J. Math. Pures Appl. 156 (2021) 351–391

Contents lists available at ScienceDirect

Journal de Mathématiques Pures et Appliquées

www.elsevier.com/locate/matpur

The L^p-to-L^q boundedness of commutators with applications to the Jacobian operator

Tuomas P. Hytönen [1]

Department of Mathematics and Statistics, P.O. Box 68 (Pietari Kalmin katu 5), FI-00014 University of Helsinki, Finland

A R T I C L E   I N F O

Article history:
Received 11 February 2021
Available online 12 October 2021

MSC:
42B20
42B25
42B37
35F20
47B47

Keywords:
Commutator
Singular integral
Jacobian determinant

A B S T R A C T

Supplying the missing necessary conditions, we complete the characterisation of the $L^p \to L^q$ boundedness of commutators $[b, T]$ of pointwise multiplication and Calderón–Zygmund operators, for arbitrary pairs of $1 < p, q < \infty$ and under minimal non-degeneracy hypotheses on T.

For $p \leq q$ (and especially $p = q$), this extends a long line of results under more restrictive assumptions on T. In particular, we answer a recent question of Lerner, Ombrosi, and Rivera-Ríos by showing that $b \in \mathrm{BMO}$ is necessary for the L^p-boundedness of $[b, T]$ for any non-zero homogeneous singular integral T. We also deal with iterated commutators and weighted spaces.

For $p > q$, our results are new even for special classical operators with smooth kernels. As an application, we show that every $f \in L^p(\mathbb{R}^d)$ can be represented as a convergent series of normalised Jacobians $Ju = \det \nabla u$ of $u \in \dot{W}^{1,dp}(\mathbb{R}^d)^d$. This extends, from $p = 1$ to $p > 1$, a result of Coifman, Lions, Meyer and Semmes about $J : \dot{W}^{1,d}(\mathbb{R}^d)^d \to H^1(\mathbb{R}^d)$, and supports a conjecture of Iwaniec about the solvability of the equation $Ju = f \in L^p(\mathbb{R}^d)$.

R É S U M É

En fournissant les conditions nécessaires manquantes, nous complétons la caractérisation de la bornitude $L^p \to L^q$ des commutateurs $[b, T]$ de multiplication ponctuelle et des opérateurs de Calderón–Zygmund, pour des paires arbitraires de $1 < p, q < \infty$ et avec des hypothèses de non-dégénérescence minimales sur T.

Pour $p \leq q$ (et en particulier $p = q$), cela étend une longue ligne de résultats avec des hypothèses plus restrictives sur T. En particulier, nous répondons à une question récente de Lerner, Ombrosi et Rivera-Ríos en montrant que $b \in \mathrm{BMO}$ est nécessaire pour la L^p-bornitude de $[b, T]$ pour toute intégrale singulière homogène non nulle T. Nous traitons également des commutateurs itérés et des espaces avec des poids.

Pour $p > q$, nos résultats sont nouveaux même pour les opérateurs classiques spéciaux à noyaux lisses. Comme application, nous montrons que chaque $f \in L^p(\mathbb{R}^d)$ peut être représenté comme une série convergente de Jacobiens normalisés $Ju = \det \nabla u$ de $u \in \dot{W}^{1,dp}(\mathbb{R}^d)^d$. Ceci étend, de $p = 1$ à $p > 1$, un résultat de Coifman,

E-mail address: tuomas.hytonen@helsinki.fi.

[1] The author was supported by the Academy of Finland via project Nos. 314829 (Frontiers of singular integrals) and 307333 (Centre of Excellence in Analysis and Dynamics Research).

https://doi.org/10.1016/j.matpur.2021.10.007

Lions, Meyer et Semmes concernant $J : \dot{W}^{1,d}(\mathbb{R}^d)^d \to H^1(\mathbb{R}^d)$, et supporte une conjecture d'Iwaniec sur la résolvabilité de l'équation $Ju = f \in L^p(\mathbb{R}^d)$.

1. Introduction

The first goal of this paper is to complete the following picture of the $L^p(\mathbb{R}^d)$-to-$L^q(\mathbb{R}^d)$ boundedness properties of commutators of pointwise multiplication and singular integral operators:

Theorem 1.0.1. *Let $1 < p, q < \infty$, let T be a "non-degenerate" Calderón–Zygmund operator on $\mathbb{R}^d$, and let $b \in L^1_{\text{loc}}(\mathbb{R}^d)$. Then the commutator*

$$[b, T] : f \mapsto bTf - T(bf)$$

defines a bounded operator $[b, T] : L^p(\mathbb{R}^d) \to L^q(\mathbb{R}^d)$ if and only if:

- *$p = q$ and b has bounded mean oscillation, or*
- *$p < q \le p^* = \dfrac{pd}{(d-p)_+}$ and b is α-Hölder continuous for $\alpha = \Big(\dfrac{1}{p} - \dfrac{1}{q}\Big)d$, or*
- *$q > p^*$ and b is constant, or*
- *$p > q$ and $b = a + c$, where $a \in L^r(\mathbb{R}^d)$ for $\dfrac{1}{r} = \dfrac{1}{q} - \dfrac{1}{p}$, and c is constant.*

To be explicit, the definition of the Sobolev exponent p^* above is $pd/(d-p)$, if $p < d$, and ∞ otherwise; thus $p < q \le p^*$ is precisely the condition that the Hölder exponent satisfies $\alpha \in (0, 1]$. We say that a Calderón–Zygmund operator $Tf(x) = \int K(x, y)f(y)\,dy$, with usual (or weaker) assumptions on the kernel K recalled in Section 2.1, is "non-degenerate" provided that, for some $c_0 > 0$,

$$\text{for every } y \in \mathbb{R}^d \text{ and } r > 0, \text{ there is } x \in B(y, r)^c \text{ with } |K(x, y)| \ge \frac{1}{c_0 r^d}; \tag{1.1}$$

i.e., uniformly over all positions and length-scales, the kernel takes some values that are as big as they are allowed to be by the standard upper bound for $K(x, y)$. When $K(x, y) = \dfrac{\Omega(x-y)}{|x-y|^d}$ is a (possibly rough) homogeneous kernel, this requirement simply says that Ω is not identically zero.

1.1. Sufficient conditions for boundedness

We note that all the "if" parts of Theorem 1.0.1 are either well known or easy. The cases when b is constant are completely trivial, since in this case the commutator vanishes. If $b \in L^r(\mathbb{R}^d)$ with $\frac{1}{r} = \frac{1}{q} - \frac{1}{p}$, the boundedness is also immediate simply from the boundedness of T on both $L^p(\mathbb{R}^d)$ and $L^q(\mathbb{R}^d)$ (taking this as part of the definition of a "Calderón–Zygmund operator"), together with Hölder's inequality:

$$\begin{aligned}\|[b, T]f\|_q &\le \|bTf\|_q + \|T(bf)\|_q \le \|b\|_r\|Tf\|_p + \|T\|_{L^q\to L^q}\|bf\|_q \\ &\le \|b\|_r\big(\|T\|_{L^p\to L^p} + \|T\|_{L^q\to L^q}\big)\|f\|_p.\end{aligned}$$

In particular, no mutual cancellation between the two terms of the commutator is involved in this estimate. This computation is also valid when $p = q$ and $r = \infty$, showing the trivial sufficiency of $b \in L^\infty(\mathbb{R}^d)$ for the boundedness of $[b, T]$ on $L^p(\mathbb{R}^d)$. The fact that the larger space $\mathrm{BMO}(\mathbb{R}^d)$ is still admissible for this boundedness is a celebrated theorem of Coifman, Rochberg and Weiss [7] and the only truly nontrivial result among the "if" statements of Theorem 1.0.1.

If b is α-Hölder continuous, using only the standard pointwise bound for Calderón–Zygmund kernels, we see that

$$|[b,T]f(x)| = \Big|\int_{\mathbb{R}^d} (b(x)-b(y))K(x,y)f(y)\,\mathrm{d}y\Big| \lesssim \int_{\mathbb{R}^d} |x-y|^{\alpha}\frac{1}{|x-y|^d}|f(y)|\,\mathrm{d}y$$

is pointwise dominated by the usual fractional integral operator, whose $L^p(\mathbb{R}^d)$-to-$L^q(\mathbb{R}^d)$ bounds are classical and well known.

1.2. Necessary conditions for boundedness

Let us then discuss the "only if" parts of Theorem 1.0.1. For $p = q$, already Coifman, Rochberg and Weiss [7] proved the necessity of $b \in \mathrm{BMO}(\mathbb{R}^d)$ for the $L^p(\mathbb{R}^d)$-boundedness of $[b,T]$ for *all* d Riesz transforms R_j, $j = 1,\dots,d$. (This reduces to just the Hilbert transform when $d = 1$.) Their argument made explicit use of the special algebraic form of the relevant kernels.

Janson [20] and Uchiyama [35], independently, extended the necessity part of the Coifman–Rochberg–Weiss theorem to more general classes of homogeneous Calderón–Zygmund kernels with "sufficient" smoothness. In particular, their results contain the fact that the boundedness of $[b,R_j]$ for just *one* (instead of *all*) $j = 1,\dots,d$ already implies that $b \in \mathrm{BMO}(\mathbb{R}^d)$. Janson's argument may be viewed as an analytic extension of that of Coifman et al., in that he used the smoothness to guarantee absolute convergence of the Fourier expansion of the inverse $1/K$ of the kernel, where the individual frequency components could then be treated by the algebraic method. Janson also proves the "only if" part of Theorem 1.0.1 for $p < q$ (and in fact for more general Orlicz norms) for the same class of smooth homogeneous kernels. Uchiyama's argument is different, but still dependent on both smoothness and homogeneity of the kernel.

A recent advance was made by Lerner, Ombrosi and Rivera-Ríos [26], who identified sufficient *local positivity* (lack of sign change in a nonempty open set) as a workable replacement of the previous smoothness assumptions on the (still homogeneous) kernel to deduce the necessity of $b \in \mathrm{BMO}(\mathbb{R}^d)$ for the $L^p(\mathbb{R}^d)$-boundedness of $[b,T]$. Similar results in the case of not necessarily homogeneous Calderón–Zygmund kernels were subsequently obtained by Guo, Lian and Wu [13]; see also Duong, Li, Li and Wick [10] for the concrete case when T is a Riesz transform related to the sub-Laplacian on a stratified nilpotent Lie group.

In the present work, we take the final step in generalising the class of admissible kernels, showing that any non-degenerate Calderón–Zygmund kernel is admissible for the "only if" conclusions of Theorem 1.0.1. In particular, our result applies to both two-variable kernels $K(x,y)$ (with very little smoothness) and rough homogeneous kernels $\dfrac{\Omega(x-y)}{|x-y|^d}$, under a minimal non-degeneracy assumption. In the case of homogeneous kernels we merely need that $\Omega \in L^1(S^{d-1})$ does not vanish identically. This answers positively a question raised by Lerner et al. [26, Remark 4.1]; as discussed below, we also address the more general two-weight bounds and higher commutators as considered in [26]. Also in the case of two-variable kernels, our non-degeneracy hypothesis seems to be at least as general as anything found in the literature; in contrast to [13] in particular, we allow in (1.1) that the point of non-degeneracy x may lie in any direction from the reference point y.

1.3. The case $p > q$ and applications to the Jacobian operator

The case $p > q$ of Theorem 1.0.1 is completely new even for special Calderón–Zygmund operators like the Riesz transforms, for which the complementary range $p \le q$ was understood for a long time. The result in this new range is perhaps surprising, in that it says that there is essentially no cancellation between bT and Tb in this regime. (An initial working hypothesis before discovering this result was that the role of BMO in the commutator boundedness in this regime of exponents could be taken by another space JN_r, which was implicitly introduced by John and Nirenberg [21, §3] and recently studied in [9]. However, the obtained result disproves this hypothesis.)

Technically, this is the hardest case of the proof, which is somewhat explained by the fact that membership in $L^r(\mathbb{R}^d)$ is a "global" condition, in contrast to the "uniform local" conditions defining both $\mathrm{BMO}(\mathbb{R}^d)$ and α-Hölder continuous functions. Incidentally, a similar dichotomy between "global" conditions characterising L^p-to-L^q (or similar) boundedness for $p > q$, and "uniform local" conditions in the case $p \le q$, has also been recently discovered in a couple of other settings as well:

1. In the context of two-weight norm inequalities for certain discrete positive operators, the characterisation for $p \le q$ by Lacey, Sawyer and Uriarte-Tuero [24] is in terms of local "testing conditions" uniform over all dyadic cubes, while the characterisation for $p > q$ due to Tanaka [34] involves the L^r membership of a "discrete Wolff potential"; see also [14] for a unified approach to both cases.
2. The boundedness of certain Toeplitz type operators between the holomorphic Hardy spaces H^p and H^q of the unit ball was characterised by Pau and Perälä [31] in both regimes of the exponents, in terms of a uniform local Carleson measure condition for $p \le q$, and in terms of the global L^r membership of a certain auxiliary function for $p > q$. These results, analogous to our present ones but in a different context, were found independently at almost the same time: the first arXiv versions of [31] and the present work came out within two weeks of each other.

It might be of interest for general operator theory in L^p spaces to find further examples of, and/or a broader context for, this phenomenon.

A part of the motivation to study this regime of exponents for commutator inequalities came from a recent observation of Lindberg [28] about the connections of such bounds, in the particular case when T is the Ahlfors–Beurling transform, to the Jacobian equation

$$Ju := \det \nabla u := \det(\partial_i u_j)_{i,j=1}^d = f \in L^p(\mathbb{R}^d).$$

It has been conjectured by Iwaniec [19] that, for $p \in (1, \infty)$, the (obviously bounded) map $J : \dot{W}^{1,pd}(\mathbb{R}^d)^d \to L^p(\mathbb{R}^d)$, where $\dot{W}^{1,pd}$ is the homogeneous Sobolev space, has a continuous right inverse and in particular is surjective. As a variant of our estimates for commutators, we will provide partial positive evidence by showing that the closed linear span of the range of J is all of $L^p(\mathbb{R}^d)$. This is an L^p-analogue of a result of Coifman, Lions, Meyer and Semmes [6, p. 258] who obtained a similar conclusion for $J : \dot{W}^{1,d}(\mathbb{R}^d)^d \to H^1(\mathbb{R}^d)$, which corresponds to the case $p = 1$, with the usual replacement of L^1 by the Hardy space H^1.

Recently, Lindberg [28, p. 739] proposed an approach to the planar ($d = 2$) case of the Jacobian operator via the complex-variable framework

$$Ju = |\partial h|^2 - |\bar{\partial} h|^2 = |S(\bar{\partial} h)|^2 - |\bar{\partial} h|^2,$$

where $h = u_1 + iu_2$, $\partial = \frac{1}{2}(\partial_1 - i\partial_2)$, $\bar{\partial} = \frac{1}{2}(\partial_1 + i\partial_2)$, and S is the Ahlfors–Beurling operator. This led him to a question about the boundedness of the commutator $[b, S] : L^{2p} \to L^{(2p)'}$, which is solved as a particular case of Theorem 1.0.1; observe that $2p > 2 > (2p)'$ here. Following Lindberg's outline [28, p. 739], conclusions about the planar Jacobian could then be obtained as corollaries to Theorem 1.0.1; but it turns out that a combination of some elements of its proof, together with the techniques of Coifman, Lions, Meyer and Semmes [6], actually allows to prove such results in any dimension; see Section 3.

1.4. A priori assumptions on b, T and $[b,T]$

In general it takes some effort to define precisely what is meant by "Tf", when T is a singular integral operator, or by saying that such an operator "is bounded" from one space to another. In our approach to the "only if" statements of Theorem 1.0.1, we avoid all this subtlety; in fact, our assumptions may be formulated entirely in terms of the kernel K without ever having to define the operator T or $[b,T]$, although we still use these symbols as convenient abbreviations. All we need is estimates for the bilinear form

$$\langle [b,T]f,g\rangle = \iint \big(b(x)-b(y)\big)K(x,y)f(y)g(x)\,\mathrm{d}y\,\mathrm{d}x, \tag{1.2}$$

where the functions $f,g\in L^\infty(\mathbb{R}^d)$ have bounded supports separated by a positive distance; we refer to such estimates as *off-support* bounds for $[b,T]$. Under the standard estimates for a Calderón–Zygmund kernel, the above integral exists as an absolutely convergent Lebesgue integral when $b\in L^1_{\mathrm{loc}}(\mathbb{R}^d)$, as in Theorem 1.0.1.

For $p\leq q$, we only need the bound

$$\begin{gathered} |\langle [b,T]f,g\rangle| \leq C\|f\|_\infty\|g\|_\infty|B|^{1/p}|\tilde{B}|^{1/p'}, \quad \text{whenever} \\ \operatorname{spt} f\subset B, \quad \operatorname{spt} g\subset \tilde{B}, \quad r_B=r_{\tilde{B}}\leq c\operatorname{dist}(B,\tilde{B}), \end{gathered} \tag{1.3}$$

where B and $\tilde{B}$ denote arbitrary balls of radius r_B and $r_{\tilde{B}}$, respectively. This is weaker than a *restricted weak type* (p,q) *estimate* in two ways: the bound involves the bigger quantities $|B|$ and $|\tilde{B}|$ in place of $|\operatorname{spt} f|$ and $|\operatorname{spt} g|$ on the right, and it is only required to hold under the quantitative off-support condition above. (A certain technical strengthening, but still formally weaker than the global boundedness of $[b,T]:L^p\to L^q$, and involving off-support bounds only, is needed when $p>q$.)

We note that Liaw and Treil [27] have provided a framework to interpret the boundedness of a singular integral operator (an issue that we have chosen to avoid) via off-support conditions of a similar flavour. However, the off-support conditions that we impose on f and g are significantly stronger (and hence the resulting estimate on the operator restricted to such pairs of functions much weaker) than those of [27]; in particular, the quantitative separation of supports in (1.3) efficiently prevents approximating a form with arbitrary f,g (as done in [27]) by the off-support forms above.

The fact that one only needs off-support estimates in the "only if" directions of Theorem 1.0.1 is already implicit in the argument of Uchiyama [35, proof of Theorem 1], but not in all recent works, and it seems not to have been explicitly stated in the literature. On the other hand, Lerner et al. [26] use a *restricted strong type* assumption, while Guo et al. [13] state one of their results under a weak type hypothesis. Our condition (1.3) simultaneously relaxes both these assumptions.

Note that the a priori assumption that $b\in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ is essentially the weakest possible to make sense of the commutator $[b,T]$, even in the off-support sense as above. While many earlier results related to Theorem 1.0.1 are obtained under this same minimal assumption, some others assume $b\in \mathrm{BMO}(\mathbb{R}^d)$ *qualitatively* to begin with, and then prove the *quantitative* bound $\|b\|_{\mathrm{BMO}}\lesssim \|[b,T]\|_{L^p\to L^p}$; see e.g. [10, Theorem 1.2]. A simplification brought by this stronger a priori assumption is that one can absorb error terms of the form $\varepsilon\|b\|_{\mathrm{BMO}}$ in the argument. We will also use absorption, but only to quantities whose finiteness is guaranteed by $b\in L^1_{\mathrm{loc}}(\mathbb{R}^d)$.

1.5. Methods and scope

We will prove versions of Theorem 1.0.1 by two methods of somewhat different scopes. The first method is based on the well-known connection of commutator estimates to *weak factorisation*, which has been widely used since the pioneering work [7]. (In contrast to proper factorisation, where an object is expressed as a product of other objects, weak factorisation refers to decompositions in terms of sums, or possibly infinite series, of products.) This depends on the basic identity

$$\langle [b,T]f,g\rangle = \langle b, gTf - fT^*g\rangle,$$

where each term is well-defined as a Lebesgue integral for disjointly supported f and g. Hence, if an arbitrary h in (a dense subspace of) a predual of the space hoped to contain b can be expanded as

$$h = \sum_i (g_i T f_i - f_i T^* g_i), \tag{1.4}$$

then we can hope to estimate

$$|\langle b,h\rangle| \leq \sum_i |\langle b, g_i T f_i - f_i T^* g_i\rangle| = \sum_i |\langle [b,T]f_i, g_i\rangle|$$

in order to bound $\|b\|$ in terms of $\|[b,T]\|$. An inherent difficulty is that, even with good convergence properties of the expansion (1.4) in the predual space, lacking the a priori knowledge that b should be in the relevant space, it may be difficult to justify the "$\leq$" above. We circumvent this problem by replacing (1.4) by an *approximate weak factorisation*, where the sum over i is finite, but there is an additional error term $\tilde{h}$ that will be eventually absorbed.

