

Sharp weighted estimates for multilinear commutators

C. Pérez* and R. Trujillo-González†

Journal of the London mathematical society **65** (2002), 672–692.

Abstract

We consider multilinear commutators with vector symbol $\vec{b} = (b_1, \dots, b_m)$ defined by

$$T_{\vec{b}}(f)(x) = \int_{\mathbb{R}^n} \left[\prod_{j=1}^m (b_j(x) - b_j(y)) \right] K(x, y) f(y) dy,$$

where K is a Calderón-Zygmund kernel. We prove the following a priori estimates for $w \in A_\infty$. For $0 < p < \infty$ there exists a constant C such that

$$\|T_{\vec{b}}(f)\|_{L^p(w)} \leq C \|\vec{b}\| \|M_{L(\log L)^{1/r}}(f)\|_{L^p(w)}$$

and

$$\sup_{t>0} \frac{1}{\Phi(\frac{1}{t})} w(\{y \in \mathbb{R}^n : |T_{\vec{b}}f(y)| > t\}) \leq C \sup_{t>0} \frac{1}{\Phi(\frac{1}{t})} w(\{y \in \mathbb{R}^n : M_{L(\log L)^{1/r}}(\|\vec{b}\|f)(y) > t\})$$

where $\|\vec{b}\| = \prod_{j=1}^m \|b_j\|_{osc_{expL^{r_j}}}$, $\Phi(t) = t \log^{1/r}(e+t)$, $\frac{1}{r} = \frac{1}{r_1} + \dots + \frac{1}{r_m}$, and $M_{L(\log L)^\alpha}$ is an Orlicz type maximal operator. This extends, with a different approach, classical results by Coifman [5] (see also [6] [24]).

As a corollary we deduce that the operators $T_{\vec{b}}$ are bounded on $L^p(w)$ when $w \in A_p$ and that they satisfy corresponding weighted $L(\log L)^{1/r}$ type estimates with $w \in A_1$.

*Partially supported by DGESIC Grant PB980106.

†Partially supported by Gobierno de Canarias PI1999/105.

1991 Mathematics Subject Classification: 42B20, 42B25.

Keywords: Calderón-Zygmund singular integral operators, commutators, A_p weights, maximal functions

1 Introduction

The main purpose of this article is to prove sharp estimates for multilinear commutators involving nonstandard symbols. This will be established by means of appropriate maximal operators which somehow control the commutators. This illustrates the classical Calderón–Zygmund principle which roughly states that any singular integral operator is controlled by a suitable maximal operator.

Background

Motivated by the work of Calderón on commutators, Coifman, Rochberg and Weiss introduced in [7] the operator

$$T_b f(x) = \int_{\mathbb{R}^n} (b(x) - b(y))K(x, y)f(y) dy, \quad (1.1)$$

where K is a kernel satisfying the standard Calderón-Zygmund estimates (see section 3.1) and where b , the “symbol” of the operator, is any locally integrable function. The operator is called commutator since $T_b = [b, T] = bT - T(b \cdot)$ where T is the Calderón-Zygmund singular integral operator associated to K . The main result from [7] states that $[b, T]$ is a bounded operator on $L^p(\mathbb{R}^n)$, $1 < p < \infty$, when the symbol b is a *BMO* function.

These commutators have proved to be of interest in many situations. We shall only mention the recent results in the theory of non divergence elliptic equations with discontinuous coefficients [2] [3] [9]. There is also an interesting connection, as pointed out in [21], with the nonlinear commutator considered by R. Rochberg and G. Weiss in [23] and defined by

$$f \rightarrow Nf = T(f \log |f|) - Tf \log |Tf|.$$

This in turn is related to the Jacobian mapping of vector function and with nonlinear P.D.E. as shown in [13] [14].

A natural generalization of the commutator $[b, T]$ is given by

$$T_b^m f(x) = \int_{\mathbb{R}^n} (b(x) - b(y))^m K(x, y)f(y) dy, \quad (1.2)$$

where $m \in \mathbb{N}$. The case $m = 0$ recaptures the Calderón-Zygmund singular integral operator.

