settled in the mentioned papers, the proof of the scalar-valued A_2 theorem raised the natural question of its extension to the matrix-weighted case. This remains open, but several related results have been achieved.

While sparse domination is perfectly applicable to vector-valued (even Banach space valued) functions, as we have seen in this chapter, it loses essential directional information, which makes it ill-suited for matrix-weighted considerations. To address this drawback, Nazarov, Petermichl, Treil, and Volberg [2017] invented a refined notion of *convex body domination*, where the averages $\langle \|f\| \rangle_Q$ are replaced by the related convex bodies

$$\Big\{ \langle \phi f \rangle_Q : \|\phi\|_{L^\infty(Q)} \leqslant 1 \Big\} \subseteq \mathbb{R}^N, \quad f \in L^1(Q;\mathbb{R}^N).$$

Convex body domination of T is most easily stated in its bilinear form, as an elaboration of the sparse form domination

$$\begin{aligned} |\langle Tf, g\rangle| &\leqslant c_{d,T} \sum_{Q\in\mathscr{S}} |Q| \langle |f| \rangle_{5Q} \langle |g| \rangle_{5Q} \\ &= c'_{d,T} \sum_{Q\in\mathscr{S}} \frac{1}{|Q|} \iint_{5Q\times 5Q} |f(s)||g(t)| \,\mathrm{d}s \,\mathrm{d}t. \end{aligned} \tag{11.57}$$

Convex body domination of T can now be stated in the form

$$|\langle Tf, g\rangle| \leqslant \sum_{Q\in\mathscr{S}} \frac{c_{d,N,T}}{|Q|} \sup_{\substack{\|\phi\|_\infty \leqslant 1 \\ \|\psi\|_\infty \leqslant 1}} \Big| \iint_{5Q\times 5Q} \phi(s)(s) \cdot \psi(t) g(t) \,\mathrm{d}s \,\mathrm{d}t \Big|, \tag{11.58}$$

with the important difference that we take the dot product of $f(s), g(t) \in \mathbb{R}^n$ first, and only then the absolute value of the result; this allows for critical directional cancellation compared to (11.57).

The proof of Nazarov, Petermichl, Treil, and Volberg [2017] (arXiv 1/2017), that standard Calderón–Zygmund operators satisfy (11.58), follows the same lines as the proof of Theorem 11.3.15 but with important elaborations at a few selected points, making again use of the ellipsoid theorem of John [1948]. On the other hand, with (11.58) available, Nazarov et al. [2017] can prove the bound

$$\|T\|_{\mathscr{L}(L^2(W))} \leqslant c_{d,T} [W]_{A_2}^{3/2},$$

which remains the best available matrix-weighted estimate for Calderón–Zygmund operators (or even just for the Hilbert transform) at the time of writing. A variant of the same results was also obtained by Culiuc, Di Plinio, and Ou [2018b], seemingly earlier (arXiv 10/2016) but not independently; according to their acknowledgment, the concept of domination by convex body averages was introduced to these authors by Sergei Treil during his seminar talk at Brown University in the Spring of 2016.

Since then, further applications and extensions of convex body domination have been explored by Cruz-Uribe, Isralowitz, and Moen [2018], Di Plinio, Hytönen, and Li [2020a], Isralowitz, Pott, and Rivera-Ríos [2021], Isralowitz, Pott, and Treil [2022], and Muller and Rivera-Ríos [2022]. Importantly, Bownik and Cruz-Uribe [2022] extended the Rubio de Francia algorithm (Proposition J.2.2), and its key application to weighted extrapolation (Theorem J.2.1), to matrix-valued weights, by further development of the convex body philosophy.

An abstract framework for convex body domination has been proposed by Hytönen [2023], allowing also Banach space valued functions in the theory. While genuinely operator-valued weights in infinite dimensions seem to be out of reach, this framework allows the treatment of $\mathbb{R}^{N\times N}$-valued weights on spaces of X^N-valued functions. In particular, the following simultaneous extensions of the boundedness of the Hilbert transform on the Banach space valued $L^2(\mathbb{R};X)$ by Burkholder [1983], and on the matrix-weighted $L^2(W)$ by Treil and Volberg [1997], is obtained there.

Theorem 11.4.2. *Let X be a UMD space, and $W:\mathbb{R}^d\to\mathbb{R}^{N\times N}$ be a matrix A_2 weight. Then the Hilbert transform H extends boundedly to*

$$L^2(W;X^N) := \Big\{ f:\mathbb{R}\to X^N : \|f\|_{L^2(W;X^N)} := \|W^{\frac{1}{2}} f\|_{L^2(\mathbb{R};X^N)} < \infty \Big\}$$

and satisfies $\|H\|_{\mathscr{L}(L^2(W;X^N))} \leqslant c_N \hbar_{2,X}[W]_{A_2}^{3/2} \leqslant c_N \beta_{2,X}^2 [W]_{A_2}^{3/2}$, where $\hbar_{2,X} = \|H\|_{L^2(\mathbb{R};X)}$ and $\beta_{2,X}$ is the UMD constant.