This method is strong enough for proving Theorem 1.0.1 as stated, where both the function b and the kernel $K(x,y)$ of T are allowed to be *complex-valued*. Besides completeness of the theory, achieving this level of generality was initially motivated by the applications to the Jacobian operator via the Ahlfors–Beurling transform, as discussed above. The kernel of this operator, $K(z,w) = -\pi^{-1}/(z-w)^2$ for $z,w\in\mathbb{C}$, is genuinely complex-valued, and it is only natural to view it as acting on (and forming commutators with) complex-valued functions. While this is hardly exotic, it should be stressed that some of the recent contributions, like our second method, are inherently restricted to real-valued b.

Our second approach could be called the *median method*, and it is a close cousin of the recent work [26]. It makes explicit use of the order structure of the real line as the range of the function b. The advantage of this method is that, with little additional effort, it can also handle the higher order commutators

$$T_b^k = [b, T_b^{k-1}], \qquad T_b^1 = [b,T].$$

As before, we only need the off-support bilinear form

$$\langle T_b^k f, g\rangle = \iint (b(x)-b(y))^k K(x,y) f(y) g(x)\,\mathrm{d}u\,\mathrm{d}x$$

of these operators for $f,g\in L^\infty(\mathbb{R}^d)$ with bounded supports separated by a positive distance, and $b\in L^k_{\mathrm{loc}}(\mathbb{R}^d)$ is a sufficient a priori assumption to make sense of this. We also apply this method to two-weight commutator inequalities in Section 4.3.

1.6. Extensions to other settings

While the present work concentrates on commutators $[b,T]$ (and their iterates of the form $[b,[b,T]]$ etc.) of pointwise multipliers and linear Calderón–Zygmund operators between L^p spaces, we record a number of extensions of either the operators or the spaces under consideration:

1. Bilinear Calderón–Zygmund operators map $T: L^{p_1}\times L^{p_2}\to L^p$ with $\frac{1}{p}=\frac{1}{p_1}+\frac{1}{p_2}$, and one can ask about conditions for

$$[b, T]_1 : (f, g) \mapsto bT(f, g) - T(bf, g) : L^{p_1} \times L^{p_2} \to L^q.$$

The necessity of $b \in \mathrm{BMO}$ when $p = q$ was first obtained by Chaffee [4] under Janson-type assumptions and methods involving the Fourier expansion of the inverse kernel $1/K$. Since the circulation of our results, Oikari [29] has extended the present hypotheses and methods to bilinear operators, obtaining a close analogue of Theorem 1.0.1 in this setting.

2. Iterated commutators of the form $[[b, T_1], T_2]$, where b is a function of two variable x_1, x_2, and each T_i is a Calderón–Zygmund operator acting in the variable x_i, play an important role in the theory of singular integrals and function spaces on product domains. A characterisation of the $L^p \to L^p$ boundedness of these bi-commutators remains open, after the recent discovery of a gap [23] in the celebrated Ferguson–Lacey theorem [11] and its extensions; nevertheless, various mixed-norm $L^{p_1}(L^{p_2}) \to L^{q_1}(L^{q_2})$ bounds with $(p_1, p_2) \neq (q_1, q_2)$ have been recently characterised by Airta et al. [2], by extending the methods of the present paper.
3. The necessity of $b \in \mathrm{BMO}(\mathbb{R}^d)$ for the boundedness of commutators of both linear and bilinear singular integrals between more general Banach functions spaces (in place of L^p spaces) has been obtained by Chaffee and Cruz-Uribe [5], again under Janson-type assumptions and approach. It might be of interest to revisit their results with our (more general) conditions and methods. Among other examples, Chaffee and Cruz-Uribe also consider weighted bounds, in which case our Theorem 4.3.2 is a significant generalisation of their [5, Corollary 2.3]. It seems plausible that similar extensions would be available for other results of [5] as well.

1.7. About notation

We will make extensive use of the notation "$\lesssim$" to indicate an inequality up to an unspecified multiplicative constant. Such constants are always allowed to depend on the underlying dimension d, any of the Lebesgue space exponents $p, q, r, \ldots$, and also on the Calderón–Zygmund operator T and its kernel K, as well as on the order k of an iterated commutator; these are regarded as fixed throughout the argument. The implied constants may never depend on any of the functions under consideration (neither on the function b appearing in the commutator $[b, T]$ itself nor on any of the functions $f, g, \ldots$ on which the commutator acts), nor points or subsets (balls, cubes, etc.) of their domain $\mathbb{R}^d$. Many arguments involve an auxiliary (large) parameter A, and dependence on it is also indicated explicitly until a suitable value of A (depending only on the admissible quantities) is fixed once and for all for the rest of the argument.

The subscript zero of a Lebesgue space indicates vanishing integral, i.e., $L^p_0(Q) = \{f \in L^p(Q) : \int f = 0\}$. The subscript zero of a Sobolev space $W^{1,p}_0(\Omega)$ (which will be only mentioned in passing) indicates vanishing boundary values in the Sobolev sense. Compact support is indicated by the subscript c, mainly in the context of the test function space $C^\infty_c(\mathbb{R}^d)$. We denote by $\fint_E f := |E|^{-1} \int_E f$ the average of a function over a set E of finite positive measure.

2. Complex commutators and approximate weak factorisation

In this section we prove the "only if" claims of Theorem 1.0.1.

2.1. Non-degenerate Calderón–Zygmund kernels

We begin by describing the precise class of singular integral kernels that we study. We consider two-variable Calderón–Zygmund kernels under the standard conditions

$$K(x, y) \leq \frac{c_K}{|x - y|^d} \qquad \forall x \neq y,$$

$$|K(x,y)-K(x',y)|+|K(y,x)-K(y,x')| \leq \frac{1}{|x-y|^d}\omega\Big(\frac{|x-x'|}{|x-y|}\Big),$$

whenever $|x-x'| < \frac{1}{2}|x-y|$, where the modulus of continuity $\omega : [0,1) \to [0,\infty)$ is increasing. We refer to such a kernel as an *ω-Calderón–Zygmund kernel*. A common assumption is that $\omega(t) = c_\alpha t^\alpha$ for some $\alpha \in (0,1]$, or a more general Dini-condition $\int_0^1 \omega(t)\frac{\mathrm{d}t}{t} < \infty$, but we need even significantly less, namely that $\omega(t) \to 0$ as $t \to 0$.

We also consider *rough homogeneous kernels*

$$K(x,y) = K(x-y) = \frac{\Omega(x-y)}{|x-y|^d},$$

where $\Omega \in L^1(S^{d-1})$ and $\Omega(tx) = \Omega(x)$ for all $t > 0$ and $x \in \mathbb{R}^d$. We note that the off-support bilinear form (1.2) is also well defined (absolutely integrable) for this type of kernels: the integrals of $y \mapsto |K(x-y)f(y)|$ are uniformly bounded over $x \in \operatorname{spt} g$, and $x \mapsto |b(x)g(x)|$ is integrable; the term involving $b(y)$ can be estimated similarly by carrying the iterated integrals in a different order.

In either case, the $L^p(\mathbb{R}^d)$-boundedness of an integral operator T associated with K neither follows from these assumptions, nor is assumed as a separate condition, as this is not needed. The story is different for the "if" directions of Theorem 1.0.1, but our present goal is to prove the "only if" directions with minimal assumptions.

Definition 2.1.1. We say that K is a non-degenerate Calderón–Zygmund kernel, if (at least) one of the following two conditions holds:

1. K is an ω-Calderón–Zygmund kernel with $\omega(t) \to 0$ as $t \to 0$ and for every $y \in \mathbb{R}^d$ and $r > 0$, there exists $x \in B(y,r)^c$ with
$$|K(x,y)| \geq \frac{1}{c_0 r^d}.$$
2. K is a homogeneous Calderón–Zygmund kernel with $\Omega \in L^1(S^{d-1}) \setminus \{0\}$. In particular, there exists a Lebesgue point $\theta_0 \in S^{d-1}$ of Ω such that
$$\Omega(\theta_0) \neq 0.$$

Remark 2.1.2 *(Comparison with non-degenerate kernels in the sense of Stein).* Suppose that K is an ω-Calderón–Zygmund kernel of the convolution form $K(x,y) = K(x-y)$. Then the non-degeneracy condition (1) of Definition 2.1.1 simplifies into the following form: for every $r > 0$, we have

$$|K(x)| \geq \frac{1}{c_0 r^d} \quad \text{for some } x \in B(0,r)^c. \tag{2.1}$$

For convolution kernels, there is also the following well-known non-degeneracy condition introduced by Stein [33, IV.4.6]: there exists a constant $a > 0$, and a unit vector u_0, so that

$$|K(t\cdot u_0)| \geq a\cdot |t|^{-d}, \quad \text{for all } t \in \mathbb{R}\setminus\{0\}. \tag{2.2}$$

It is immediate that Stein's non-degeneracy implies our version. In fact, assume (2.2) and fix some $c_1 \geq 1$. Given $r > 0$, we find that any $x = t\cdot u_0$, where $|t| \in [r, c_1 r]$, satisfies (2.1) with $c_0 = c_1^d/a$. Thus, while (2.1) requires just the existence of one x, Stein's condition provides two symmetric line segments of admissible x that, moreover, have simple explicit dependence on r and are always located on the same fixed ray through the origin. It is not surprising that (2.1) is easily satisfied even when (2.2) is not, and we provide some examples below.

Note that it is assumed in the discussion of non-degeneracy in [33, IV.4.6], but not in Definition 2.1.1, that K should be the kernel of a bounded operator on $L^2(\mathbb{R}^d)$, and this would offer a source of cheap examples in terms of kernels of unbounded operators. To make clear that this is not a decisive difference between the two conditions, we take the slight additional trouble of making our examples correspond to bounded operators on $L^2(\mathbb{R}^d)$.

Several of our examples to follow will exploit a standard resolution of unity

$$\sum_{j\in\mathbb{Z}} \varphi(2^{-j}x) \equiv 1 \quad \forall x>0, \quad \varphi \in C_c^\infty(\frac{1}{2},2), \tag{2.3}$$

where we note in particular that $\varphi(1)=1$ under these conditions.

Example 2.1.3 *(Stein's non-degeneracy violated at one or two points).* When $d=1$, Stein's condition (2.2) simply says that $|K(x)| \geq a|x|^{-1}$, so any K that vanishes even at one point of $\mathbb{R}\setminus\{0\}$ is not admissible. Let us fix some K_0 that does satisfy (2.2), say the Hilbert kernel $K_0(x)=1/x$. We then define

$$K(x) = K_0(x) - K_0(1)\varphi(x).$$

It is immediate that this perturbation of K_0 neither destroys the Calderón–Zygmund kernel bounds nor the $L^2(\mathbb{R})$-boundedness of the operator. But $K(1)=0$, so (2.2) is clearly violated. In contrast, (2.1) trivially holds; we can e.g. take $x=-r$ for any given $r>0$. If we also subtract $K_0(-1)\varphi(-x)$, so as to violate Stein's condition at both $x=\pm 1$, we still have (2.1), where we can e.g. take $x=\pm r$ when $r\in(0,\frac{1}{2}]\cup[2,\infty)$ and $x=\pm 2$ when $r\in(\frac{1}{2},2)$.

Example 2.1.4 *(Stein's non-degeneracy violated in a half-space, $d\geq 2$).* Let $d\geq 2$ and consider a homogeneous convolution kernel

$$K(x) = \frac{x_i x_j}{|x|^{d+2}} \cdot 1_{(0,\infty)}(x_j) = \frac{1}{|x|^d}\Omega\big(\frac{x}{|x|}\big), \quad \Omega(x) = x_i \cdot x_j \cdot 1_{(0,\infty)}(x_j);$$

this is the truncation of a second order Riesz transform to a half space. Since $s\mapsto s\cdot 1_{(0,\infty)}(s)$ is Lipschitz-continuous, $K(x-y)$ is an ω-Calderón–Zygmund kernel with $\omega(t)=t$, and it also satisfies $\int_{S^{d-1}}\Omega(u)\,\mathrm{d}\sigma(u) = 0$. Under these conditions, it is classical that K is the convolution kernel of a bounded operator on $L^2(\mathbb{R}^d)$ (see e.g. [12, Proposition II.5.5]). For any unit vector u_0, it is clear that $K(t\cdot u_0)$ must vanish for either all $t\in(0,\infty)$ or all $t\in(-\infty,0)$, so that Stein's condition (2.2) is impossible. On the other hand, for any $r>0$, choosing $x=2^{-1/2}(e_i+e_j)r$, where e_i,e_j are standard unit vectors, we have $x\in B(0,r)^c$ and $K(x)=(2^{-1/2}r)^2/r^{d+2}=2^{-1}r^{-d}$, so that K satisfied Definition 2.1.1(1).

Example 2.1.5 *(Stein's non-degeneracy violated on a half-line, $d=1$).* In dimension $d=1$, it takes a bit more effort to construct an analogue of Example 2.1.4; but this pays off, as it allows us to connect the example to the theory of *one-sided singular integrals* introduced by Aimar, Forzani and Martín-Reyes [1]. These are simply convolution-type ω-Calderón–Zygmund kernel K supported on $(0,\infty)$. The basic example of a non-trivial one-sided kernel provided in [1, (1.5)], $K(x)=1_{(0,\infty)}(x)\cdot x^{-1}\cdot\sin(\log x)/\log x$, decays a bit too fast at 0 and ∞ to satisfy Definition 2.1.1(1), but a non-degenerate example can be given as follows: With the resolution of unity (2.3), let

$$K(x) := \sum_{j\in\mathbb{Z}} \varphi(2^{-j}x)\frac{(-1)^j}{x}.$$

It is immediate that this satisfies the higher order Calderón–Zygmund estimates $|x|^n|D^n k(x)| \le c_n$ for all $n = 0, 1, 2, \ldots$, and in particular the ω-Calderón–Zygmund estimates with $\omega(t) = t$. Since consecutive bumps in the series of K have equal integral with opposite signs, K also satisfies the usual cancellation condition $|\int_{\varepsilon<|x|<N} K(x)\,\mathrm{d}x| \le C$ for all $0 < \varepsilon < N < \infty$. However, it fails to satisfy the existence of the limit $\lim_{\varepsilon\to 0}\int_{\varepsilon<|x|<1} K(x)\,\mathrm{d}x$, which is needed to define the associated principal value convolution operator in the classical theory. But if we take the limit $\varepsilon \to 0$ only along the powers $\varepsilon = 4^{-n}$, $n \in \mathbb{Z}$ (so as to proceed in steps of two consecutive bumps of opposite signs), then the relevant limit exists, and a trivial modification of the standard theory (see e.g. [12, Proposition II.5.5]) shows that $Tf(x) := \lim_{n\to\infty}\int_{|x-y|>4^{-n}} K(x-y)f(y)\,\mathrm{d}y$ defines a bounded operator on $L^2(\mathbb{R}^d)$ with convolution kernel K. Finally, this K easily satisfies Definition 2.1.1(1): Given $r > 0$, let $r \le 2^j < 2r$ so that $x = 2^j \in B(0,r)^c$ satisfies $|K(x)| = |(-1)^j x^{-1}| \ge (2r)^{-1}$.

Recall that Stein's non-degeneracy condition was introduced for the following result [33, IV.4.6, Proposition 7]: If a convolution operator with non-degenerate kernel in Stein's sense acts boundedly on a weighted space $L^p(w)$, then the weight w must belong to Muckenhoupt's class A_p. On the other hand, Aimar et al. [1] show that their one-sided operators, and hence in particular the example that we just gave, act boundedly on $L^p(w)$ for a strictly larger weight class A_p^-. As we will show in this paper, non-degeneracy in the sense of Definition 2.1.1(1) (which is satisfied by the said example) is enough to imply various necessary conditions on b for the boundedness of the commutator $[b,T]$. In particular, a weaker notion of non-degeneracy of a singular integral T is needed to deduce that $b \in \mathrm{BMO}$ from the L^p-boundedness of $[b,T]$, than what is needed to deduce that $w \in A_p$ from the $L^p(w)$-boundedness of T. This is perhaps unexpected in view of the many known connections between the two questions.

Example 2.1.6 *(Stein's non-degeneracy violated all over the place).* In the two previous examples, a variant of Stein condition would still be satisfied, if we only demanded (2.2) for $t \in (0,\infty)$. This final (arguably somewhat artificial) example shows that we can make (2.2) fail for a significantly larger set of $t \in \mathbb{R}$, while still retaining non-degeneracy in the sense of Definition 2.1.1(1).

Let $d \ge 2$ and φ be as in (2.3). Let $(w_k)_{k\in\mathbb{Z}}$ be a sequence of unit vectors that is dense in the unit sphere S^{d-1} of $\mathbb{R}^d$, and let $(v_j)_{j\in\mathbb{Z}}$ be a sequence that, for each w_k, contains arbitrarily long subsequences of constant value w_k. Fixing a resolution as in (2.3), let finally

$$K(x) := \sum_{j\in\mathbb{Z}} \varphi(2^{-j}|x|)\frac{x\cdot v_j}{|x|^{d+1}}.$$

It is immediate that K satisfies not only the ω-Calderón–Zygmund estimates with $\omega(t) = t$, but in fact the higher Calderón–Zygmund estimates $|\partial^\alpha K(x)| \le c_\alpha|x|^{-d-|\alpha|}$ of any order, and also that K has vanishing integral over any sphere centred at the origin. It is well-known (see again [12, Proposition II.5.5]) that, under these conditions, K is the convolution kernel of a singular integral operator bounded on $L^2(\mathbb{R}^d)$.

To see that K satisfies Definition 2.1.1(1), given $r > 0$, let $r \le 2^k < 2r$, and $x := 2^k v_k \in B(0,r)^c$. Then

$$K(x) = \frac{2^k v_k \cdot v_k}{|2^k v_k|^{d+1}} = 2^{-kd} > (2r)^{-d}.$$

On the other hand, let us fix some candidate unit-vector u_0 and $a > 0$ for Stein's condition (2.2), and choose another unit vector $u_1 \perp u_0$. By density, we can find some w_k with $|w_k - u_1| < \frac{1}{2}a$. Given $N > 0$ (large), we can find $v_n, v_{n+1}, \ldots, v_{n+N} \equiv w_k$. Then

$$K(x) \equiv \frac{x\cdot w_k}{|x|^{d+1}} \quad \text{whenever } 2^n \le |x| \le 2^{n+N},$$

and in particular

$$|K(t\cdot u_0)| = \frac{|tu_0\cdot w_k|}{|tu_0|^{d+1}} = |t|^{-d}|u_0\cdot(w_k-u_1)| \leq \frac{1}{2}a|t|^{-d} \quad \text{when } |t|\in[2^n,2^{n+N}].$$

So not only is (2.2) violated, but it is violated on symmetric line-segments that may be arbitrarily long relative to their distance from 0. On the other hand, the points of non-degeneracy for Definition 2.1.1(1), $x=2^k v_k$ where v_k are dense in S^{d-1}, have a rather wild distribution in the underlying space $\mathbb{R}^d$.

2.2. Consequences of non-degeneracy

We will use the assumption of non-degeneracy through the following result:

Proposition 2.2.1. *Let K be a non-degenerate Calderón–Zygmund kernel. Then for every $A\geq 3$ and every ball $B=B(y_0,r)$, there is a disjoint ball $\tilde B = B(x_0,r)$ at distance $\operatorname{dist}(B,\tilde B)\eqsim Ar$ such that*

$$|K(x_0,y_0)| \eqsim \frac{1}{A^d r^d}, \tag{2.4}$$

and for all $y_1\in B$ and $x_1\in\tilde B$, we have

$$\int_B |K(x_1,y)-K(x_0,y_0)|\,\mathrm{d}y + \int_{\tilde B} |K(x,y_1)-K(x_0,y_0)|\,\mathrm{d}x \lesssim \frac{\varepsilon_A}{A^d}, \tag{2.5}$$

where $\varepsilon_A\to 0$ as $A\to\infty$.

The implied constants can depend at most on c_K,ω and d, as well as c_0 or $|\Omega(\theta_0)|$ from Definition 2.1.1. If K is homogeneous, we can take $x_0=y_0+Ar\theta_0$.