Observing that $T_b^m = \overbrace{[b, \cdot, [b, T]]}^{(m \text{ times})}$ we have that for $m \in \mathbb{N}$, T_b^m is also a bounded operator on $L^p(\mathbb{R}^n)$, $1 < p < \infty$, when b is a *BMO* function.

It was shown in [21] that there is an intimate connection between the commutator T_b^m and iterations of the maximal operators. Indeed, the main theorem from [7] was sharpened in [21] as follows: for any $0 < p < \infty$ and any $w \in A_\infty$ there is a constant C such that

$$\int_{\mathbb{R}^n} |T_b^m f(x)|^p w(x) dx \leq C \|b\|_{BMO}^{mp} \int_{\mathbb{R}^n} (M^{m+1} f(x))^p w(x) dx, \quad (1.3)$$

where M^{m+1} denotes the $m + 1$ iterations of the Hardy-Littlewood maximal operator, namely $M^m = \overbrace{M \circ \dots \circ M}^{(m \text{ times})}$. Furthermore, this inequality is sharp since, by the Lebesgue differentiation theorem, M^{m+1} can not be replaced by the smaller operator M^m .

Estimate (1.3) can also be seen as a generalization of a by now classical result of Coifman for Calderón-Zygmund singular integral operators. See [5] and also [6].

There is a weak version of inequality (1.3) obtained in [20]. Indeed, if we let $\Phi_m(t) = t \log^m(e + t)$, for any $w \in A_\infty$ and $b \in BMO$ there is a constant C such that

$$\sup_{t>0} \frac{1}{\Phi_m(\frac{1}{t})} w(\{y \in \mathbb{R}^n : |T_b^m f(y)| > t\}) \leq C \sup_{t>0} \frac{1}{\Phi_m(\frac{1}{t})} w(\{y \in \mathbb{R}^n : M^{m+1} f(y) > t\}). \quad (1.4)$$

A priori inequalities of the form (1.3) (or (1.4)) encode a good amount of information about the behavior of the operator. The first observation is that in any case they reflect the higher degree of singularity of T_b^m , as compared with T , since a larger operator than M , namely M^{m+1} , is needed to balance the inequality. As a second instance, if $p > 1$ we can apply $m + 1$ times Muckenhoupt's theorem to conclude from (1.3) the well known fact that high order commutators are bounded on $L^p(w)$ whenever $w \in A_p$. It should be mentioned that this A_p estimate follows, as it is well known, from an estimate due to Strömberg (see [15, p. 268], [25, p. 417]). This method is powerful and can be applied to different situations. As a sample we remit to [10] for a very nice application to commutators with strongly singular integral operators. However, this estimate of Strömberg is not sharp enough neither to derive (1.3) nor (1.4). In fact it can be shown from (1.4) the following ($L \log L$ type) endpoint estimate first deduced in [20]: let $w \in A_1$ and $b \in BMO$, then there exists a constant C such that for all $\lambda > 0$

$$w(\{y \in \mathbb{R}^n : |T_b^m f(y)| > \lambda\}) \leq C \int_{\mathbb{R}^n} \frac{|f(y)|}{\lambda} (1 + \log^+(\frac{|f(y)|}{\lambda}))^m w(y) dy. \quad (1.5)$$

As a third instance, it was shown in [21] that estimate (1.3) implies a sharp two weighted inequalities for the commutator of the form

$$\int_{\mathbb{R}^n} |T_b^m f(x)|^p w(x) dx \leq C \|b\|_{BMO}^{mp} \int_{\mathbb{R}^n} |f(x)|^p M^{[(m+1)p]+1} w(x) dx,$$

where no condition on the weight w is assumed. This result is an extension of the case $m = 0$ proved in [18] generalizing some previous partial results by M. Wilson [26]. The approach considered in [18] is different from that in [26] and it combines (1.3) with certain sharp two weighted estimates for the Hardy-Littlewood maximal function derived in [19].