The stated quantitative formulation in terms of $\hbar_{2,X}$ is not explicit in Hytönen [2023], but can be tracked in the proof, in a similar way as in Corollary 11.3.27 in the text.

A summary of sharp weighted bounds for classical operators

Our discussion above has been focused on norms of Calderón–Zygmund singular integrals and their various extensions, viewed as operators on a weighted $L^p(w)$ (or matrix-weighted $L^p(W)$) space; these are referred to as strong-type bounds. We will briefly summarise results in two closely related directions. First, one may inquire about the corresponding weak-type bounds, i.e., operator norms in $\mathscr{L}(L^p(w), L^{p,\infty}(w))$. These are obviously dominated by the strong-type norms, but the point is that the optimal weak-type norms may be significantly smaller in some cases, which gives these questions an independent interest. Second, one may pose the same questions for various square-functions, which could be viewed as part of the extended family of (vector-valued, when suitably interpreted) Calderón–Zygmund operators; however, it turns out that these operators are actually slightly "better" in terms of the dependence of their norms on the weight constant. A basic example is the *dyadic square function*

$$Sf(x) := \Big(\sum_{Q \in \mathscr{D}} |\mathbb{D}_Q f(x)|^2 \Big)^{1/2},$$

(where the operators $\mathbb{D}_Q$ are defined in (12.1) and discussed extensively in Chapter 12), but several other classical square functions satisfy the same weighted bounds; we refer the reader to the papers quoted below for details.

A summary of the sharp bounds known for these operators is as follows:

Singular integrals:

For $p \in (1, \infty)$ and $w \in A_p$, the sharp estimates in $L^p(w)$ are:

(1) the strong-type bound is $[w]_{A_p}^{\max(1, \frac{1}{p-1})}$ (Hytönen [2012]);
(2) the weak-type bound is $[w]_{A_p}$ (Hytönen, Lacey, Martikainen, Orponen, Reguera, Sawyer, and Uriarte-Tuero [2012]);
(3) the weak-type $L^1(w)$ bound is $[w]_{A_1}(1 + \log[w]_{A_1})$ (the upper bound was proved by Lerner, Ombrosi, and Pérez [2009], its sharpness is due to Lerner, Nazarov, and Ombrosi [2020]).

A speculative linear-in-$[w]_{A_1}$ bound in (3) was known as the A_1 conjecture, or the weak Muckenhoupt–Wheeden conjecture. The original conjecture, disproved by Reguera [2011] and Reguera and Thiele [2012], was about the boundedness of $T : L^1(Mw) \to L^{1,\infty}(w)$ for any weight w. This holds for M in place of T (Theorem 3.2.27), which motivated the conjecture.

Square functions:

For the range of p as specified and $w \in A_p$, the sharp estimates in $L^p(w)$ are:

(4) the strong-type bound is $[w]_{A_p}^{\max(\frac{1}{2}, \frac{1}{p-1})}$ for $p \in (1, \infty)$ (Lerner [2011]);
(5) the weak-type bound is $[w]_{A_p}^{\max(\frac{1}{2}, \frac{1}{p})}$ for $p \in [1, \infty) \setminus \{2\}$ ($p = 1$: Chanillo and Wheeden [1987], Wilson [2007, 2008]; $p \in (1, 2)$: Lacey and Scurry [2012]; $p > 2$: Hytönen and Li [2018]);
(6) the weak-type $L^2(w)$ bound is at most $[w]_{A_2}^{\frac{1}{2}}(1 + \log[w]_{A_1})^{\frac{1}{2}}$ (Domingo-Salazar, Lacey, and Rey [2016]), but its sharpness seems to remain open (see Ivanisvili and Volberg [2018] for partial related results).

In contrast to singular integrals, the bounds at $p = 1$ above are consequences of the stronger statement that $S : L^1(Mw) \to L^{1,\infty}(w)$ is bounded for any weight w, i.e., the Muckenhoupt–Wheeden conjecture holds for square functions. This also explains the (implicit) appearance of sharp weighted bounds in Chanillo and Wheeden [1987], long before this became a fashionable topic.

For matrix-weights, the only known sharp estimates among these examples, at the time of writing, seem to be the square function bounds (4) for $p \in (1,2]$; this was proved by Hytönen, Petermichl, and Volberg [2019b] for $p = 2$ and extended by Isralowitz [2020] to $p \in (1,2)$.