Proof of Proposition 2.2.1, case (1). We assume that K is as in Definition 2.1.1(1). Fix a ball $B=B(y_0,r)$ and $A\geq 3$. We apply the assumption with y_0 in place of y and Ar in place of r. This produces a point $x_0\in B(y_0,Ar)^c$ such that

$$\frac{1}{c_0(Ar)^d} \leq |K(x_0,y_0)| \leq \frac{c_K}{|x_0-y_0|^d}.$$

Let $\tilde B := B(x_0,r)$. Then

$$Ar \leq |x_0-y_0| \leq (c_0c_K)^{1/d}Ar, \quad \operatorname{dist}(B,\tilde B)\eqsim |x_0-y_0|.$$

Moreover, if $x\in\tilde B$ and $y\in B$, then

$$\begin{aligned}
|K(x,y)-K(x_0,y_0)| &\leq |K(x,y)-K(x,y_0)| + |K(x,y_0)-K(x_0,y_0)| \\
&\leq \frac{1}{|x-y_0|^d}\omega\Big(\frac{|y-y_0|}{|x-y_0|}\Big) + \frac{1}{|x_0-y_0|^d}\omega\Big(\frac{|x-y_0|}{|x_0-y_0|}\Big) \\
&\leq \frac{1}{(Ar-r)^d}\omega\Big(\frac{r}{Ar-r}\Big) + \frac{1}{(Ar)^d}\omega\Big(\frac{r}{Ar}\Big) \\
&= \frac{1}{(Ar)^d}\Big[\frac{1}{(1-A^{-1})^d}\omega\Big(\frac{1}{A-1}\Big) + \omega\Big(\frac{1}{A}\Big)\Big] = \frac{\varepsilon_A}{(Ar)^d},
\end{aligned}$$

where $\varepsilon_A\to 0$ as $A\to\infty$ by the condition that $\omega(t)\to 0$ as $t\to 0$. Integrating this over $x\in\tilde B$ or $y\in B$, which both have measure $|\tilde B|=|B|\eqsim r^d$, we obtain (2.5). □

Proof of Proposition 2.2.1, case (2). We assume that K is as in Definition 2.1.1(2). Fix a ball $B = B(y_0, r)$ and $A \geq 3$. Let $x_0 = y_0 + Ar\theta_0$ and $\tilde{B} = B(x_0, r)$. Clearly $\operatorname{dist}(\tilde{B}, B) = (A-2)r \eqsim Ar$ and

$$|K(x_0, y_0)| = \frac{|\Omega(x_0 - y_0)|}{|x_0 - y_0|^d} = \frac{|\Omega(Ar\theta_0)|}{|Ar\theta_0|^d} = \frac{|\Omega(\theta_0)|}{(Ar)^d} \eqsim \frac{1}{(Ar)^d},$$

recalling that the implied constant was allowed to depend on $|\Omega(\theta_0)|$.

We then consider the integrals in (2.5). Writing $x \in B(x_0, r) = B(y_0 + Ar\theta_0, r)$ as $x = y_0 + Ar\theta_0 + ru$ and $y \in B(y_0, r)$ as $y = y_0 + rv$, where $u, v \in B(0, 1)$, and using the homogeneity of Ω, we have

$$\begin{aligned} K(x, y) - K(x_0, y_0) &= \frac{\Omega(x - y)}{|x - y|^d} - \frac{\Omega(x_0 - y_0)}{|x_0 - y_0|^d} \\ &= \frac{\Omega(Ar\theta_0 + r(u - v))}{|Ar\theta_0 + r(u - v)|^d} - \frac{\Omega(Ar\theta_0)}{|Ar\theta_0|^d} \\ &= \frac{1}{(Ar)^d}\Big(\frac{\Omega(\theta_0 + A^{-1}(u - v))}{|\theta_0 + A^{-1}(u - v)|^d} - \Omega(\theta_0)\Big), \end{aligned}$$

where

$$\begin{aligned} &\frac{\Omega(\theta_0 + A^{-1}(u - v))}{|\theta_0 + A^{-1}(u - v)|^d} - \Omega(\theta_0) \\ &= \frac{\Omega(\theta_0 + A^{-1}(u - v)) - \Omega(\theta_0)}{|\theta_0 + A^{-1}(u - v)|^d} + \Omega(\theta_0)\Big(\frac{1}{|\theta_0 + A^{-1}(u - v)|^d} - 1\Big) =: I + II. \end{aligned}$$

Here it is immediate that $|II| \lesssim A^{-1}$, and hence the integral of $(Ar)^{-d}II$ over either $x \in \tilde{B}$ or $y \in B$ is bounded by $A^{-d-1} = A^{-d}\varepsilon_A$.

We turn to term I. Keeping either $x \in \tilde{B}$ fixed and varying $y \in B$, or the other way round, the difference $u - v$ varies over a subset of $B(0, 2)$. Hence both $\int_{\tilde{B}}(Ar)^{-d}|I|\,\mathrm{d}x$ and $\int_B (Ar)^{-d}|I|\,\mathrm{d}y$ are dominated by

$$\begin{aligned} &\frac{1}{(Ar)^d}\int_{B(0,2)} \Big|\frac{\Omega(\theta_0 + A^{-1}z) - \Omega(\theta_0)}{|\theta_0 + A^{-1}z|^d}\Big| r^d\,\mathrm{d}z \\ &\lesssim A^{-d} \fint_{B(0,2/A)} |\Omega(\theta_0 + s) - \Omega(\theta_0)|\,\mathrm{d}s = A^{-d}\varepsilon_A \end{aligned}$$

by the assumption that θ_0 is a Lebesgue point of Ω. $\square$

2.3. Approximate weak factorisation

For the class of non-degenerate Calderón–Zygmund operators just described, we prove certain "weak factorisation" type results that are pivotal in our proof of Theorem 1.0.1. These results have a technical flavour and may fail to have an "independent interest", but they are precisely what we need below. For a ball $B \subset \mathbb{R}^d$, we denote

$$\begin{aligned} L^\infty(B) &= \{f \in L^\infty(\mathbb{R}^d) : f = 1_B f\}, \\ L_0^\infty(B) &= \{f \in L^\infty(B) : \int_B f = 0\}, \\ L_+^\infty(B) &= \{f \in L^\infty(B) : f \geq 0\}. \end{aligned}$$

Lemma 2.3.1. *Let T be a non-degenerate Calderón–Zygmund operator. Using the notation of Proposition 2.2.1, if $f \in L_0^\infty(B)$ and $g \in L_+^\infty(\tilde{B})$ is such that $\|g\|_\infty \lesssim \fint_{\tilde{B}} g$, then there is a decomposition*

$$f = gTh - hT^*g + \tilde{f},$$

where $\tilde{f} \in L_0^\infty(\operatorname{spt} g)$ and $h \in L^\infty(\operatorname{spt} f)$ satisfy

$$\|g\|_\infty \|h\|_\infty \lesssim A^d \|f\|_\infty, \qquad \|\tilde{f}\|_\infty \lesssim \varepsilon_A \|f\|_\infty,$$

provided that A is chosen large enough so that $\varepsilon_A \ll 1$.

Proof. The decomposition is given by

$$f = \frac{f}{T^*g} T^*g =: -hT^*g = -hT^*g + gTh - gTh =: -hT^*g + gTh + \tilde{f},$$

where we need to justify that the definition of $h := -f/T^*g$ does not involve division by zero. However, if $y \in B$, then

$$\begin{aligned} T^*g(y) &= \int_{\tilde{B}} K(x,y)g(x)\,\mathrm{d}x \\ &= K(x_0,y_0)\int_{\tilde{B}} g(x)\,\mathrm{d}x + \int_{\tilde{B}} [K(x,y) - K(x_0,y_0)]g(x)\,\mathrm{d}x = I + II, \end{aligned}$$

where, using Proposition 2.2.1,

$$|I| \eqsim \frac{1}{A^d r^d} \int_{\tilde{B}} g \eqsim \frac{1}{A^d} \fint_{\tilde{B}} g \eqsim \frac{1}{A^d}\|g\|_\infty$$

and

$$|II| \lesssim \int_{\tilde{B}} |K(x,y) - K(x_0,y_0)|\,\mathrm{d}x \|g\|_\infty \lesssim \frac{\varepsilon_A}{A^d}\|g\|_\infty,$$

so that

$$|T^*g(y)| = |I + II| \geq |I| - |II| \gtrsim \frac{1}{A^d}\|g\|_\infty,$$

recalling that A was chosen large enough so that $\varepsilon_A \ll 1$. This justifies the well-definedness of the decomposition, and we turn to the quantitative bounds.

From the previous considerations it directly follows that

$$\|g\|_\infty \|h\|_\infty \lesssim \|g\|_\infty \frac{\|f\|_\infty}{A^{-d}\|g\|_\infty} = A^d \|f\|_\infty.$$

It is also immediate that

$$-\int_{\tilde{B}} \tilde{f} = \int gTh = \int hT^*g = \int \frac{f}{T^*g} T^*g = \int_B f = 0.$$

Let us then estimate

$$Th = T\Big(\frac{f}{T^*g}\Big) = T\Big(\frac{f}{T^*g} - \frac{f}{K(x_0,y_0)\int_{\tilde B} g}\Big) + \frac{1}{K(x_0,y_0)\int_{\tilde B} g} Tf =: I' + II'.$$

For $y \in B$,

$$\begin{aligned}
&\Big|\frac{1}{T^*g(y)} - \frac{1}{K(x_0,y_0)\int_{\tilde B} g(y)}\Big| = \Big|\frac{K(x_0,y_0)\int_{\tilde B} g - T^*g(y)}{T^*g(y)K(x_0,y_0)\int_{\tilde B} g}\Big| \\
&\lesssim \frac{1}{(A^{-d}\|g\|_\infty)^2}\int_{\tilde B} |K(x_0,y_0) - K(x,y)||g(x)|\,\mathrm{d}x \\
&\lesssim \frac{1}{(A^{-d}\|g\|_\infty)^2}\frac{\varepsilon_A}{A^d}\|g\|_\infty = \frac{A^d\varepsilon_A}{\|g\|_\infty}.
\end{aligned}$$

Hence for $x \in \tilde B$,

$$\begin{aligned}
|I'(x)| &\le \int_B |K(x,y)|\|f\|_\infty \frac{A^d\varepsilon_A}{\|g\|_\infty}\,\mathrm{d}y \\
&\lesssim \fint_B \frac{1}{A^d}\|f\|_\infty \frac{A^d\varepsilon_A}{\|g\|_\infty}\,\mathrm{d}y = \varepsilon_A\frac{\|f\|_\infty}{\|g\|_\infty}.
\end{aligned}$$

On the other hand, recalling that $f \in L_0^\infty(B)$,

$$\begin{aligned}
|Tf(x)| &= \Big|\int_B [K(x,y) - K(x_0,y_0)]f(y)\,\mathrm{d}y\Big| \\
&\le \int_B |K(x,y) - K(x_0,y_0)|\,\mathrm{d}y\|f\|_\infty \lesssim \frac{\varepsilon_A}{A^d}\|f\|_\infty,
\end{aligned}$$

and thus

$$|II'(x)| \lesssim \frac{|Tf(x)|}{A^{-d}\|g\|_\infty} \lesssim \varepsilon_A\frac{\|f\|_\infty}{\|g\|_\infty}.$$

It is then immediate that

$$\|\tilde f\|_\infty = \|gTh\|_\infty \lesssim \|g\|_\infty\varepsilon_A\frac{\|f\|_\infty}{\|g\|_\infty} = \varepsilon_A\|f\|_\infty. \quad \square$$

By iterating the previous decomposition (but just once more), we achieve the useful additional property that the error term is supported on the same set as the original function. In the following lemma and below, we will make use of the following notion:

Definition 2.3.2 *(Major subset).* If $E \subset F \subset \mathbb{R}^d$ are sets of finite measure, we say that E is a *major subset* of F if $|E| \ge c|F|$ for some fixed constant $c \in (0,1)$ that depends only on the admissible parameters, as described in Section 1.7.

In the following lemma, we denote certain major subsets by the suggestive letter Q, since the main subsequent application deals with the case, where these sets are cubes; however, the lemma itself does not require assuming this.

Lemma 2.3.3. *Let T be a non-degenerate Calderón–Zygmund operator. Let B and $\tilde{B}$ be as in Proposition 2.2.1, and $Q \subset B$, $\tilde{Q} \subset \tilde{B}$ be their major subsets, i.e., $|Q| \gtrsim |B|$ and $|\tilde{Q}| \gtrsim |\tilde{B}|$.*

If $f \in L^\infty_0(Q)$, there is a decomposition

$$f = \sum_{i=1}^{2} (g_i T h_i - h_i T^* g_i) + \tilde{\tilde{f}}, \tag{2.6}$$

where $\tilde{\tilde{f}} \in L^\infty_0(Q)$, $g_i \in L^\infty(\tilde{Q})$ and $h_i \in L^\infty(Q)$ satisfy

$$\|g_i\|_\infty \|h_i\|_\infty \lesssim A^d \|f\|_\infty, \qquad \|\tilde{\tilde{f}}\|_\infty \lesssim \varepsilon_A \|f\|_\infty,$$

provided that A is chosen large enough so that $\varepsilon_A \ll 1$.

Proof. We first apply Lemma 2.3.1 to f and $g_1 := 1_{\tilde{Q}} \in L^\infty_+(\tilde{B})$, which clearly satisfies the condition $\|g\|_\infty = 1 \lesssim |\tilde{Q}|/|\tilde{B}| = \fint_{\tilde{B}} g$. Thus Lemma 2.3.1 yields a decomposition

$$f = g_1 T h_1 - h_1 T^* g_1 + \tilde{f},$$

where $\tilde{f} \in L^\infty_0(\operatorname{spt} g_1) = L^\infty_0(\tilde{Q})$, $g_1 \in L^\infty(\tilde{Q})$ and $h_1 \in L^\infty(\operatorname{spt} f) \subset L^\infty(Q)$ with the estimates

$$\|g_1\|_\infty \|h_1\|_\infty \lesssim A^d \|f\|_\infty, \qquad \|\tilde{f}\|_\infty \lesssim \varepsilon_A \|f\|_\infty.$$

We then wish to apply Lemma 2.3.1 again, this time to the functions $\tilde{f}$ and $\tilde{g} := 1_Q \in L^\infty_+(B)$, and the adjoint operator T^* in place of T. For this, we notice that the conclusions of Proposition 2.2.1 are preserved under the replacement of $(B, \tilde{B}, T)$ by $(\tilde{B}, B, T^*)$. Hence Lemma 2.3.1 provides a decomposition

$$\tilde{f} = \tilde{g} T^* \tilde{h} - \tilde{h} T \tilde{g} + \tilde{\tilde{f}},$$

where $\tilde{\tilde{f}} \in L^\infty_0(\operatorname{spt} \tilde{g}) = L^\infty_0(Q)$, $\tilde{g} \in L^\infty(Q)$ and $\tilde{h} \in L^\infty(\operatorname{spt} \tilde{f}) \subset L^\infty(\tilde{Q})$ with the estimates

$$\|\tilde{g}\|_\infty \|\tilde{h}\|_\infty \lesssim A^d \|\tilde{f}\|_\infty \lesssim A^d \|f\|_\infty, \qquad \|\tilde{\tilde{f}}\|_\infty \lesssim \varepsilon_A \|\tilde{f}\|_\infty \lesssim \varepsilon_A \|f\|_\infty.$$

(We could write ε_A^2 in the ultimate right, but since $\varepsilon_A \to 0$ at an unspecified rate anyway, this is irrelevant.) It remains to define $g_2 := -\tilde{h} \in L^\infty(\tilde{Q})$, $h_2 := \tilde{g} \in L^\infty(Q)$ so that

$$\tilde{g} T^* \tilde{h} - \tilde{h} T \tilde{g} = -h_2 T^* g_2 + g_2 T h_2,$$

and we get the required decomposition (2.6). □

2.4. Necessary conditions for $[b, T] : L^p \to L^q$ when $p \le q$

We now come to the proof of some of the "only if" directions of Theorem 1.0.1. Assuming a weak form of the boundedness of the commutator $[b, T]$, we wish to derive the membership of b in a suitable function space, with estimates for its norm. The relevant spaces here will be the functions of bounded mean oscillation,

$$\operatorname{BMO}(\mathbb{R}^d) := \Big\{ b \in L^1_{\mathrm{loc}}(\mathbb{R}^d) \Big| \|b\|_{\mathrm{BMO}} := \sup_B \fint_B |b - \langle b \rangle_B| < \infty \Big\},$$

where the supremum is over all balls $B \subset \mathbb{R}^d$, and the homogeneous Hölder spaces

$$\dot{C}^{0,\alpha}(\mathbb{R}^d) := \Big\{ b : \mathbb{R}^d \to \mathbb{C} \Big| \|b\|_{\dot{C}^{0,\alpha}} := \sup_{x \neq y} \frac{|b(x)-b(y)|}{|x-y|^\alpha} < \infty \Big\}.$$

Note that we do not impose any boundedness condition on b; this would lead to the inhomogeneous Hölder space $C^{0,\alpha}$, which does not play any role in our results.

Theorem 2.4.1. *Let K be a non-degenerate Calderón–Zygmund kernel, and $b \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$. Let further*

$$1 < p \leq q < \infty, \qquad \alpha := d\Big(\frac{1}{p} - \frac{1}{q}\Big) \geq 0,$$

and suppose that $[b,T]$ satisfies the following weak form of $L^p \to L^q$ boundedness:

$$\begin{aligned} |\langle [b,T]f, g\rangle| &= \Big| \iint (b(x)-b(y))K(x,y)f(y)g(x)\,\mathrm{d}y\,\mathrm{d}x \Big| \\ &\leq \Theta \cdot \|f\|_\infty |B|^{1/p} \cdot \|g\|_\infty |\tilde{B}|^{1/q'}, \end{aligned} \tag{2.7}$$

whenever $f \in L^\infty(B)$, $g \in L^\infty(\tilde{B})$ for any two balls of equal radius r and distance $\mathrm{dist}(B,\tilde{B}) \gtrsim r$. Then

- *if $\alpha = 0$, equivalently $p = q$, we have $b \in \mathrm{BMO}(\mathbb{R}^d)$, and $\|b\|_{\mathrm{BMO}} \lesssim \Theta$;*
- *if $\alpha \in (0,1]$, we have $b \in \dot{C}^{0,\alpha}(\mathbb{R}^d)$, and $\|b\|_{\dot{C}^{0,\alpha}} \lesssim \Theta$;*
- *if $\alpha > 1$, the function b is constant, so in fact $[b,T] = 0$.*

Proof. Let us consider a fixed ball $B \subset \mathbb{R}^d$ of radius r. Then

$$\fint_B |b - \langle b\rangle_B| \eqsim \sup_{\substack{f \in L^\infty_0(B) \\ \|f\|_\infty \leq 1}} \Big| \fint_B bf \Big|$$

is finite by the assumption that $b \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$. Given $f \in L^\infty_0(B)$, we apply Lemma 2.3.3 to write

$$f = \sum_{i=1}^2 (g_i T h_i - h_i T^* g_i) + \tilde{\tilde{f}},$$

where $\tilde{\tilde{f}} \in L^\infty_0(B)$, $g_i \in L^\infty(\tilde{B})$ and $h_i \in L^\infty(B)$ satisfy

$$\|g_i\|_\infty \|h_i\|_\infty \lesssim A^d \|f\|_\infty, \qquad \|\tilde{\tilde{f}}\|_\infty \lesssim \varepsilon_A \|f\|_\infty,$$

and $\tilde{B}$ is another ball of radius r such that $\mathrm{dist}(B,\tilde{B}) \eqsim Ar$.