Finally, there is a relationship between (1.4) and the endpoint behavior of the operator. Indeed, an interesting observation that occurs in the development of the theory of commutators is that the L^p theory, $p > 1$, was developed without appealing to any endpoint estimate. On the other hand, it is well known that in the classical Calderón-Zygmund L^p theory, $p > 1$, a crucial step is to show that the operators are of weak type $(1, 1)$. However, this is not the case of the commutator $[b, T]$ when $b \in BMO$ as shown in [20] and estimate (1.5) is the right replacement.

This can be seen as another way of expressing the fact that commutators have a higher degree of singularity as compared with the Calderón-Zygmund singular integral operators.

Results of this paper

In this paper we obtain similar estimates for a wider class of commutators. Given a Calderón-Zygmund singular integral operator T with kernel K in $\mathbb{R}^n \times \mathbb{R}^n$ and a vector $\vec{b} = (b_1, \dots, b_m)$ of locally integrable functions, we define the multilinear operator

$$T_{\vec{b}}f(x) = \int_{\mathbb{R}^n} \left[\prod_{j=1}^m (b_j(x) - b_j(y)) \right] K(x, y) f(y) dy, \quad (1.6)$$

generalizing the commutator (1.2). We remit the reader to [6] for an extensive study of multilinear operators.

We will be considering the following class of symbols: for $r \geq 1$ and for any locally integrable function f , we define

$$\|f\|_{osc_{expL^r}} = \sup_Q \|f - f_Q\|_{expL^r, Q}$$

being the supremum taken over all the cubes Q with sides parallels to the axes. Here $\|g\|_{expL^r, Q}$ is the mean value of g on the cube Q with respect to the Young function $\Phi(t) = e^{t^r} - 1$. See Section 2.2 for the precise definition. As usual, f_Q denotes the average of f on Q . For $r \geq 1$ we define the space Osc_{expL^r} by

$$Osc_{expL^r} = \{f \in L^1_{loc}(\mathbb{R}^n) : \|f\|_{osc_{expL^r}} < \infty\}.$$

In the particular case of $r = 1$, Osc_{expL^1} coincides with the BMO space by John-Nirenberg's theorem. Also we have $Osc_{expL^r} \subsetneq BMO$ for any $r > 1$. Other examples are provided by Trudinger's inequality for Riesz potentials. To be more precise, for any $0 < \alpha < n$ and any $f \in L^{n/\alpha}(\mathbb{R}^n)$, the Riesz potential $I_\alpha f$ of order α belongs to $Osc_{expL^{(n/\alpha)'}}$ (cf. [12, 27])

For the statement of our results we introduce some notation that will simplify the presentation. Along this paper m will always be the number of symbols of the operator $T_{\vec{b}}$ where $\vec{b} = (b_1, \dots, b_m)$ is a family of m locally integrable functions. If r_1, \dots, r_m are m positive real numbers, we denote

$$\frac{1}{r} = \frac{1}{r_1} + \dots + \frac{1}{r_m}$$

and

$$\|\vec{b}\| = \prod_{j=1}^m \|b_j\|_{osc_{expL^{r_j}}}.$$

Our main results are the following.

Theorem 1.1 *Let $0 < p < \infty$, $w \in A_\infty$. Suppose that $T_{\vec{b}}f$ is the commutator (1.6) where b is as above such that $b_i \in Osc_{expL^{r_i}}$, $r_i \geq 1$, $1 \leq i \leq m$. Then there exists a constant $C > 0$ such that*

$$\int_{\mathbb{R}^n} |T_{\vec{b}}(f)(x)|^p w(x) dx \leq C \|\vec{b}\|^p \int_{\mathbb{R}^n} (M_{L(\log L)^{1/r}}(f)(x))^p w(x) dx \quad (1.7)$$

for all bounded functions f with compact support.

For the precise definition of the operator $M_{L(\log L)^\alpha}$ see Section 2.2.