Then

$$\begin{aligned} \int bf &= \sum_{i=1}^2 \int b(g_i T h_i - h_i T^* g_i) + \int b\tilde{\tilde{f}} \\ &= \sum_{i=1}^2 \int [g_i b T h_i - g_i T(b h_i)] + \int b\tilde{\tilde{f}} = \sum_{i=1}^2 \int g_i [b,T] h_i + \int b\tilde{\tilde{f}}, \end{aligned}$$

where, by assumption (2.7),

$$\Big| \int g_i [b,T] h_i \Big| \leq \Theta \cdot \|g_i\|_\infty \|h_i\|_\infty \cdot |B|^{1/p + 1/q'} \lesssim \Theta \cdot A^d \|f\|_\infty \cdot |B| \cdot r^\alpha.$$

Thus

$$\Big|\fint_B bf\Big| \lesssim \Theta \cdot A^d \|f\|_\infty r^\alpha + \Big|\fint_B b\tilde{\tilde{f}}\Big|,$$

where

$$\Big|\fint_B b\tilde{\tilde{f}}\Big| \leq \Big(\fint_B |b - \langle b\rangle_B|\Big)\|\tilde{\tilde{f}}\|_\infty \lesssim \Big(\fint_B |b - \langle b\rangle_B|\Big)\varepsilon_A \|f\|_\infty.$$

Taking the supremum over $f \in L^\infty_0(B)$ of norm one, we deduce that

$$\fint_B |b - \langle b\rangle_B| \lesssim \Theta A^d r^\alpha + \varepsilon_A \fint_B |b - \langle b\rangle_B|,$$

and the last term can be absorbed if A is fixed large enough, depending only on the implied constants. Thus

$$\fint_B |b - \langle b\rangle_B| \lesssim \Theta r^\alpha.$$

If $\alpha = 0$, this is precisely the condition $b \in \mathrm{BMO}(\mathbb{R}^d)$ with the claimed estimate.

For $\alpha > 0$, this is also a well-known reformulation of $b \in \dot{C}^{0,\alpha}(\mathbb{R}^d)$ (which consists only of constants for $\alpha > 1$). We recall the argument for completeness.

Let x_i, $i = 1, 2$, be two Lebesgue points of b, and let $r := |x_1 - x_2|$. Then

$$b(x_i) = \lim_{t\to 0} \langle b\rangle_{B(x_i,t)} = \sum_{k=0}^\infty \big(\langle b\rangle_{B(x_i,2^{-k-1}r)} - \langle b\rangle_{B(x_i,2^{-k}r)}\big) + \langle b\rangle_{B(x_i,r)}.$$

If $B \subset B^*$ are two balls of radius comparable to R, then

$$|\langle b\rangle_B - \langle b\rangle_{B^*}| = \Big|\fint_B (b - \langle b\rangle_{B^*})\Big| \leq \fint_B |b - \langle b\rangle_{B^*}| \lesssim \fint_{B^*} |b - \langle b\rangle_{B^*}| \lesssim \Theta R^\alpha, \tag{2.8}$$

and thus

$$\sum_{k=0}^\infty |\langle b\rangle_{B(x_i,2^{-k-1}r)} - \langle b\rangle_{B(x_i,2^{-k}r)}| \lesssim \sum_{k=0}^\infty \Theta(2^{-k}r)^\alpha \lesssim \Theta r^\alpha.$$

Hence

$$|b(x_1) - b(x_2)| \lesssim \Theta r^\alpha + |\langle b\rangle_{B(x_1,r)} - \langle b\rangle_{B(x_2,r)}|,$$

where another application of (2.8) shows that

$$|\langle b\rangle_{B(x_i,r)} - \langle b\rangle_{B(\frac{1}{2}(x_1+x_2),2r)}| \lesssim \Theta r^\alpha.$$

Thus altogether

$$|b(x_1) - b(x_2)| \lesssim \Theta r^\alpha = \Theta |x_1 - x_2|^\alpha,$$

and this can be extended to all x_1, x_2 by redefining b in a set of measure zero. This is the required bound for $\|b\|_{\dot{C}^{0,\alpha}}$ if $\alpha \in (0, 1]$.

If $\alpha > 1$, we let $y_k := x_1 + N^{-1}k(x_2 - x_1)$ for $k = 0, 1, \ldots, N$ to deduce that

$$|b(x_1) - b(x_2)| \leq \sum_{k=1}^{N} |b(y_k) - b(y_{k-1})| \lesssim \sum_{k=1}^{N} \Theta(N^{-1}r)^\alpha = N^{1-\alpha}\Theta r^\alpha.$$

With $N \to \infty$, this shows that $b(x_1) = b(x_2)$, and hence b is constant. □

2.5. Necessary condition for $[b, T] : L^p \to L^q$ when $p > q$

We now come to the more exotic case of Theorem 1.0.1, which is precisely restated in the following:

Theorem 2.5.1. *Let K be a non-degenerate Calderón–Zygmund kernel, and $b \in L^1_{\text{loc}}(\mathbb{R}^d)$. Let*

$$1 < q < p < \infty, \qquad r = \frac{pq}{p - q} \in (1, \infty),$$

and suppose that $[b, T]$ satisfies the following weak form of $L^p \to L^q$ boundedness:

$$\sum_{i=1}^{N} |\langle [b, T]f_i, g_i\rangle| \leq \Theta \Big\| \sum_{i=1}^{N} \|f_i\|_\infty 1_{Q_i} \Big\|_p \Big\| \sum_{i=1}^{N} \|g_i\|_\infty 1_{\tilde{Q}_i} \Big\|_{q'}, \tag{2.9}$$

whenever, for each $i = 1, \ldots, N$, we have $f_i \in L^\infty(Q_i)$ and $g_i \in L^\infty(\tilde{Q}_i)$ for cubes Q_i and $\tilde{Q}_i$ such that $\operatorname{dist}(Q_i, \tilde{Q}_i) \gtrsim \operatorname{diam}(Q_i) = \operatorname{diam}(\tilde{Q}_i)$.

Then $b = a + c$ for some $a \in L^r(\mathbb{R}^d)$ and some constant $c \in \mathbb{C}$, where $\|a\|_r \lesssim \Theta$.

Note that each term on the left of (2.9) can be defined as in (2.7). In order to better understand the assumption (2.9), we include:

Lemma 2.5.2.

1. *For any $p, q \in (1, \infty)$, (2.9) follows if $[b, T]$ exists as a bounded linear operator $[b, T] : L^p \to L^q$, and $\Theta(2.9) \leq \|[b, T]\|_{L^p \to L^q}$.*
2. *If $p \leq q$, then (2.9) follows from (2.7), and $\Theta(2.9) \leq \Theta(2.7)$.*

Proof. For certain fixed signs σ_i, and random signs ε_i on some probability space with expectation denoted by $\mathbb{E}$, we have

$$\begin{aligned}\sum_{i=1}^{N} |\langle [b, T]f_i, g_i\rangle| &= \sum_{i=1}^{N} \sigma_i \langle [b, T]f_i, g_i\rangle = \mathbb{E}\Big\langle [b, T] \sum_{i=1}^{N} \varepsilon_i \sigma_i f_i, \sum_{j=1}^{N} \varepsilon_j g_j \Big\rangle \\ &\leq \mathbb{E}\|[b, T]\|_{L^p \to L^q} \Big\| \sum_{i=1}^{N} \varepsilon_i \sigma_i f_i \Big\|_p \Big\| \sum_{j=1}^{N} \varepsilon_j g_j \Big\|_{q'} \\ &\leq \|[b, T]\|_{L^p \to L^q} \Big\| \sum_{i=1}^{N} |f_i| \Big\|_p \Big\| \sum_{j=1}^{N} |g_j| \Big\|_{q'}.\end{aligned}$$

If $p \leq q$, using (2.7) followed by Hölder's inequality and several applications of $\| \ \|_{\ell^s} \leq \| \ \|_{\ell^t}$ if $t \leq s$, we find that

$$\begin{aligned}\sum_{i=1}^N |\langle [b,T]f_i, g_i\rangle| &\le \sum_{i=1}^N \Theta \|f_i\|_p \|g_i\|_{q'} \le \Theta\Big(\sum_{i=1}^N \|f_i\|_p^q\Big)^{1/q}\Big(\sum_{i=1}^N \|g_i\|_{q'}^{q'}\Big)^{1/q'} \\ &\le \Theta\Big(\sum_{i=1}^N \|f_i\|_p^p\Big)^{1/p}\Big(\sum_{i=1}^N \|g_i\|_{q'}^{q'}\Big)^{1/q'} \\ &= \Theta\Big\|\Big(\sum_{i=1}^N |f_i|^p\Big)^{1/p}\Big\|_p \Big\|\Big(\sum_{i=1}^N |g_i|^{q'}\Big)^{1/q'}\Big\|_{q'} \le \Theta\Big\|\sum_{i=1}^N |f_i|\Big\|_p \Big\|\sum_{i=1}^N |g_i|\Big\|_{q'},\end{aligned}$$

where $\Theta = \Theta(2.7)$. $\square$

For the proof of Theorem 2.5.1, we need the following lemma. Given a cube $Q_0 \subset \mathbb{R}^d$, we denote by $\mathscr{D}(Q_0)$ the collection of its dyadic subcubes (obtained by repeatedly bisecting each side of the initial cube any finite number of times).

Lemma 2.5.3. *Let $f \in L_0^\infty(Q_0)$ for some cube $Q_0 \subset \mathbb{R}^d$. Then it has a decomposition*

$$f = \sum_{n=0}^N f_n, \qquad f_n = \sum_{k=0}^\infty f_{n,k},$$

where $f_{n,k} \in L_0^\infty(Q_{n,k})$ and $\|f_{n,k}\|_\infty \lesssim \langle |f|\rangle_{Q_{n,k}}$, and $Q_{n,k} \in \mathscr{D}(Q_0)$ are disjoint in k for each n. Moreover, for all n and k we have $Q_{n,k} \subset Q_{n-1,j}$ for a unique j, and $\langle |f|\rangle_{Q_{n,k}} > 2\langle |f|\rangle_{Q_{n-1,j}}$.

Proof. Let $\mathcal{F}_0 = \{Q_0\}$ and

$$\mathcal{F}_{n+1} := \bigcup_{F\in\mathcal{F}_n} \operatorname{ch}_{\mathcal{F}} F, \quad \operatorname{ch}_{\mathcal{F}} F := \{Q \in \mathscr{D}(F) \text{ maximal with } \langle |f|\rangle_Q > 2\langle |f|\rangle_F\}.$$

Since $f \in L^\infty(Q_0)$, there is a finite N such that $\mathcal{F}_n = \varnothing$ for all $n > N$. For $F \in \mathcal{F} = \bigcup_{n=0}^N \mathcal{F}_n$, we define

$$E(F) := F \setminus \bigcup_{F'\in \operatorname{ch}_{\mathcal{F}} F} F'$$

and then

$$f_F := 1_{E(F)} f + \sum_{F'\in\operatorname{ch}_{\mathcal{F}} F} 1_{F'}\langle f\rangle_{F'} - 1_F\langle f\rangle_F,$$

so that $\int f_F = 0$ and

$$|f_F| \le 1_{E(F)} 2\langle |f|\rangle_F + \sum_{F'\in\operatorname{ch}_{\mathcal{F}} F} 1_{F'} 2\cdot 2^d\langle |f|\rangle_F + 1_F\langle f\rangle_F \lesssim 1_F\langle |f|\rangle_F.$$

Letting $(f_{n,k}, Q_{n,k})_{k=0}^\infty$ be some enumeration of $(f_F, F)_{F\in\mathcal{F}_n}$, the claimed properties are easily checked. $\square$

Lemma 2.5.4. *Let Q_k be cubes, and $E_k \subset Q_k$ their subsets with $|E_k| \ge \eta|Q_k|$ for some $\eta \in (0,1)$. Let $\lambda_k \ge 0$ and $p \in [1,\infty)$. Then*

$$\Big\|\sum_{k=0}^\infty \lambda_k 1_{Q_k}\Big\|_p \lesssim \frac{1}{\eta}\Big\|\sum_{k=0}^\infty \lambda_k 1_{E_k}\Big\|_p$$

For $A \ge 1$, the bound is also true with A^d in place of $1/\eta$, if $E_k = \tilde{Q}_k$ is another cube with $\operatorname{dist}(Q_k, \tilde{Q}_k) \le$ *$A\ell(Q_k) = A\ell(\tilde{Q}_k)$.*

In the second claim, a more delicate argument could be given to improve the bound A^d to $\log A$, but this is unnecessary for the present purposes.

Proof. Dualising the left side with $\phi \in L^{p'}$, we find that

$$\int \Big(\sum_{k=0}^{\infty} \lambda_k 1_{Q_k}\Big)\phi = \sum_{k=0}^{\infty} \lambda_k \int_{Q_k} \phi \leq \sum_{k=0}^{\infty} \lambda_k \frac{|E_k|}{\eta} \fint_{Q_k} \phi = \frac{1}{\eta} \int \Big(\sum_{k=0}^{\infty} \lambda_k 1_{E_k}\Big) M\phi,$$

and the first claim by the boundedness of the maximal operator on $L^{p'}$.

For the second claim, let Q_k^* be a cube that contains both Q_k and $\tilde{Q}_k$, with $\ell(Q_k^*) \lesssim A\ell(Q_k)$. Then we first use the trivial bound $1_{Q_k} \leq 1_{Q_k^*}$, and then the first part of the lemma with Q_k^* in place of Q_k, and $\tilde{Q}_k \subset Q_k^*$ in place of E_k, observing that $|\tilde{Q}_k| \gtrsim A^{-d}|Q_k^*|$. □

Proof of Theorem 2.5.1. We fix a cube $Q_0 \subset \mathbb{R}^d$ and consider the quantity

$$C_R := \sup\Big\{\Big|\int_{Q_0} bf\Big| : f \in L_0^\infty(Q_0), \|f\|_\infty \leq R, \|f\|_{r'} \leq 1\Big\}.$$

This has the trivial a priori upper bound $C_R \leq \|b\|_{L^1(Q_0)} R < \infty$, since $b \in L^1_{\mathrm{loc}}(R^d)$, but we wish to deduce a bound independent of R. To this end, we fix an $f \in L_0^\infty(Q_0)$ and make the decomposition given by Lemma 2.5.3. Then

$$\int_{Q_0} bf = \sum_{n=0}^{N} \int_{Q_0} bf_n = \sum_{n=0}^{N} \sum_{k=0}^{\infty} \int_{Q_{n,k}} bf_{n,k};$$

the last step follows since $bf_n = \sum_{k=0}^{\infty} bf_{n,k}$ is integrable and the terms $bf_{n,k}$ are disjointly supported. For each (k,n), we apply the decomposition of Lemma 2.3.3 to write

$$f_{n,k} = \sum_{i=1}^{2} (g_{n,k}^i T h_{n,k}^i - h_{n,k}^i T^* g_{n,k}^i) + \tilde{\tilde{f}}_{n,k},$$

where $\tilde{\tilde{f}}_{n,k} \in L_0^\infty(Q_{n,k})$, $g_{n,k}^i \in L^\infty(\tilde{Q}_{n,k})$ and $h_{n,k}^i \in L^\infty(Q_{n,k})$ for some cubes $\tilde{Q}_{n,k}$ of the same size as $Q_{n,k}$ and distance $\mathrm{dist}(Q_{n,k}, \tilde{Q}_{n,k}) \eqsim A\,\mathrm{diam}(Q_{n,k})$. In particular, the functions $\tilde{\tilde{f}}_{n,k} \in L_0^\infty(Q_{n,k})$ are again disjointly supported with respect to k, for each fixed n. Thus

$$\int_{Q_{n,k}} bf_{n,k} = \sum_{i=1}^{2} \int g_{n,k}^i [b,T] h_{n,k}^i + \int_{Q_{n,k}} b\tilde{\tilde{f}}_{n,k}.$$

Since both the left side and the second term on the right is summable over k, so is the first term on the right, and we have

$$\int_{Q_0} bf_n = \sum_{k=0}^{\infty} \sum_{i=1}^{2} \int g_{n,k}^i [b,T] h_{n,k}^i + \int_{Q_0} b\tilde{\tilde{f}}_n, \qquad \tilde{\tilde{f}}_n := \sum_{k=0}^{\infty} \tilde{\tilde{f}}_{n,k}.$$

Summing over $n = 0, 1, \ldots, N$, we further deduce that

$$\int_{Q_0} bf = \sum_{n=0}^{N}\sum_{k=0}^{\infty}\sum_{i=1}^{2}\int g_{n,k}^i[b,T]h_{n,k}^i + \int_{Q_0} b\tilde{\tilde{f}}, \qquad \tilde{\tilde{f}} := \sum_{n=0}^{N}\tilde{\tilde{f}}_n.$$

We notice that

$$\begin{aligned}|\tilde{\tilde{f}}| &\leq \sum_{n=0}^{N}\sum_{k=0}^{\infty}|\tilde{\tilde{f}}_{n,k}| \leq \sum_{n=0}^{N}\sum_{k=0}^{\infty}\|\tilde{\tilde{f}}_{n,k}\|_\infty 1_{Q_{n,k}} \lesssim \varepsilon_A \sum_{n=0}^{N}\sum_{k=0}^{\infty}\|f_{n,k}\|_\infty 1_{Q_{n,k}} \\ &\lesssim \varepsilon_A \sum_{n=0}^{N}\sum_{k=0}^{\infty}\langle|f_{n,k}|\rangle_{Q_{n,k}} 1_{Q_{n,k}} \lesssim \varepsilon_A Mf.\end{aligned}$$

This pointwise maximal function bound proves both

$$\|\tilde{\tilde{f}}\|_\infty \lesssim \varepsilon_A\|f\|_\infty \leq \varepsilon_A R, \qquad \|\tilde{\tilde{f}}\|_{r'} \lesssim \varepsilon_A\|f\|_{r'} \leq \varepsilon_A,$$

so that

$$\Big|\int_{Q_0} b\tilde{\tilde{f}}\Big| \lesssim \varepsilon_A C_R.$$

On the other hand, by the definition of convergent series, we have

$$\sum_{n=0}^{N}\sum_{k=0}^{\infty}\sum_{i=1}^{2}\int g_{n,k}^i[b,T]h_{n,k}^i = \lim_{K\to\infty}\sum_{n=0}^{N}\sum_{k=0}^{K}\sum_{i=1}^{2}\int g_{n,k}^i[b,T]h_{n,k}^i.$$

Recalling that $g_{n,k}^i \in L^\infty(\tilde{Q}_{n,k})$ and $h_{n,k}^i \in L^\infty(Q_{n,k})$, where $\operatorname{dist}(Q_{n,k},\tilde{Q}_{n,k}) \eqsim A\operatorname{diam}(Q_{n,k}) = A\operatorname{diam}(\tilde{Q}_{n,k})$, the finite triple sum has exactly the form appearing in (2.9), and we can estimate

$$\Big|\sum_{n=0}^{N}\sum_{k=0}^{K}\sum_{i=1}^{2}\int g_{n,k}^i[b,T]h_{n,k}^i\Big| \leq \Theta\Big\|\sum_{n,k,i}\|g_{n,k}^i\|_\infty 1_{\tilde{Q}_{n,k}}\Big\|_{q'}\Big\|\sum_{n,k,i}\|h_{n,k}^i\|_\infty 1_{Q_{n,k}}\Big\|_p.$$

Note that $g_{n,k}^i$ and $h_{n,k}^i$ appear in the decomposition of $f_{n,k}^i$ in a bilinear way so that we are free to multiply these functions by any $\alpha > 0$ and α^{-1}, respectively. In particular, since $1/r = 1/q - 1/p$ implies that $1/r' = 1/q' + 1/p$, we may arrange the bound

$$\|g_{n,k}^i\|_\infty\|h_{n,k}^i\|_\infty \lesssim A^d\|f_{n,k}\|_\infty \lesssim A^d\langle|f|\rangle_{Q_{n,k}}$$

into the form

$$\|g_{n,k}^i\|_\infty \lesssim A^d\langle|f|\rangle_{Q_{n,k}}^{r'/q'}, \qquad \|h_{n,k}^i\|_\infty \lesssim \langle|f|\rangle_{Q_{n,k}}^{r'/p}.$$

Thus

$$\sum_{n,k,i}\|h_{n,k}^i\|_\infty 1_{Q_{n,k}} \lesssim \sum_{n,k}\langle|f|\rangle_{Q_{n,k}}^{r'/p} 1_{Q_{n,k}}.$$

At a fixed point $x \in Q_{N,k_N} \subset \ldots \subset Q_{1,k_1} \subset Q_0$, the averages $\langle|f|\rangle_{Q_{n,k_n}}$ satisfy $\langle|f|\rangle_{Q_{n+1,k_{n+1}}} > 2\langle|f|\rangle_{Q_{n,k_n}}$; thus $\langle|f|\rangle_{Q_{n,k_n}} \leq 2^{n-N}\langle|f|\rangle_{Q_{N,k_N}}$, and hence

$$\sum_{n=0}^{N}\langle|f|\rangle_{Q_{n,k_n}}^{r'/p}\leq\sum_{n=0}^{N}2^{(n-N)r'/p}\langle|f|\rangle_{Q_{N,k_N}}^{r'/p}\lesssim\langle|f|\rangle_{Q_{N,k_N}}^{r'/p}\leq(Mf(x))^{r'/p},$$

so that

$$\sum_{n,k}\langle|f|\rangle_{Q_{n,k}}^{r'/p}1_{Q_{n,k}}\lesssim(Mf)^{r'/p},$$

and hence

$$\Big\|\sum_{n,k,i}\|h_{n,k}^i\|_\infty 1_{Q_{n,k}}\Big\|_p\lesssim\|(Mf)^{r'/p}\|_p=\|Mf\|_{r'}^{r'/p}\lesssim\|f\|_{r'}^{r'/p}\leq 1.$$

For the similar term involving the $g_{n,k}^i$, we need in addition Lemma 2.5.4:

$$\Big\|\sum_{n,k,i}\|g_{n,k}^i\|_\infty 1_{\tilde{Q}_{n,k}}\Big\|_{q'}\lesssim A^d\Big\|\sum_{n,k,i}\|g_{n,k}^i\|_\infty 1_{Q_{n,k}}\Big\|_{q'}\lesssim A^{2d}\Big\|\sum_{n,k,i}\langle f_{n,k}\rangle_{Q_{n,k}}^{r'/q'}1_{Q_{n,k}}\Big\|_{q'},$$

where, as before,

$$\Big\|\sum_{n,k,i}\langle f_{n,k}\rangle_{Q_{n,k}}^{r'/q'}1_{Q_{n,k}}\Big\|_{q'}\lesssim\|(Mf)^{r'/q'}\|_{q'}=\|Mf\|_{r'}^{r'/q'}\lesssim\|f\|_{r'}^{r'/q'}\leq 1.$$

Collecting the bounds, we have proved that

$$\begin{aligned}\Big|\int_{Q_0}bf\Big|&\leq\lim_{K\to\infty}\Big|\sum_{n=0}^{N}\sum_{k=0}^{K}\sum_{i=1}^{2}\int g_{n,k}^i[b,T]h_{n,k}^i\Big|+\Big|\int_{Q_0}b\tilde{\tilde{f}}\Big|\\&\lesssim A^{2d}\Theta+\varepsilon_A C_R\end{aligned}$$

for all $f\in L_0^\infty(Q_0)$ with $\|f\|_{r'}\leq 1$ and $\|f\|_\infty\leq R$, and thus

$$C_R\lesssim A^{2d}\Theta+\varepsilon_A C_R.$$

Fixing A large enough so that $\varepsilon_A\ll 1$, we can absorb the last term and conclude that $C_R\lesssim\Theta$.

Let $L_{c,0}^\infty(\mathbb{R}^d):=\bigcup_{Q\subset\mathbb{R}^d}L_0^\infty(Q)$. Since every $f\in L_{c,0}^\infty(\mathbb{R}^d)$ satisfies $\|f\|_\infty\leq R$ for some R, we conclude that

$$\Big|\int bf\Big|\lesssim\Theta\|f\|_{r'}\qquad\forall f\in L_{c,0}^\infty(\mathbb{R}^d).$$

As this is a dense subspace of $L^{r'}(\mathbb{R}^d)$, there exists a unique bounded linear functional $\Lambda\in(L^{r'}(\mathbb{R}^d))^*$ such that

$$\|\Lambda\|_{(L^{r'}(\mathbb{R}^d))^*}\lesssim\Theta,\qquad\Lambda(f)=\int bf\quad\forall f\in L_0^\infty(\mathbb{R}^d).$$

By the Riesz representation theorem, such a $\Lambda\in(L^{r'}(\mathbb{R}^d))^*$ is represented by a unique function $a\in L^r(\mathbb{R}^d)$ of the same norm, and hence

$$\|a\|_r\lesssim\Theta,\qquad\int af=\int bf\quad\forall f\in L_0^\infty(\mathbb{R}^d).$$

Let $\Delta := b - a \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$. We have $\int \Delta \cdot f = 0$ for all $f \in L^\infty_0(\mathbb{R}^d)$. Taking $f = t^{-d}(1_{B(x,t)} - 1_{B(y,t)})$ and letting $t \to 0$, we deduce that $\Delta(x) = \Delta(y)$ for all Lebesgue points x and y of Δ. Thus $\Delta(x) \equiv c$ is a constant, and $b = a + c$ with $\|a\|_r \lesssim \Theta$, as claimed. □

3. Applications to the Jacobian operator

We now discuss applications of the previous methods towards the problem of finding an unknown function u with the prescribed Jacobian

$$Ju = \det \nabla u = \det(\partial_i u_j)^d_{i,j=1} = f.$$

The Jacobian equation has been quite extensively studied in the form of a Dirichlet boundary value problem in a bounded, sufficiently smooth domain $\Omega \subset \mathbb{R}^d$,

$$\begin{cases} Ju = f & \text{in } \Omega, \\ u = g & \text{on } \partial\Omega. \end{cases} \tag{3.1}$$

There are several works dealing with datum f in Hölder [8] or Sobolev spaces [36]; in a different direction, a recent result [22, Theorem 6.3] addresses $f \in L^p(\Omega)$ with $p \in (\frac{1}{d}, 1)$.

Our interest is in the conjecture of Iwaniec [19] discussed in Section 1.3; besides being set on the full space $\mathbb{R}^d$, it deals with datum f in the spaces $L^p(\mathbb{R}^d)$, $p \in (1, \infty)$, which fall in some sense “between” the higher regularity classes considered by [8,36], and the sub-integrability classes in [22]. The closest analogue of our results in the existing literature is the Hardy space $H^1(\mathbb{R}^d)$ results of Coifman, Lions, Meyer and Semmes [6].

3.1. Norming properties of Jacobians

We prove that the norm of a function b in various function spaces can be computed by dualising against functions in the range of the Jacobian operator. The following lemma, a variant of considerations used in [6, p. 263], already gives a flavour of such results:

Lemma 3.1.1. *Let $b \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$. For each $q \in (1, \infty)$ we have*

$$\fint_Q |b - \langle b\rangle_Q| \lesssim \sup\Big\{\Big|\fint_{2Q} bJ(u)\Big| : u \in C_c^\infty(2Q)^d, \fint_{2Q} |\nabla u|^q \le 1\Big\}.$$

Proof. We can find $g \in L^\infty_0(Q')$, supported in a slightly smaller cube $Q' = (1 - \delta)Q$, and with $\|g\|_\infty \le 1$ such that

$$\fint_Q |b - \langle b\rangle_Q| \lesssim \Big|\fint_Q (b - \langle b\rangle_Q)g\Big| = \Big|\fint_Q bg\Big|. \tag{3.2}$$

Now $g \in L^q_0(Q')$ for every $q \in (1, \infty)$. By [32, Lemma II.2.1.1], we can find at least one $v \in W^{1,q}_0(Q')^d$ (Sobolev space with zero boundary values) satisfying

$$\operatorname{div} v = g, \qquad \|\nabla v\|_q \lesssim \|g\|_q.$$

In fact, [32, Lemma II.2.1.1] proves this with an unspecified dependence on the cube (or more generally, a Lipschitz domain) Q'; we apply this in the unit cube Q_0 first, and then obtain the stated estimate in an arbitrary cube by a change of variables. So we have

$$\fint_Q |b-\langle b\rangle_Q| \lesssim \Big|\fint_Q b\operatorname{div} v\Big|, \qquad v\in W_0^{1,q}(Q')^d, \qquad \fint_Q |\nabla v|^q \lesssim 1.$$

If we now replace v by a standard mollification $\phi_\varepsilon * v$ and note that $\nabla(\phi_\varepsilon * v) = \phi_\varepsilon * \nabla v$, we observe that the above display remains valid for small enough $\varepsilon > 0$, except that $v \in W_0^{1,q}(Q')^d$ is replaced by $\phi_\varepsilon * v \in C_c^\infty(Q)^d$. We now proceed with this replacement, writing $w = \phi_\varepsilon * v$.

Next, at least one of the integrals $\int_Q b\partial_k w_k$, $k = 1,\ldots,d$, has to be at least as big as their average $d^{-1}\fint_Q b\operatorname{div} w$, so in fact

$$\fint_Q |b-\langle b\rangle_Q| \lesssim \Big|\fint_Q b\partial_k w_k\Big|, \qquad w_k\in C_c^\infty(Q), \qquad \fint_Q |\nabla w_k|^q \lesssim 1.$$

We now define a vector-valued function $u = (u_i)_{i=1}^d \in C_c^\infty(2Q)$ as follows. For $i = k$, let $u_k = w_k$. For all $i \neq k$, let $u_i(x) = (x_i - c_i)\varphi_Q(x)$, where c is the centre of Q, we write x_i (resp. c_i) for the ith component of x (resp. c), and $\varphi_Q \in C_c^\infty(2Q)$ is a usual bump such that $1_Q \le \varphi_Q \le 1_{2Q}$ and $|\nabla\varphi_Q| \lesssim 1/\ell(Q)$. Then

$$\nabla u_i(x) = e_i\varphi_Q(x) + (x_i - c_i)\nabla\varphi_Q(x), \qquad |\nabla u_i(x)| \lesssim 1_{2Q}(x),$$

where e_i is the ith coordinate vector. Thus $\fint_{2Q} |\nabla u_i|^q \lesssim 1$ for $i \neq k$, and we already knew this for $i = k$. Since $u_k = w_k$ is compactly supported inside Q, so is $J(u)$, and for $x \in Q$, we simply have $\nabla u_i(x) = e_i$ for $i \neq k$. Hence

$$J(u)(x) = \det(\partial_i u_j(x))_{i,j=1}^d = \sum_{\sigma\in S_d} \operatorname{sgn}(\sigma)\prod_{i\neq k}\delta_{i,\sigma(i)} \times \partial_{\sigma(k)} w_k(x) = \partial_k w_k(x),$$

since only the identity permutation gives a contribution. We have shown that

$$\fint_Q |b-\langle b\rangle_Q| \lesssim \Big|\fint_Q b\partial_k w_k\Big| = \Big|\fint_Q bJ(u)\Big|,$$

for a certain $u \in C_c^\infty(2Q)^d$ such that $\fint_{2Q} |\nabla u|^q \lesssim 1$, and this proves the lemma. □

For the passage from the local estimate of Lemma 3.1.1 to global function space norms, we need two further lemmas that have nothing to do with the Jacobian, and will also be used in the next section.

Lemma 3.1.2. *Let $b \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$ and let $Q_0 \subset \mathbb{R}^d$ be a cube. Then there is collection $\mathscr{Q}$ of dyadic subcubes of Q_0 such that, at almost every $x \in Q_0$,*

$$1_{Q_0}(x)|b-\langle b\rangle_{Q_0}| \lesssim \sum_{Q\in\mathscr{Q}} 1_Q(x)\fint_Q |b-\langle b\rangle_Q|,$$

and $\mathscr{Q}$ is sparse in the sense that each $Q \in \mathscr{Q}$ has a major subset $E(Q)$ such that $|E(Q)| \ge \frac{1}{2}|Q|$ and the subsets $E(Q)$ are pairwise disjoint.

Proof. This is a more elementary variant of Lerner's oscillation formula [25]; we recall the idea of the proof. For any disjoint subcubes Q_j^1 of Q_0, we have

$$\begin{aligned} 1_{Q_0}(b-\langle b\rangle_{Q_0}) &= 1_{Q_0\setminus\bigcup_j Q_j^1}(b-\langle b\rangle_{Q_0}) \\ &\quad + \sum_j 1_{Q_j^1}(\langle b\rangle_{Q_j^1} - \langle b\rangle_{Q_0}) + \sum_j 1_{Q_j^1}(b-\langle b\rangle_{Q_j^1}). \end{aligned} \tag{3.3}$$

If the Q_j^1 are chosen to be the maximal dyadic subcubes $Q \subset Q_0$ such that

$$\fint_Q |b - \langle b \rangle_{Q_0}| > 2 \fint_{Q_0} |b - \langle b \rangle_{Q_0}|,$$

then $\sum_j |Q_j^1| \leq \frac{1}{2}|Q_0|$ so that $E(Q_0) = Q_0 \setminus \bigcup_j Q_j^1$ qualifies for a major subset. Moreover, the sum of the first two terms on the right of (3.3) is dominated by $1_{Q_0} \fint_{Q_0} |b - \langle b \rangle_{Q_0}|$ and the last term is a sum over disjointly supported terms of the same form as where we started, and we can iterate. $\square$

We borrow the following observation from [9, Remark 2.4]:

Lemma 3.1.3. *Suppose that $b \in L^r_{\mathrm{loc}}(\mathbb{R}^d)$, $r \in [1, \infty)$, satisfies*

$$\|b - \langle b \rangle_Q\|_{L^r(Q)} \leq \Theta$$

for every cube $Q \subset \mathbb{R}^d$. Then $b = a + c$, where c is a constant, $a \in L^r(\mathbb{R}^d)$, and

$$\|a\|_{L^r(\mathbb{R}^d)} \leq \Theta.$$

Proof. Let us consider a sequence of cubes $Q_0 \subset Q_1 \subset \ldots$ with $\bigcup_{n=0}^\infty Q_n = \mathbb{R}^d$. For $m \leq n$, we have

$$|\langle b \rangle_{Q_n} - \langle b \rangle_{Q_m}| = |(b(x) - \langle b \rangle_{Q_m}) - (b(x) - \langle b \rangle_{Q_n})|$$

and hence, taking the L^r average over $x \in Q_m$,

$$\begin{aligned} |\langle b \rangle_{Q_n} - \langle b \rangle_{Q_m}| &\leq |Q_m|^{-1/r}(\|b - \langle b \rangle_{Q_m}\|_{L^r(Q_m)} + \|b - \langle b \rangle_{Q_n}\|_{L^r(Q_m)}) \\ &\leq |Q_m|^{-1/r}(\Theta + \|b - \langle b \rangle_{Q_n}\|_{L^r(Q_n)}) \leq 2\Theta |Q_m|^{-1/r}. \end{aligned}$$

Thus $(\langle b \rangle_{Q_n})_{n=0}^\infty$ is a Cauchy sequence, and hence converges to some c. We conclude by Fatou's lemma that

$$\int_{\mathbb{R}^d} |b - c|^r = \int_{\mathbb{R}^d} \lim_{n\to\infty} 1_{Q_n} |b - \langle b \rangle_{Q_n}|^r \leq \liminf_{n\to\infty} \int_{Q_n} |b - \langle b \rangle_{Q_n}|^r \leq \Theta^r. \quad \square$$

We are now ready for the main result of this section:

Theorem 3.1.4. *Let $b \in L^1_{\mathrm{loc}}(\mathbb{R}^d)$, let $r_i \in (1, \infty)$ for $i = 1, \ldots, d$, and $\frac{1}{r} := \sum_{i=1}^d \frac{1}{r_i}$. Then*

$$\Gamma := \sup \Big\{ \Big| \int bJ(u) \Big| : u = (u_i)_{i=1}^d \in C_c^\infty(\mathbb{R}^d)^d, \|\nabla u_i\|_{r_i} \leq 1 \ \forall i = 1, \ldots, d \Big\} \tag{3.4}$$

is finite, if and only if

- $r = 1$ *and* $b \in \mathrm{BMO}(\mathbb{R}^d)$, *or*
- $r \in [\frac{d}{d+1}, 1)$ *and b is $d(\frac{1}{r} - 1)$-Hölder continuous, or*
- $r < \frac{d}{d+1}$ *and b is constant, or*
- $r > 1$ *and $b = a + c$, where c is constant and $a \in L^{r'}(\mathbb{R}^d)$.*

Moreover, in each case the respective function space norm is comparable to (3.4).

Proof. Let us first consider the "if" directions. The constant cases follow from the fact that $\int J(u) = 0$, and it is immediate from Hölder's inequality that

$$\Big| \int bJ(u) \Big| \lesssim \|b\|_{r'} \prod_{i=1}^{d} \|\nabla u_i\|_{r_i}, \qquad r > 1.$$

We then deal with $r \in [\frac{d}{d+1}, 1]$. Let us first check that there is at least one k such that $1/r - 1/r_k < 1$. Suppose for contradiction that we have $1/r - 1/r_k \geq 1$ for all $k = 1, \ldots, d$. Summing over k, this gives $d/r - 1/r \geq d$, and thus $r \leq \frac{d-1}{d}$. But we are also assuming that $\frac{d}{d+1} \leq r$, thus $d^2 \leq (d-1)(d+1) = d^2 - 1$, a contradiction. Without loss of generality, we may assume that $k = 1$, thus

$$\frac{1}{s} := \sum_{i=2}^{d} \frac{1}{r_i} = \frac{1}{r} - \frac{1}{r_1} \in (0, 1)$$

so that $s \in (1, \infty)$. We can then write

$$J(u) = \nabla u_1 \cdot \sigma = Rf \cdot \sigma,$$

where $\sigma \in C_c^\infty(\mathbb{R}^d)^d$ satisfies $\operatorname{div} \sigma = 0$ and

$$\|\sigma\|_s \lesssim \prod_{i=2}^{d} \|\nabla u_i\|_{r_i} \leq 1,$$

and $R = (R_i)_{i=1}^d = \nabla(-\Delta)^{-1/2}$ is the vector of the Riesz transforms, and finally $f = (-\Delta)^{1/2} u_1$ satisfies $\|f\|_{r_1} \lesssim \|\nabla u_1\|_{r_1} \leq 1$. Then

$$-\int R(bf) \cdot \sigma = \int bf(R \cdot \sigma) = \int bf(-\Delta)^{-1/2} \operatorname{div} \sigma = 0,$$

and thus

$$\Big| \int bJ(u) \Big| = \Big| \int bRf \cdot \sigma \Big| = \Big| \int [b, R]f \cdot \sigma \Big| \leq \|[b, R]\|_{L^{r_1} \to L^{s'}} \|f\|_{r_1} \|\sigma\|_s.$$

The last two norms are bounded by one, and $1/s' = 1 - 1/r + 1/r_1$, so that

$$\Gamma \leq \|[b, R]\|_{L^{r_1} \to L^{s'}} \lesssim \begin{cases} \|b\|_{\mathrm{BMO}}, & \text{if } r = 1, \\ \|b\|_{\dot{C}^{0,\alpha}}, & \text{if } \alpha := d(\frac{1}{r_1} - \frac{1}{s'}) = d(\frac{1}{r} - 1) \in (0, 1] \end{cases}$$

by Theorem 1.0.1.

We turn to the "only if" parts of the theorem. Recall the definition of Γ from (3.4). We apply Lemma 3.1.1 with some $q > \max_{i=1,\ldots,d} r_i$. If $Q \subset \mathbb{R}^d$ is any cube, then for some $u \in C_c^\infty(2Q)^d$ with $\fint_{2Q} |\nabla u|^q \leq 1$ we have

$$\begin{aligned}\fint_Q |b-\langle b\rangle_Q| &\lesssim \Big|\fint_{2Q} bJ(u)\Big| \lesssim \frac{\Gamma}{|Q|}\prod_{i=1}^d \|\nabla u_i\|_{r_i} \\ &\lesssim \Gamma\frac{|Q|^{1/r}}{|Q|}\prod_{i=1}^d \Big(\fint_{2Q}|\nabla u_i|^{r_i}\Big)^{1/r_i} \\ &\lesssim \Gamma\ell(Q)^{d(1/r-1)}\prod_{i=1}^d\Big(\fint_{2Q}|\nabla u_i|^q\Big)^{1/q} \leq \Gamma\ell(Q)^{d(1/r-1)}.\end{aligned} \tag{3.5}$$

If $r=1$, this is precisely the condition that $\|b\|_{\mathrm{BMO}} \lesssim \Gamma$. For $r<1$, the conclusion follows as in the proof of Theorem 2.4.1.

Let us then consider $r>1$. Let $Q_0 \subset \mathbb{R}^d$ be an arbitrary cube. We apply Lemma 3.1.2 and monotone convergence to see that

$$\|b-\langle b\rangle_{Q_0}\|_{L^{r'}(Q_0)} \lesssim \lim_{N\to\infty}\Big\|\sum_{k=1}^N 1_{Q_k}\fint_{Q_k}|b-\langle b\rangle_{Q_k}|\Big\|_{L^{r'}(Q_0)}$$

where $\{Q_k\}_{k=1}^\infty$ is an enumeration of the collection $\mathcal{Q}$ given by Lemma 3.1.2.

We then dualise with some $\|\phi\|_r \leq 1$, and apply just the first step of (3.5) to each Q_k in place of Q. Note that this produces a possibly different $u^k=(u_i^k)_{i=1}^d \in C_c^\infty(2Q_k)^d$ for each k. Thus we end up estimating

$$\begin{aligned}\sum_{k=1}^N\int_{Q_k}\phi\fint_{Q_k}|b-\langle b\rangle_{Q_k}| &\lesssim \sum_{k=1}^N\int_{Q_k}\phi\fint_{2Q_k} bJ(u^k) \\ &\lesssim \int b\sum_{k=1}^N J(u^k)\lambda_k, \qquad \lambda_k := \fint_{Q_k}\phi.\end{aligned}$$

In order to proceed, we make a randomisation trick. Due to the d-linear nature of the Jacobian, we invoke a sequence $(\zeta_k)_{k=1}^N$ of independent random dth roots of unity, i.e. the ζ_k's are independent random variables on some probability space, distributed so that $\mathbb{P}(\zeta_k = e^{i2\pi a/d}) = 1/d$ for each $a=0,1,\ldots,d-1$. The case $d=2$ thus corresponds to the familiar random signs. The important feature of these random variables is that, denoting by $\mathbb{E}$ the expectation,

$$\mathbb{E}\prod_{j=1}^d \zeta_{k_j} = \begin{cases} 1, & \text{if } k_1=\ldots=k_d, \\ 0, & \text{else.}\end{cases} \tag{$*$}$$

Indeed, if $k_1=\ldots=k_d=k$, then $\prod_{j=1}^d\zeta_{k_j}=\zeta_k^d\equiv 1$, so also its expectation is equal to 1. Otherwise, we have $\prod_{j=1}^d\zeta_{k_j}=\prod_{j=1}^r\zeta_{m_j}^{n_j}$ for some *distinct* values $m_1,\ldots,m_r\in\{1,\ldots,N\}$ and exponents $n_1,\ldots,n_r\in\{1,\ldots,d-1\}$. By independence, it then follows that

$$\mathbb{E}\prod_{j=1}^d\zeta_{k_j} = \mathbb{E}\prod_{j=1}^r\zeta_{m_j}^{n_j} = \prod_{j=1}^r\mathbb{E}\zeta_{m_j}^{n_j},$$

where each factor satisfies

$$\mathbb{E}\zeta_{m_j}^{n_j} = \frac{1}{d}\sum_{a=0}^{d-1}e^{i2\pi a n_j/d} = 0,$$

noting that $e^{i2\pi n_j/d} \neq 1$ since $0 < n_j < d$.

Using $(*)$, we can now continue the computation from above with

$$\begin{aligned}\int b\sum_{k=1}^N J(u^k)\lambda_k &\overset{(*)}{=} \mathbb{E}\int bJ(\sum_{k_1=1}^N \varepsilon_{k_1}\lambda_{k_1}^{r/r_1}u_1^{k_1},\cdots,\sum_{k_d=1}^N \varepsilon_{k_d}\lambda_{k_d}^{r/r_d}u_d^{k_d}) \\ &\leq \Gamma\mathbb{E}\prod_{i=1}^d \Big\|\sum_{k=1}^N \varepsilon_k\lambda_k^{r/r_i}\nabla u_i^k\Big\|_{r_i} \\ &\leq \Gamma\prod_{i=1}^d \Big\|\sum_{k=1}^N \lambda_k^{r/r_i}|\nabla u_i^k|\Big\|_{r_i}.\end{aligned}$$

To estimate each L^{r_i} norm above, we dualise with $\|\psi\|_{r_i'} \leq 1$. Recalling that $u_i^k \in C_c^\infty(2Q_k)$ satisfies the bound for u in Lemma 3.1.1, and using the definition of λ_k and the disjoint major subsets $E(Q_k)$ from Lemma 3.1.2, we have

$$\begin{aligned}\int \psi\sum_{k=1}^N \lambda_k^{r/r_i}|\nabla u_i^k| &= \sum_{k=1}^N \lambda_k^{r/r_i}\int_{2Q_k}|\nabla u_i^k|\psi \\ &\lesssim \sum_{k=1}^N \lambda_k^{r/r_i}|Q_k|\Big(\fint_{2Q_k}|\nabla u_i^k|^q\Big)^{1/q}\Big(\fint_{2Q_k}\psi^{q'}\Big)^{1/q'} \\ &\leq \sum_{k=1}^N \Big(\fint_{Q_k}\phi\Big)^{r/r_i}|Q_k|\Big(\fint_{2Q_k}\psi^{q'}\Big)^{1/q'} \\ &\lesssim \sum_{k=1}^N |E(Q_k)|\Big(\inf_{x\in Q_k}M\phi(x)\Big)^{r/r_i}\Big(\inf_{y\in Q_k}M(\psi^{q'})(y)\Big)^{1/q'} \\ &\leq \int (M\phi)^{r/r_i}(M(\psi^{q'}))^{1/q'} \leq \|(M\phi)^{r/r_i}\|_{r_i}\|(M(\psi^{q'}))^{1/q'}\|_{r_i'} \\ &= \|M\phi\|_r^{r/r_i}\|M(\psi^{q'})\|_{r_i'/q'}^{1/q'} \lesssim \|\phi\|_r^{r/r_i}\|\psi\|_{r_i'} \leq 1,\end{aligned}$$

by the boundedness of the maximal operator and the choice of $q > r_i$ so that $r_i' > q'$.

Substituting back, we have checked that

$$\|b - \langle b\rangle_{Q_0}\|_{L^{r'}(Q_0)} \lesssim \Gamma$$

for an arbitrary cube Q_0; by Lemma 3.1.3, this completes the proof of the theorem in the remaining case that $r > 1$. $\square$

Remark 3.1.5. The "if" parts of the cases $r \in (\frac{d}{d+1}, 1]$ of Theorem 3.1.4 could also be deduced from a result of [6, cf. Theorem II.3], which says that $J(u)$ belongs to the Hardy space $H^r(\mathbb{R}^d)$ under the same assumptions, together with the H^1-BMO duality when $r = 1$ or the H^r-$\dot{C}^{0,d(1/r-1)}$-duality for $r \in (\frac{d}{d+1}, 1)$. However, a separate argument would be required for the end-point $r = \frac{d}{d+1}$ any way: in fact, $J : C_c^\infty(\mathbb{R}^d) \not\to H^{d/(d+1)}(\mathbb{R}^d)$, since $J(u)$ fails, in general, to satisfy the required moment conditions $\int x_i a = 0$ of an $H^{d/(d+1)}$-atom a. This follows e.g. from the proof of Lemma 3.1.1, which contains the observation that any $\partial_k w$, with $w \in C_c^\infty(\mathbb{R}^d)$, can arise as the Jacobian $J(u)$ of a suitable $u \in C_c^\infty(\mathbb{R}^d)^d$. However, we have $\int x_k\partial_k w = -\int w\partial_k x_k = -\int w$, which can easily be nonzero. The departure from the Hardy-Hölder duality is also reflected by the fact that the condition for b in Theorem 3.1.4 corresponding to $r = \frac{d}{d+1}$ is the usual Lipschitz-continuity, $|b(x) - b(y)| \lesssim |x - y|$, and not the Zygmund class condition arising from the Hardy space duality.

On the other hand, one can also give a different proof of the "if" part of Theorem 3.1.4 in this special case $r = \frac{d}{d+1}$. Using the notation from the previous proof, where $1 > \frac{1}{s} = \frac{1}{r} - \frac{1}{r_1} = 1 + \frac{1}{d} - \frac{1}{r_1}$, we find that $r_1 \in (1, d)$. Writing, as before, $J(u) = \nabla u_1 \cdot \sigma$, we have

$$\int bJ(u) = \int b\nabla u_1 \cdot \sigma = -\int u_1 \operatorname{div}(b\sigma) = -\int u_1(\nabla b) \cdot \sigma,$$

since $\operatorname{div} \sigma = 0$. But then we can estimate

$$\Big| \int u_1(\nabla b) \cdot \sigma \Big| \leq \|u_1\|_{s'} \|\nabla b\|_\infty \|\sigma\|_s,$$

where $\|\nabla b\|_\infty$ is bounded by the Lipschitz constant, $\|\sigma\|_s \leq 1$, and

$$\frac{1}{s'} = 1 - \frac{1}{s} = 1 - \Big(\frac{1}{r} - \frac{1}{r_1}\Big) = \frac{1}{r_1} - \frac{1}{d},$$

so that $s' = r_1 d/(d - r_1) = r_1^*$ is the Sobolev exponent. Thus

$$\|u_1\|_{s'} = \|u_1\|_{r_1^*} \lesssim \|\nabla u_1\|_{r_1} \leq 1,$$

by Sobolev's inequality, and this completes the alternative proof.

3.2. The linear span of Jacobians

Here we will obtain the following consequence of Theorem 3.1.4:

Theorem 3.2.1. *Let $d \geq 2$ and $p \in [1, \infty)$. Then*

$$\Big\{ \sum_{j=1}^\infty Ju^j : u^j \in \dot{W}^{1,pd}(\mathbb{R}^d)^d, \ \sum_{j=1}^\infty \|\nabla u^j\|^d_{L^{pd}(\mathbb{R}^d)^{d\times d}} < \infty \Big\} = \begin{cases} L^p(\mathbb{R}^d), & p \in (1,\infty), \\ H^1(\mathbb{R}^d), & p = 1. \end{cases}$$

In fact, each $f \in L^p(\mathbb{R}^d)$ (resp. $f \in H^1(\mathbb{R}^d)$) admits a representation

$$f = \sum_{j=1}^\infty Ju^j, \quad \sum_{j=1}^\infty \|\nabla u^j\|^d_{L^{pd}(\mathbb{R}^d)^{d\times d}} \lesssim \begin{cases} \|f\|_{L^p(\mathbb{R}^d)}, & p \in (1,\infty), \\ \|f\|_{H^1(\mathbb{R}^d)}, & p = 1, \end{cases}$$

where each $u^j \in C_c^\infty(\mathbb{R}^d)^d$.

The power d in the series is related to the d-homogeneity of the Jacobian, so that $\|\nabla u^j\|^d_{L^{pd}(\mathbb{R}^d)^{d\times d}}$ is (up to constant) an upper bound for $\|Ju^j\|_{L^p(\mathbb{R}^d)}$ or $\|Ju^j\|_{H^1(\mathbb{R}^d)}$ for $p = 1$. The case $p = 1$ is already due to Coifman et al. [6]; they explicitly formulate a similar result [6, Theorem III.2] for the "div-curl example" but point out that "this type of answer applies also to other examples like the Jacobian". Our proof of the full Theorem 3.2.1 depends on the same functional analytic lemma as used in [6] for the case $p = 1$. The formulation below combines [6, Lemmas III.1, III.2] and is taken from [28]. We recall the short proof for the sake of recording a precise quantitative relation between the equivalent qualitative conditions:

Lemma 3.2.2. *Let $V \subset \bar{B}_X(0,1)$ be a symmetric subset of the unit-ball of a Banach space X. Then the following conditions are equivalent:*

1. *There is $\alpha > 0$ such that $\sup_{x\in V} |\langle \lambda, x\rangle| \geq \alpha \|\lambda\|_{X^*}$ for all $\lambda \in X^*$.*
2. *The closed convex hull $\overline{\mathrm{conv}}(V)$ contains a ball $\bar{B}_X(0,\beta)$ of radius $\beta > 0$.*
3. *The s-*convex hull

$$s(V) := \Big\{ \sum_{j=1}^{\infty} \lambda_j x_j : x_j \in V, \lambda_j \geq 0, \sum_{j=1}^{\infty} \lambda_j = 1 \Big\}$$

contains an open ball $B_X(0,\gamma)$ of radius $\gamma > 0$.

Moreover, the largest admissible values of α, β, γ satisfy $\alpha = \beta = \gamma$.

Proof. If $\lambda \in X^*$ and $\beta' < \beta$, we can find $x_0 \in \bar{B}_X(0,\beta)$ such that $\beta'\|\lambda\|_{X^*} \leq |\langle\lambda, x_0\rangle|$. Writing $x_0 = \lim_n x_n$, where $x_n \in \mathrm{conv}(V)$, we easily check that $\sup_{x\in V} |\langle \lambda, x\rangle| \geq \beta'\|\lambda\|_{X^*}$, and hence $\alpha \geq \beta$. On the other hand, if $y_0 \notin \overline{\mathrm{conv}}(V)$, then by the Hahn–Banach theorem there exists $\lambda \in X^*$ such that $\mathrm{Re}\langle\lambda, x\rangle \leq \eta < \mathrm{Re}\langle\lambda, y_0\rangle$ for some $\eta \in \mathbb{R}$ and all $x \in \overline{\mathrm{conv}}(V)$, in particular for $x \in V$, and thus, by the symmetry of V, also $|\langle\lambda, x\rangle| \leq \eta < |\langle\lambda, y_0\rangle| \leq \|\lambda\|_{X^*}\|y_0\|_X$ for all $x \in V$. Taking the supremum over $x \in V$ and using (1) it follows that $\alpha\|\lambda\|_{X^*} \leq \eta < \|\lambda\|_{X^*}\|y_0\|_X$. Since clearly $\lambda \neq 0$, it follows that $\|y_0\|_X > \alpha$, and thus $\beta \geq \alpha$.

Clearly $s(V) \subset \overline{\mathrm{conv}}(V)$, and hence $B_X(0,\gamma) \subset s(V)$ implies $\bar{B}_X(0,\gamma) \subset \overline{\mathrm{conv}}(V)$ so that $\beta \geq \gamma$. On the other hand, suppose that $x \in \bar{B}(0,\beta) \subset \overline{\mathrm{conv}}(V)$. Fix $\varepsilon > 0$. Suppose that we have already found $x_k \in \mathrm{conv}(V)$ such that

$$\|x - \sum_{k=0}^{n-1} \varepsilon^k x_k\|_X \leq \varepsilon^n \beta \tag{3.6}$$

(this is vacuous for $n = 0$). Then $\varepsilon^{-n}(x - \sum_{k=0}^{n-1} \varepsilon^k x_k) \in \bar{B}_X(0,\beta) \subset \overline{\mathrm{conv}}(V)$, and thus we can pick $x_n \in \mathrm{conv}(V)$ with $\|\varepsilon^{-n}(x - \sum_{k=0}^{n-1} \varepsilon^k x_k) - x_n\|_X \leq \varepsilon\beta$. But this is the same as (3.6) with $n+1$ in place of n. By induction it follows that $x = \sum_{k=0}^{\infty} \varepsilon_k x_k$ with $x_k \in \mathrm{conv}(V)$. Since $\sum_{k=0}^{\infty} \varepsilon^k = (1-\varepsilon)^{-1}$, this means that $(1-\varepsilon)x \in s(V)$. As $x \in \bar{B}_X(0,\beta)$ and $\varepsilon > 0$ were arbitrary, we have $\bar{B}_X(0,(1-\varepsilon)\beta) \subset s(V)$, and hence $B_X(0,\beta) \subset s(V)$. Thus $\gamma \geq \beta$. □

Proof of Theorem 3.2.1. We apply Lemma 3.2.2 with $X = L^p(\mathbb{R}^d)$ if $p \in (1,\infty)$, or $X = H^1(\mathbb{R}^d)$ if $p = 1$. In either case, let

$$V = \{Ju : u \in C_c^\infty(\mathbb{R}^d)^d, \|\nabla u\|_{L^{pd}(\mathbb{R}^d)^{d\times d}} \leq 1\}.$$

It is immediate that V is symmetric, and that $V \subset \bar{B}_X(0,1)$ if $p > 1$. For $p = 1$, this last inclusion is nontrivial but well known from [6, Theorem II.1].

The assertion of Theorem 3.2.1 is clearly the same as (3) of Lemma 3.2.2 for these choices of X and V. By Lemma 3.2.2, it hence suffices to verify (1) of the same lemma, i.e., that

$$\|b\|_{X^*} \lesssim \sup\Big\{\Big|\int bJ(u)\Big| : u \in C_c^\infty(\mathbb{R}^d)^d : \|\nabla u\|_{L^p(\mathbb{R}^d)^{d\times d}} \leq 1\Big\},$$

$$\forall b \in X^* = \begin{cases} L^{p'}(\mathbb{R}^d), & p \in (1,\infty), \\ \mathrm{BMO}(\mathbb{R}^d), & p = 1. \end{cases}$$

But this is precisely the statement of Theorem 3.1.4 for $r = p \in [1, \infty)$ and $r_1 = \ldots = r_d = pd$. The a priori condition that $b \in L^{p'}(\mathbb{R}^d)$ guarantees that the additive constant present in Theorem 3.1.4 for $r > 1$ does not appear here. $\square$

Remark 3.2.3.

1. Lindberg [28, Lemma 3.1] shows that another equivalent condition in Lemma 3.2.2 is that $\bigcup_{n=1}^{\infty} n \cdot s(V)$ has second category in X. Hence, if any of these conditions fails, then $\bigcup_{n=1}^{\infty} n \cdot s(V)$ has first category in X. Lindberg uses this to show [28, Theorems 1.2, 7.4] that the set
$$\Big\{ \sum_{j=1}^{\infty} Ju^j : u^j \in W^{1,pd}(\mathbb{R}^d)^d, \sum_{j=1}^{\infty} \big(\|u^j\|_{L^{pd}(\mathbb{R}^d)^d} + \|\nabla u^j\|_{L^{pd}(\mathbb{R}^d)^{d\times d}} \big)^d < \infty \Big\}$$
has first category in $L^p(\mathbb{R}^d)$ if $p \in (1, \infty)$, or in $H^1(\mathbb{R}^d)$ if $p = 1$.
2. Lindberg [28, p. 739] also sketches how to deduce the special case $d = 2$ of Theorems 3.1.4 and 3.2.1 from the special case of (then unknown) Theorem 1.0.1, where T is the Ahlfors–Beurling operator. Since a more general result is proved above by working directly with the Jacobian, we do not repeat his argument here, but the interested reader may consult the companion paper [18] for this approach. Nevertheless, the strategy proposed by Lindberg was an important motivation for the discovery of our present results.

4. Higher order real commutators and the median method

In this section we establish the following variant of Theorem 1.0.1. In one direction, it generalises Theorem 1.0.1 by allowing iterated commutators of arbitrary order, but in another direction it imposes a more restrictive assumption by requiring the pointwise multiplier b to be real-valued. This restriction arises from the proof using the so-called median method, which takes explicit advantage of the order structure of the real line. We note, however, that this restriction is imposed on b only; the kernel K of T may still be complex-valued.

Theorem 4.0.1. *Let $1 < p, q < \infty$, let T be a non-degenerate Calderón–Zygmund operator on $\mathbb{R}^d$, and let $k \in \{1, 2, \ldots\}$ and $b \in L^k_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R})$. Then the k times iterated commutator*
$$T_b^k := [b, T_b^{k-1}], \qquad T_b^1 := [b, T],$$
defines a bounded operator $T_b^k : L^p(\mathbb{R}^d) \to L^q(\mathbb{R}^d)$ if and only if:

- *$p = q$ and b has bounded mean oscillation, or*
- *$p < q \le p_k^* = \Big(\dfrac{1}{p} - \dfrac{k}{d}\Big)_+^{-1}$ and b is $\alpha = \dfrac{d}{k}\Big(\dfrac{1}{p} - \dfrac{1}{q}\Big)$-Hölder continuous, or*
- *$q > p_k^*$ and b is constant, or*
- *$p > q$ and $b = a + c$, where $a \in L^{rk}(\mathbb{R}^d)$ for $\dfrac{1}{r} = \dfrac{1}{q} - \dfrac{1}{p}$, and c is constant.*

As in the case of Theorem 1.0.1, all the “if” statements are either classical (such as the case $p = q$ that goes back to Coifman, Rochberg and Weiss [7]) or straightforward; this applies to the remaining cases, which may be handled by easy extensions of the arguments sketched for $k = 1$ in Section 1.1. (There is also a variant of the $p < q$ case of Theorem 4.0.1 due to Paluszyński, Taibleson and Weiss [30], but for $k > 1$, it deals with operators that are related to, but not exactly the same as, the iterated commutators T_b^k that we study. This leads to a slightly different result.)

As before, our principal task is to prove the "only if" directions.

4.1. Basic estimates of the median method

We will not give a formal definition of the "median method", but the reason for this nomenclature should be fairly apparent from the considerations that follow. The broad philosophy of this method should be attributed to Lerner, Ombrosi and Rivera-Ríos [26], but we fine-tune some of its details in such a way as to be able, in particular, to answer a problem that was raised but left open in [26, Remark 4.1].

The simplest form of the median method is contained in the following lemma. Under a quantitative positivity assumption on the kernel (which may nevertheless be complex-valued!), it needs no additional "Calderón–Zygmund" structure.

Lemma 4.1.1. *Let $b \in L^k_{\mathrm{loc}}(\mathbb{R}^d;\mathbb{R})$. Suppose that, for some disjoint balls $B, \tilde{B}$ of equal radius r, we have*

$$\operatorname{Re}(\sigma K(x,y)) \gtrsim \frac{1}{|B|} \qquad \text{for all } x \in \tilde{B}, y \in B \tag{4.1}$$

for some $|\sigma| = 1$. If T has kernel K, then

$$\inf_c \int_B |b(y) - c|^k \, \mathrm{d}y \lesssim \sum_{i=1}^{2} |\langle 1_{\tilde{E}_i}, T_b^k 1_{E_i}\rangle|$$

for some subsets $E_i \subset B$, $\tilde{E}_i \subset \tilde{B}$.

Remark 4.1.2. If K is a non-degenerate two-variable Calderón–Zygmund kernel (Definition 2.1.1(1)), then for all large enough A and for every ball $B = B(y_0, r)$ there exists another ball $\tilde{B} = B(x_0, r)$ with $|x_0 - y_0| \eqsim Ar$, where the assumptions, and hence the conclusions, of Lemma 4.1.1 are satisfied.

Indeed, by Proposition 2.2.1 and case (1) of its proof, we can find an x_0 with $|x_0 - y_0| \eqsim Ar$ such that

$$|K(x_0,y_0)| \eqsim \frac{1}{(Ar)^d}, \qquad |K(x,y) - K(x_0,y_0)| \leq \frac{\varepsilon_A}{(Ar)^d} \quad \forall x \in B(x_0,r),\ \forall y \in B(y_0,r).$$

Hence, for suitable σ, we have

$$\operatorname{Re}(\sigma K(x,y)) \geq |K(x_0,y_0)| - |K(x,y) - K(x_0,y_0)| \eqsim \frac{1}{(Ar)^d}.$$

Proof of Lemma 4.1.1. The basic observation is that, if $\alpha \in \mathbb{R}$ and $x \in \tilde{B} \cap \{b \leq \alpha\}$, then

$$\begin{aligned} \int_B (b(y) - \alpha)_+^k \, \mathrm{d}y &\leq \int_{B \cap \{b \geq \alpha\}} (b(y) - b(x))^k \, \mathrm{d}y \\ &\lesssim (-1)^k |B| \operatorname{Re}\Big(\sigma \int_{B \cap \{b \geq \alpha\}} (b(x) - b(y))^k K(x,y) \, \mathrm{d}y\Big), \end{aligned}$$

and hence

$$\begin{aligned}
&|\tilde{B} \cap \{b \leq \alpha\}| \int\limits_{B} (b(y)-\alpha)_+^k \,\mathrm{d}y \\
&\lesssim |B| \Big| \int\limits_{\tilde{B}\cap\{b\leq\alpha\}} \int\limits_{B\cap\{b\geq\alpha\}} (b(x)-b(y))^k K(x,y)\,\mathrm{d}y\,\mathrm{d}x \Big| \\
&= |B| \Big| \langle 1_{\tilde{B}\cap\{b\leq\alpha\}}, T_b^k 1_{B\cap\{b\geq\alpha\}} \rangle \Big| =: |B| \cdot |\langle 1_{\tilde{E}_1}, T_b^k 1_{E_1}\rangle|.
\end{aligned}$$

In a completely analogous way, integrating over $x \in \tilde{B} \cap \{b \geq \alpha\}$, we also prove that

$$|\tilde{B} \cap \{b \geq \alpha\}| \int\limits_{B} (\alpha - b(y))_+^k \,\mathrm{d}y \lesssim |B| \Big| \langle 1_{\tilde{B}\cap\{b\geq\alpha\}}, T_b^k 1_{B\cap\{b\leq\alpha\}} \rangle \Big| = |B| \cdot |\langle 1_{\tilde{E}_2}, T_b^k 1_{E_2}\rangle|.$$

Choosing α as a median of b on $\tilde{B}$, we have

$$\min(|\tilde{B} \cap \{b \leq \alpha\}|, |\tilde{B} \cap \{b \geq \alpha\}|) \geq \frac{1}{2}|\tilde{B}| = \frac{1}{2}|B|,$$

and hence

$$\int\limits_{B} |b(y)-\alpha|^k \,\mathrm{d}y = \int\limits_{B} (b(y)-\alpha)_+^k \,\mathrm{d}y + \int\limits_{B} (\alpha - b(y))_+^k \,\mathrm{d}y \lesssim \sum_{i=1}^{2} |\langle 1_{\tilde{E}_i}, T_b^k 1_{E_i}\rangle|. \quad \square$$

We present a variant of the result for rough homogeneous kernels. While the conclusion is essentially identical, the proof requires an additional iteration of the basic argument.

Lemma 4.1.3. *Let $k \in \{1, 2, \dots\}$, let $\Omega \in L^1(S^{d-1}) \setminus \{0\}$ and $K(x) = \dfrac{\Omega(x/|x|)}{|x|^d}$. Let $\theta_0 \in S^{d-1}$ be a Lebesgue point of K, where $K(\theta_0) = \Omega(\theta_0) \neq 0$. Let T be an operator with kernel $K(x, y) = K(x - y)$.*

Then there is a (large) constant A, depending only on the above data, such that every $b \in L^k_{\mathrm{loc}}(\mathbb{R}^d)$ satisfies the following estimate for every ball B:

$$\inf_c \int\limits_{B} |b(y) - c|^k \,\mathrm{d}y \lesssim \sum_{i=1}^{4} |\langle 1_{\tilde{E}_i}, T_b^k 1_{E_i}\rangle|$$

for some subsets $E_i \subset B$ and $\tilde{E}_i \subset \tilde{B} := B + Ar_B\theta_0$.

Proof. Given $B = B(y_0, r)$, let $x_0 = y_0 + Ar\theta_0$, where the large A is yet to be chosen, and $\tilde{B} = B(x_0, r)$.

The basic observation is that, if $b(x) \leq \alpha$, then

$$\begin{aligned}
&\int\limits_{B} (b(y)-\alpha)_+^k \,\mathrm{d}y \leq \int\limits_{B\cap\{b\geq\alpha\}} (b(y)-b(x))^k \,\mathrm{d}y \\
&= \int\limits_{B\cap\{b\geq\alpha\}} (b(y)-b(x))^k \frac{(Ar)^d K(x_0-y_0)}{\Omega(\theta_0)} \,\mathrm{d}y \\
&= \frac{(-1)^k c_d |B|}{\Omega(\theta_0)} A^d \int\limits_{B\cap\{b\geq\alpha\}} (b(x)-b(y))^k [K(x-y) + K(x_0-y_0) - K(x-y)] \,\mathrm{d}y
\end{aligned}$$

Hence, taking α as the median of b on $\tilde{B}$, we have

$$
\begin{aligned}
\int_B (b(y)-\alpha)_+^k \, \mathrm{d}y &\le 2\frac{|\tilde{B}\cap\{b\le\alpha\}|}{|\tilde{B}|}\int_B (b(y)-\alpha)_+^k \, \mathrm{d}y \\
&\lesssim A^d |\langle 1_{\tilde{B}\cap\{b\le\alpha\}}, T_b^k 1_{B\cap\{b\ge\alpha\}}\rangle| + \\
&\quad + A^d \int_{\tilde{B}\cap\{b\le\alpha\}} \int_{B\cap\{b\ge\alpha\}} (b(y)-b(x))^k |K(x-y)-K(x_0-y_0)| \, \mathrm{d}y \, \mathrm{d}x.
\end{aligned}
$$

Estimating

$$
(b(y)-b(x))^k = (b(y)-\alpha+\alpha-b(x))^k \le c_k(b(y)-\alpha)^k + c_k(\alpha-b(x))^k,
$$

the double integral can be dominated by the sum of

$$
\int_{B\cap\{b\ge\alpha\}} (b(y)-\alpha)^k \Big(\int_{\tilde{B}\cap\{b\le\alpha\}} |K(x-y)-K(x_0-y_0)| \, \mathrm{d}x \Big) \, \mathrm{d}y
$$

and

$$
\int_{\tilde{B}\cap\{b\le\alpha\}} (\alpha-b(x))^k \Big(\int_{B\cap\{b\ge\alpha\}} |K(x-y)-K(x_0-y_0)| \, \mathrm{d}y \Big) \, \mathrm{d}x.
$$

Writing $x-y = x_0-y_0+(x-x_0)-(y-y_0)$, both inner integrals are seen to be bounded by

$$
\begin{aligned}
&\int_{B(0,2r)} |K(x_0-y_0+z)-K(x_0-y_0)| \, \mathrm{d}z \\
&\qquad = \int_{B(0,2r)} |K(Ar\theta_0+z)-K(Ar\theta_0)| \, \mathrm{d}z \\
&\qquad = (Ar)^{-d} \int_{B(0,2r)} |K(\theta_0+\frac{z}{Ar})-K(\theta_0)| \, \mathrm{d}z \\
&\qquad = \int_{B(0,2/A)} |K(\theta_0+u)-K(\theta_0)| \, \mathrm{d}u = \varepsilon_A A^{-d},
\end{aligned}
$$

where $\varepsilon_A \to 0$ as $A \to \infty$, by the assumption that θ_0 is a Lebesgue point of K.

Substituting back, and observing in particular the cancellation of the factors A^d and A^{-d} in the double integral, we have proved that

$$
\begin{aligned}
\int_B (b(y)-\alpha)_+^k \, \mathrm{d}y &\le c_A |\langle 1_{\tilde{B}\cap\{b\le\alpha\}}, T_b^k 1_{B\cap\{b\ge\alpha\}}\rangle| \\
&\qquad + c\varepsilon_A \int_B (b(y)-\alpha)_+^k \, \mathrm{d}y + c\varepsilon_A \int_{\tilde{B}} (\alpha-b(x))_+^k \, \mathrm{d}x,
\end{aligned}
$$

and hence

$$(1-c\varepsilon_A)\int_B (b(y)-\alpha)_+^k \,\mathrm{d}y \le c_A|\langle 1_{\tilde{B}\cap\{b\le\alpha\}}, T_b^k 1_{B\cap\{b\ge\alpha\}}\rangle| + c\varepsilon_A \int_{\tilde{B}} (\alpha-b(x))_+^k \,\mathrm{d}x.$$

Replacing (b,α) by $(-b,-\alpha)$, we also have

$$(1-c\varepsilon_A)\int_B (\alpha-b(y))_+^k \,\mathrm{d}y \le c_A|\langle 1_{\tilde{B}\cap\{b\ge\alpha\}}, T_b^k 1_{B\cap\{b\le\alpha\}}\rangle| + c\varepsilon_A \int_{\tilde{B}} (b(x)-\alpha)_+^k \,\mathrm{d}x,$$

and adding the two estimates,

$$(1-c\varepsilon_A)\int_B |b(y)-\alpha|^k \,\mathrm{d}y \le c_A \sum_{i=1}^2 |\langle 1_{\tilde{E}_i}, T_b^k 1_{E_i}\rangle| + c\varepsilon_A \int_{\tilde{B}} |b(x)-\alpha|^k \,\mathrm{d}x,$$

where $E_i \subset B$ and $\tilde{E}_i \subset \tilde{B}$ for $i=1,2$. Recall that α was the median of b on $\tilde{B}$, but since this choice of α is a quasi-minimiser for the integral on the right, we also deduce the more symmetric version

$$\inf_{\alpha\in\mathbb{R}} \int_B |b(y)-\alpha|^k \,\mathrm{d}y \le c\sum_{i=1}^2 |\langle 1_{\tilde{E}_i}, T_b^k 1_{E_i}\rangle| + \frac{1}{2}\inf_{\alpha\in\mathbb{R}} \int_{\tilde{B}} |b(x)-\alpha|^k \,\mathrm{d}x, \tag{4.2}$$

where we have also fixed an A so that $c\varepsilon_A/(1-c\varepsilon_A)\le 1/2$.

We now apply the same argument to the adjoint

$$(T_b^k)^* = (-1)^k (T^*)_b^k.$$

We note that the kernel K^* of T^* is related to the kernel K of T given by $K^*(x,y)=K(y,x)$, and hence it is also a homogeneous kernel with symbol $\Omega^*(\theta)=\Omega(-\theta)$. In particular, the point $-\theta_0$ plays the same role for T^* as θ_0 plays for T, and thus the ball $B=\tilde{B}-Ar\theta_0$ plays the same role for $\tilde{B}$ and T^* as $\tilde{B}$ plays for B and T.

This means that the analogue of (4.2) in the adjoint case reads as

$$\inf_{\alpha\in\mathbb{R}} \int_{\tilde{B}} |b(x)-\alpha|^k \,\mathrm{d}x \le c\sum_{i=3}^4 |\langle 1_{\tilde{E}_i}, T_b^k 1_{E_i}\rangle| + \frac{1}{2}\inf_{\alpha\in\mathbb{R}} \int_{B} |b(y)-\alpha|^k \,\mathrm{d}y,$$

where again $E_i \subset B$ and $\tilde{E}_i \subset \tilde{B}$ for $i=3,4$. Using (4.2) and its adjoint version above consecutively, we have

$$
\begin{aligned}
\inf_{\alpha\in\mathbb{R}} \int_B |b(y)-\alpha|^k \,\mathrm{d}y &\le c\sum_{i=1}^{2} |\langle 1_{\tilde{E}_i}, T_b^k 1_{E_i}\rangle| + \frac{1}{2}\inf_{\alpha\in\mathbb{R}} \int_{\tilde{B}} |b(x)-\alpha|^k \,\mathrm{d}x \\
&\le c\sum_{i=1}^{2} |\langle 1_{\tilde{E}_i}, T_b^k 1_{E_i}\rangle| + \frac{1}{2}\Big(c\sum_{i=3}^{4} |\langle 1_{\tilde{E}_i}, T_b^k 1_{E_i}\rangle| + \frac{1}{2}\inf_{\alpha\in\mathbb{R}} \int_B |b(y)-\alpha|^k \,\mathrm{d}y\Big) \\
&\le c\sum_{i=1}^{4} |\langle 1_{\tilde{E}_i}, T_b^k 1_{E_i}\rangle| + \frac{1}{4}\inf_{\alpha\in\mathbb{R}} \int_B |b(y)-\alpha|^k \,\mathrm{d}y,
\end{aligned}
$$

which implies that

$$
\inf_{\alpha\in\mathbb{R}} \int_B |b(y)-\alpha|^k \,\mathrm{d}y \lesssim \sum_{i=1}^{4} |\langle 1_{\tilde{E}_i}, T_b^k 1_{E_i}\rangle|
$$

as claimed. $\square$

4.2. The lower bound for higher commutators

We restate and then prove the "only if" parts of Theorem 4.0.1 in the two theorems below, dealing with the cases $p \le q$ and $p > q$, in analogy with Theorems 2.4.1 and 2.5.1.

Theorem 4.2.1. *Let K be a non-degenerate Calderón–Zygmund kernel, let $k \in \{1, 2, \dots\}$ and $b \in L^k_{\mathrm{loc}}(\mathbb{R}^d)$. Let further*

$$
1 < p \le q < \infty, \qquad \alpha := \frac{d}{k}\Big(\frac{1}{p} - \frac{1}{q}\Big) \ge 0,
$$

and suppose that T_b^k satisfies the following weak form of $L^p \to L^q$ boundedness:

$$
\begin{aligned}
|\langle T_b^k f, g\rangle| &= \Big| \iint (b(x)-b(y))^k K(x,y) f(y) g(x) \,\mathrm{d}y \,\mathrm{d}x \Big| \\
&\le \Theta \cdot \|f\|_\infty |B|^{1/p} \cdot \|g\|_\infty |\tilde{B}|^{1/q'},
\end{aligned} \tag{4.3}
$$

whenever $f \in L^\infty(B)$, $g \in L^\infty(\tilde{B})$ for any two balls of equal radius r and distance $\operatorname{dist}(B, \tilde{B}) \gtrsim r$. Then

- *if $\alpha = 0$, equivalently $p = q$, we have $b \in \mathrm{BMO}(\mathbb{R}^d)$, and $\|b\|^k_{\mathrm{BMO}} \lesssim \Theta$;*
- *if $\alpha \in (0, 1]$, we have $b \in \dot{C}^{0,\alpha}(\mathbb{R}^d)$, and $\|b\|^k_{\dot{C}^{0,\alpha}} \lesssim \Theta$;*
- *if $\alpha > 0$, the function b is constant, so in fact $T_b^k = 0$.*

Proof. Consider a ball B of radius r. From Lemma 4.1.1 and Remark 4.1.2 or Lemma 4.1.3 (depending whether K is a two-variable or rough homogeneous kernel), it is immediate that

$$
\Big(\fint_B |b - \langle b\rangle_B|\Big)^k \le \fint_B |b - \langle b\rangle_B|^k \lesssim \sum_{i=1}^{4} \frac{|\langle 1_{\tilde{E}_j}, T_b^k 1_{E_j}\rangle|}{|B|}, \tag{4.4}
$$

for some subsets $E_j \subset B$, $\tilde{E}_j \subset \tilde{B}$, where $\tilde{B}$ is a ball of the same radius r and $\operatorname{dist}(B, \tilde{B}) \gtrsim r$. By assumption (4.3), it follows that

$$
|\langle 1_{\tilde{E}_j}, T_b^k 1_{E_j}\rangle| \le \Theta |B|^{1/p} |B|^{1/q'} = \Theta |B|^{1/p - 1/q + 1} = \Theta |B| r^{\alpha k}
$$

and hence

$$\fint_B |b - \langle b\rangle_B| \lesssim \Theta^{1/k} r^\alpha.$$

From this the rest follows as in the proof of Theorem 2.4.1. □

Theorem 4.2.2. *Let K be a non-degenerate Calderón–Zygmund kernel, let $k \in \{1, 2, \ldots\}$, and $b \in L^k_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R})$. Let*

$$1 < q < p < \infty, \qquad r = \frac{pq}{p-q} \in (1, \infty),$$

and suppose that T_b^k satisfies the following weak form of $L^p \to L^q$ boundedness:

$$\sum_{i=1}^N |\langle T_b^k f_i, g_i\rangle| \leq \Theta \Big\| \sum_{i=1}^N \|f_i\|_\infty 1_{Q_i} \Big\|_p \Big\| \sum_{i=1}^N \|g_i\|_\infty 1_{\tilde{Q}_i} \Big\|_{q'}, \tag{4.5}$$

whenever, for each $i = 1, \ldots, N$, we have $f_i \in L^\infty(Q_i)$ and $g_i \in L^\infty(\tilde{Q}_i)$ for cubes Q_i and $\tilde{Q}_i$ such that $\operatorname{dist}(Q_i, \tilde{Q}_i) \gtrsim \operatorname{diam}(Q_i) = \operatorname{diam}(\tilde{Q}_i)$.

Then $b = a + c$ for some $a \in L^{rk}(\mathbb{R}^d)$ and some constant $c \in \mathbb{C}$, where $\|a\|_{rk} \lesssim \Theta$.

Proof. Let us fix some (large) cube $Q_0 \subset \mathbb{R}^d$. We apply Lemma 3.1.2 to find that

$$1_{Q_0}|b - \langle b\rangle_{Q_0}| \lesssim \sum_{Q\in\mathscr{Q}} 1_Q \fint_Q |b - \langle b\rangle_Q| \leq \sum_{j=1}^\infty 1_{Q_j}\Big(\fint_{Q_j} |b - \langle b\rangle_{Q_j}|^k\Big)^{1/k},$$

where we also introduced an enumeration of the sparse collection $\mathscr{Q}$ of dyadic subcubes of Q_0 given by Lemma 3.1.2.

By Lemma 4.1.1 and Remark 4.1.2, or Lemma 4.1.3 in the rough homogeneous case, we have

$$\Big(\fint_{Q_j} |b - \langle b\rangle_{Q_j}|^k\Big)^{1/k} \lesssim \sum_{i=1}^4 \Big(\frac{|\langle 1_{\tilde{E}_j^i}, T_b^k 1_{E_j^i}\rangle|}{|Q_j|}\Big)^{1/k},$$

for some subsets $E_j^i \subset Q_j$ and $\tilde{E}_j^i \subset \tilde{Q}_j$, where the cube $\tilde{Q}_j$ satisfies $\operatorname{dist}(Q_j, \tilde{Q}_j) \gtrsim \operatorname{diam}(Q_j) = \operatorname{diam}(\tilde{Q}_j)$. Hence

$$\|b - \langle b\rangle_{Q_0}\|_{L^{kr}(Q_0)} \lesssim \sup\Big\{ \sum_{i=1}^4 \sum_{j=1}^\infty \Big(\frac{|\langle 1_{\tilde{E}_j^i}, T_b^k 1_{E_j^i}\rangle|}{|Q_j|}\Big)^{1/k} \int_{Q_j} \phi : \|\phi\|_{L^{(kr)'}} \leq 1\Big\}.$$

It is enough to give a uniform bound for the finite sums

$$\begin{aligned}
&\sum_{j=1}^N \Big(\frac{|\langle 1_{\tilde{E}_j^i}, T_b^k 1_{E_j^i}\rangle|}{|Q_j|}\Big)^{1/k} |Q_j| \fint_{Q_j} \phi \\
&\leq \Big(\sum_{j=1}^N \Big(\frac{|\langle 1_{\tilde{E}_j^i}, T_b^k 1_{E_j^i}\rangle|}{|Q_j|}\Big)^r |Q_j|\Big)^{1/(kr)} \Big(\sum_{j=1}^N |Q_j| \Big[\fint_{Q_j} \phi\Big]^{(kr)'}\Big)^{1/(kr)'}.
\end{aligned}$$

Using the sparseness of $\mathcal{Q} \supset \{Q_j\}_{j=1}^N$, we can bound the second factor by

$$\sum_{j=1}^N |Q_j| \Big[\fint_{Q_j} \phi \Big]^{(kr)'} \lesssim \sum_{j=1}^N |E(Q_j)| \inf_{z \in Q_j} M\phi(z)^{(kr)'} \leq \int (M\phi)^{(kr)'} \lesssim \int \phi^{(kr)'} \leq 1.$$

We dualise the first factor with $\sum_{j=1}^N \lambda_j^{r'} |Q_j| \leq 1$ to end up considering

$$\begin{aligned} \sum_{j=1}^N \frac{|\langle 1_{\tilde{E}_j^i}, T_b^k 1_{E_j^i} \rangle|}{|Q_j|} \lambda_j |Q_j| &= \sum_{j=1}^N |\langle \lambda_j^{r'/q} 1_{\tilde{E}_j^i}, T_b^k (\lambda_j^{r'/p} 1_{E_j^i}) \rangle| \\ &\leq \Theta \Big\| \sum_{j=1}^N \lambda_j^{r'/q'} 1_{\tilde{Q}_j} \Big\|_{q'} \Big\| \sum_{j=1}^N \lambda_j^{r'/p} 1_{Q_j} \Big\|_p, \end{aligned} \tag{4.6}$$

where we used the assumption (4.5) in the last step.

By Lemma 2.5.4, we have, using the disjoint major subsets $E(Q_j) \subset Q_j$,

$$\Big\| \sum_{j=1}^N \lambda_j^{r'/p} 1_{Q_j} \Big\|_p \lesssim \Big\| \sum_{j=1}^N \lambda_j^{r'/p} 1_{E(Q_j)} \Big\|_p = \Big(\sum_{j=1}^N \lambda_j^{r'} |E(Q_j)| \Big)^{1/p} \leq 1.$$

For the first factor on the right of (4.6), we obtain a similar bound by starting with

$$\Big\| \sum_{j=1}^N \lambda_j^{r'/q'} 1_{\tilde{Q}_j} \Big\|_{q'} \lesssim \Big\| \sum_{j=1}^N \lambda_j^{r'/q'} 1_{Q_j} \Big\|_{q'},$$

which also follows from Lemma 2.5.4, and then finishing as before.

We have now proved that

$$\|b - \langle b \rangle_{Q_0}\|_{L^{kr}(Q_0)} \lesssim \Theta^{1/k}$$

for any cube $Q_0 \subset \mathbb{R}^d$. This shows in particular that $b \in L^{kr}_{\mathrm{loc}}(\mathbb{R}^d)$, and we conclude by Lemma 3.1.3. □

4.3. Two-weight norm inequalities of Bloom type

We finally discuss the boundedness of commutators between weighted L^p spaces with weights from the Muckenhoupt class

$$A_p(\mathbb{R}^d) = \Big\{ w \in L^1_{\mathrm{loc}}(\mathbb{R}^d) : w > 0 \text{ a.e.}, \ [w]_{A_p} := \sup_B \fint_B w \Big(\fint_B w^{-\frac{1}{p-1}} \Big)^{p-1} < \infty \Big\},$$

where the supremum is over all balls $B \subset \mathbb{R}^d$. We consider $p \in (1, \infty)$ fixed throughout this discussion, and denote by $w' := w^{-\frac{1}{p-1}}$ the *dual weight*. One checks that $w \in A_p$ if and only if $w' \in A_{p'}$. The space $L^{p'}(w')$ is the dual of $L^p(w)$ with respect to the *unweighted* duality $\langle f, g \rangle = \int fg$. We will identify a weight and its induced measure, using notation like $w(Q) := \int_Q w$.

We will be concerned with the boundedness of

$$T_b^k : L^p(\mu) \to L^p(\lambda), \qquad \mu, \lambda \in A_p(\mathbb{R}^d),$$

i.e., we allow two different weights on the domain and the target space, but (in contrast to the rest of the paper) we restrict the Lebesgue exponents to $p = q \in (1, \infty)$. This fits with the line of investigation that was started by Bloom [3] and that has been recently revived by Holmes, Lacey, and Wick [15], followed by several others as we shortly recall. Here we complete the following picture:

Theorem 4.3.1. *Let T be a non-degenerate Calderón–Zygmund operator, let $k \in \{1, 2, \ldots\}$ and $b \in L^k_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R})$. Let further $p \in (1, \infty)$ and $\lambda, \mu \in A_p(\mathbb{R}^d)$. Then T^k_b defines a bounded operator $L^p(\mu) \to L^p(\lambda)$ if and only if*

$$\|b\|_{\mathrm{BMO}(\nu^{1/k})} := \sup_B \frac{1}{\nu^{1/k}(B)} \int_B |b - \langle b\rangle_B| < \infty, \qquad \nu := (\mu/\lambda)^{1/p}.$$

The first version of Theorem 4.3.1, when $k = d = 1$ and T is the Hilbert transform, is due to Bloom [3]. Still for first order commutators ($k = 1$) but in arbitrary dimension $d \geq 1$, Holmes, Lacey, and Wick [15] proved the "if" part of Theorem 4.3.1 for all standard Calderón–Zygmund operators, and the "only if" part assuming the boundedness of each of the d Riesz transforms R_i, $i = 1, \ldots, d$, thus extending the exact scope (in terms of operators) of the classical Coifman–Rochberg–Weiss theorem [7] to the two-weight setting. The first two-weight result for iterated commutators was achieved in the "if" direction by Holmes and Wick [16] (with a simplified proof in [17]): they obtained the boundedness of $T^k_b : L^p(\mu) \to L^p(\lambda)$ for any $k \geq 1$ under the stronger condition that $b \in \mathrm{BMO}(\nu) \cap \mathrm{BMO}(\mathbb{R}^d) \subset \mathrm{BMO}(\nu^{1/k})$. (For the inclusion, which in general is strict, see [26, Lemma 4.7].) Finally, Lerner, Ombrosi, and Rivera-Ríos [26] obtained Theorem 4.3.1 almost as stated: they identified the correct BMO space with the weight $\nu^{1/k}$ depending on the order k of the commutator, and they proved the "if" part of Theorem 4.3.1 for all standard Calderón–Zygmund operators and the "only if" part for all homogeneous Calderón–Zygmund operators with the fairly general local positivity assumption discussed in Section 1.2. For us, it remains to prove this "only if" part assuming non-degeneracy only, and more precisely we prove:

Theorem 4.3.2. *Let K be a non-degenerate Calderón–Zygmund kernel, let $k \in \{1, 2, \ldots\}$ and $b \in L^k_{\mathrm{loc}}(\mathbb{R}^d; \mathbb{R})$. Let $p \in (1, \infty)$, let $\lambda, \mu \in A_p(\mathbb{R}^d)$, and suppose that T^k_b satisfies the following weak form of $L^p(\mu) \to L^p(\lambda)$ boundedness:*

$$\begin{aligned} |\langle T^k_b f, g\rangle| &= \Big| \iint (b(x) - b(y))^k K(x, y) f(y) g(x) \,\mathrm{d}y \,\mathrm{d}x \Big| \\ &\leq \Theta \cdot \|f\|_\infty \mu(B)^{1/p} \cdot \|g\|_\infty \lambda'(\tilde{B})^{1/p'}, \end{aligned} \tag{4.7}$$

whenever $f \in L^\infty(B)$, $g \in L^\infty(\tilde{B})$ for any two balls of equal radius r and distance $\mathrm{dist}(B, \tilde{B}) \gtrsim r$. Then $b \in \mathrm{BMO}(\nu^{1/k})$, where $\nu = (\mu/\lambda)^{1/p}$, and more precisely

$$\|b\|_{\mathrm{BMO}(\nu^{1/k})} \lesssim \Theta^{1/k}.$$

Let us first observe that (4.7) is indeed a weak form of the boundedness of $T^k_b : L^p(\mu) \to L^p(\lambda)$: if this boundedness holds, then

$$\begin{aligned} |\langle T^k_b f, g\rangle| &\leq \|T^k_b f\|_{L^p(\lambda)} \|g\|_{L^{p'}(\lambda')} \leq \|T^k_b\|_{L^p(\mu) \to L^p(\lambda)} \|f\|_{L^p(\mu)} \|g\|_{L^{p'}(\lambda')} \\ &\leq \|T^k_b\|_{L^p(\mu) \to L^p(\lambda)} \cdot \|f\|_\infty \mu(B)^{1/p} \cdot \|g\|_\infty \lambda'(\tilde{B})^{1/p'}, \end{aligned}$$

and thus $\Theta \leq \|T^k_b\|_{L^p(\mu) \to L^p(\lambda)}$.

Turning to the proof of Theorem 4.3.2, we need a simple lemma, which is the only place where the A_p condition is used.

Lemma 4.3.3. *Let $\lambda, \mu \in A_p$. If $B, \tilde{B}$ are balls of equal radius r with $\operatorname{dist}(B, \tilde{B}) \lesssim r$, then*

$$\mu(B)^{1/p}\lambda'(\tilde{B})^{1/p'} \lesssim \langle \nu^{1/k}\rangle_B^k \cdot |B|$$

for all $k = 1, 2, \ldots$, where $\nu = (\mu/\lambda)^{1/p}$.

Proof. We recall that all A_p weights, and then also $\lambda' \in A_{p'}$, are doubling. Hence $\lambda'(\tilde{B}) \lesssim \lambda'(B)$. We then use the A_p property of both μ and ν directly via the definition (together with some basic algebra involving p and p') to see that

$$\frac{\mu(B)^{1/p}\lambda'(\tilde{B})^{1/p'}}{|B|} \lesssim \langle \mu\rangle_B^{1/p}\langle \lambda'\rangle_B^{1/p'} \lesssim \frac{1}{\langle \mu'\rangle_B^{1/p'}\langle \lambda\rangle_B^{1/p}} \leq \frac{1}{\langle (\mu')^{1/p'}\lambda^{1/p}\rangle_B} = \frac{1}{\langle \nu^{-1}\rangle_B}$$

Finally,

$$1 = \langle \nu^{-\frac{1}{k+1}}\nu^{\frac{1}{k}\cdot\frac{k}{k+1}}\rangle_B \leq \langle \nu^{-1}\rangle_B^{\frac{1}{k+1}}\langle \nu^{\frac{1}{k}}\rangle_B^{\frac{k}{k+1}},$$

and hence $\langle \nu^{-1}\rangle_B^{-1} \leq \langle \nu^{1/k}\rangle_B^k$. $\square$

Proof of Theorem 4.3.2. As in the proof of the unweighted version in Theorem 4.2.1, we have (just copying (4.4) from the said proof)

$$\Big(\fint_B |b - \langle b\rangle_B|\Big)^k \leq \fint_B |b - \langle b\rangle_B|^k \lesssim \sum_{i=1}^4 \frac{|\langle 1_{\tilde{E}_j}, T_b^k 1_{E_j}\rangle|}{|B|},$$

for some subsets $E_j \subset B$, $\tilde{E}_j \subset \tilde{B}$, where $\tilde{B}$ is a ball of the same radius r and $\operatorname{dist}(B, \tilde{B}) \eqsim r$.

By assumption (4.7) and Lemma 4.3.3, we have

$$\frac{|\langle 1_{\tilde{E}_j}, T_b^k 1_{E_j}\rangle|}{|B|} \leq \Theta\frac{\mu(B)^{1/p}\lambda'(\tilde{B})^{1/p'}}{|B|} \lesssim \Theta\langle \nu^{1/k}\rangle_B^k.$$

Hence

$$\Big(\fint_B |b - \langle b\rangle_B|\Big)^k \lesssim \Theta\langle \nu^{1/k}\rangle_B^k,$$

which simplifies to $\|b\|_{\mathrm{BMO}(\nu^{1/k})} \lesssim \Theta^{1/k}$. $\square$

Declaration of competing interest

None.

Acknowledgements

I would like to thank Sauli Lindberg for bringing the problem about the Jacobian operator and its connection to commutators to my attention, and for pointing out some oversights in an earlier version of the manuscript. I would also like to thank Riikka Korte for discussions on the theme of the paper, and the anonymous referee for constructive suggestions on the presentation.

References

[1] H. Aimar, L. Forzani, F.J. Martín-Reyes, On weighted inequalities for singular integrals, Proc. Am. Math. Soc. 125 (7) (1997) 2057–2064.
[2] E. Airta, T. Hytönen, K. Li, H. Martikainen, T. Oikari, Off-diagonal estimates for bi-commutators, arXiv:2005.03548.
[3] S. Bloom, A commutator theorem and weighted BMO, Trans. Am. Math. Soc. 292 (1) (1985) 103–122, https://doi.org/10.2307/2000172.
[4] L. Chaffee, Characterizations of bounded mean oscillation through commutators of bilinear singular integral operators, Proc. R. Soc. Edinb., Sect. A 146 (6) (2016) 1159–1166.
[5] L. Chaffee, D. Cruz-Uribe, Necessary conditions for the boundedness of linear and bilinear commutators on Banach function spaces, Math. Inequal. Appl. 21 (1) (2018) 1–16.
[6] R. Coifman, P.-L. Lions, Y. Meyer, S. Semmes, Compensated compactness and Hardy spaces, J. Math. Pures Appl. (9) 72 (3) (1993) 247–286.
[7] R.R. Coifman, R. Rochberg, G. Weiss, Factorization theorems for Hardy spaces in several variables, Ann. Math. (2) 103 (3) (1976) 611–635.
[8] B. Dacorogna, J. Moser, On a partial differential equation involving the Jacobian determinant, Ann. Inst. Henri Poincaré, Anal. Non Linéaire 7 (1) (1990) 1–26, https://doi.org/10.1016/S0294-1449(16)30307-9.
[9] G. Dafni, T. Hytönen, R. Korte, H. Yue, The space JN_p: nontriviality and duality, J. Funct. Anal. 275 (3) (2018) 577–603, https://doi.org/10.1016/j.jfa.2018.05.007.
[10] X.T. Duong, H.-Q. Li, J. Li, B.D. Wick, Lower bound of Riesz transform kernels and commutator theorems on stratified nilpotent Lie groups, J. Math. Pures Appl. (9) 124 (2019) 273–299, https://doi.org/10.1016/j.matpur.2018.06.012.
[11] S.H. Ferguson, M.T. Lacey, A characterization of product BMO by commutators, Acta Math. 189 (2) (2002) 143–160.
[12] J. García-Cuerva, J.L. Rubio de Francia, Weighted Norm Inequalities and Related Topics, North-Holland Mathematics Studies, vol. 116 Notas de Matemática, vol. 104, North-Holland Publishing Co., Amsterdam, 1985.
[13] W. Guo, J. Lian, H. Wu, The unified theory for the necessity of bounded commutators and applications, J. Geom. Anal. 30 (4) (2020) 3995–4035.
[14] T.S. Hänninen, T.P. Hytönen, K. Li, Two-weight L^p-L^q bounds for positive dyadic operators: unified approach to $p \leq q$ and $p > q$, Potential Anal. 45 (3) (2016) 579–608.
[15] I. Holmes, M.T. Lacey, B.D. Wick, Commutators in the two-weight setting, Math. Ann. 367 (1–2) (2017) 51–80, https://doi.org/10.1007/s00208-016-1378-1.
[16] I. Holmes, B.D. Wick, Two weight inequalities for iterated commutators with Calderón-Zygmund operators, J. Oper. Theory 79 (1) (2018) 33–54.
[17] T.P. Hytönen, The Holmes-Wick theorem on two-weight bounds for higher order commutators revisited, Arch. Math. (Basel) 107 (4) (2016) 389–395, https://doi.org/10.1007/s00013-016-0956-5.
[18] T.P. Hytönen, Of commutators and Jacobians, arXiv:1905.00814, 2019.
[19] T. Iwaniec, Nonlinear commutators and Jacobians, J. Fourier Anal. Appl. 3 (1997) 775–796.
[20] S. Janson, Mean oscillation and commutators of singular integral operators, Ark. Mat. 16 (2) (1978) 263–270.
[21] F. John, L. Nirenberg, On functions of bounded mean oscillation, Commun. Pure Appl. Math. 14 (1961) 415–426.
[22] K. Koumatos, F. Rindler, E. Wiedemann, Differential inclusions and Young measures involving prescribed Jacobians, SIAM J. Math. Anal. 47 (2) (2015) 1169–1195, https://doi.org/10.1137/140968860.
[23] M.T. Lacey, S. Petermichl, J.C. Pipher, B.D. Wick, Notification of error: multiparameter Riesz commutators, Am. J. Math. 143 (2) (2021) 333–334.
[24] M.T. Lacey, E.T. Sawyer, I. Uriarte-Tuero, Two weight inequalities for discrete positive operators, arXiv:0911.3437, 2010.
[25] A.K. Lerner, A pointwise estimate for the local sharp maximal function with applications to singular integrals, Bull. Lond. Math. Soc. 42 (5) (2010) 843–856, https://doi.org/10.1112/blms/bdq042.
[26] A.K. Lerner, S. Ombrosi, I.P. Rivera-Ríos, Commutators of singular integrals revisited, Bull. Lond. Math. Soc. 51 (1) (2019) 107–119, https://doi.org/10.1112/blms.12216.
[27] C. Liaw, S. Treil, Regularizations of general singular integral operators, Rev. Mat. Iberoam. 29 (1) (2013) 53–74.
[28] S. Lindberg, On the Hardy space theory of compensated compactness quantities, Arch. Ration. Mech. Anal. 224 (2) (2017) 709–742.
[29] T. Oikari, Approximate weak factorizations and bilinear commutators, arXiv:2102.13535.
[30] M. Paluszyński, M. Taibleson, G. Weiss, Characterization of Lipschitz spaces via the commutator operator of Coifman, Rochberg, and Weiss, Rev. Unión Mat. Argent. 37 (1–2) (1991) 142–144 (1992), X Latin American School of Mathematics (Spanish) (Tanti, 1991).
[31] J. Pau, A. Perälä, A Toeplitz-type operator on Hardy spaces in the unit ball, Trans. Am. Math. Soc. 373 (5) (2020) 3031–3062, https://doi.org/10.1090/tran/8053.
[32] H. Sohr, The Navier-Stokes Equations. An Elementary Functional Analytic Approach, Birkhäuser Advanced Texts: Basler Lehrbücher, Birkhäuser Verlag, Basel, 2001.
[33] E.M. Stein, Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals, with the assistance of Timothy S. Murphy, in: Monographs in Harmonic Analysis, III, in: Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, NJ, 1993.
[34] H. Tanaka, A characterization of two-weight trace inequalities for positive dyadic operators in the upper triangle case, Potential Anal. 41 (2) (2014) 487–499.
[35] A. Uchiyama, On the compactness of operators of Hankel type, Tohoku Math. J. (2) 30 (1) (1978) 163–171.
[36] D. Ye, Prescribing the Jacobian determinant in Sobolev spaces, Ann. Inst. Henri Poincaré, Anal. Non Linéaire 11 (3) (1994) 275–296, https://doi.org/10.1016/S0294-1449(16)30185-8.