Inequality (1.7) shows that the maximal operator $M_{L(\log L)^{1/r}}$ is the one that controls the multilinear commutators T_b . Since $r_i \geq 1$ for each i , it follows that $M_{L(\log L)^{1/r}}$ is pointwise smaller than $M_{L(\log L)^m}$ and this in turn is known to be equivalent to $m + 1$ iterations of the Hardy–Littlewood maximal operator M^{m+1} . Hence we can use again Muckenhoupt’s theorem and deduce the boundedness of T_b on $L^p(w)$ for $p > 1$ and any $w \in A_p$:

Corollary 1.2 *Let $1 < p < \infty$ and $w \in A_p$. Suppose that $T_{\vec{b}}f$ is the commutator (1.6) where b is as above such that $b_i \in Osc_{expL^{r_i}}$, $r_i \geq 1$, $1 \leq i \leq m$. Then there exists a constant $C > 0$ such that*

$$\int_{\mathbb{R}^n} |T_{\vec{b}}(f)(x)|^p w(x) dx \leq C \|\vec{b}\|^p \int_{\mathbb{R}^n} |f(x)|^p w(x) dx \quad (1.8)$$

for all bounded functions f with compact support.

As we mentioned above this estimate is a generalized version of Coifman’s result in [5]. However, our approach is different and will be based on a pointwise inequality that roughly can be expressed as

$$M_\delta^\#(T_{\vec{b}}f)(x) \leq C \|\vec{b}\| M_{L(\log L)^{1/r}} f(x) + R(f)(x) \quad (1.9)$$

for an appropriate positive but small enough number δ . R denotes certain “remainder” operator which is smoother, less singular, than $M_{L(\log L)^{1/r}}$ in some suitable sense. See Lemma 3.1 for a complete and precise statement of this estimate.

Moreover, reasoning as in [21], Theorem 2, from this theorem we deduce the following estimate for general weights. As usual $[\alpha]$ will denote the integer part of α .

Theorem 1.3 *Let $1 < p < \infty$ and let $T_{\vec{b}}$ be as in Theorem 1.1. Then there exists a constant C such that for any weight w*

$$\int_{\mathbb{R}^n} |T_{\vec{b}}(f)(x)|^p w(x) dx \leq C \|\vec{b}\|^p \int_{\mathbb{R}^n} |f(x)|^p M^{[(\frac{1}{r}+1)p]+1}(w)(x) dx$$

for all bounded functions f with compact support.

Remark 1.4 We remark that the number of iterations of the maximal function needed in the Theorem is optimal as can be seen in [21], Section 5. In fact it follows from the proof of Theorem 1.3 that there is a sharper estimate:

$$\int_{\mathbb{R}^n} |T_{\vec{b}}(f)(x)|^p w(x) dx \leq C \|\vec{b}\|^p \int_{\mathbb{R}^n} |f(x)|^p M_{L(\log L)^{(\frac{1}{r}+1)p-1+\epsilon}}(w)(x) dx$$

where $\epsilon > 0$, being the result false for $\epsilon = 0$.

The key pointwise estimate (1.9) is also used to derive the next endpoint result. Recall that commutators with BMO functions are not of weak type $(1, 1)$.

Recall that we denote $\|\vec{b}\| = \prod_{j=1}^m \|b_j\|_{osc_{expL^{r_j}}}$ and $\Phi(t) = t \log^{1/r}(e+t)$, where $\frac{1}{r} = \frac{1}{r_1} + \dots + \frac{1}{r_m}$.

Theorem 1.5 *Let $w \in A_1$. There exists a constant $C > 0$ such that for all $\lambda > 0$*

$$w(\{y \in \mathbb{R}^n : |T_b f(y)| > \lambda\}) \leq C \int_{\mathbb{R}^n} \Phi\left(\frac{\|\vec{b}\| |f(y)|}{\lambda}\right) w(y) dy, \quad (1.10)$$

for all bounded functions f with compact support and for all \vec{b} .

The proof is based on the following result which generalizes inequality (1.4).

Theorem 1.6 *Let $w \in A_\infty$. Then, there exists a positive constant C such that*

$$\sup_{t>0} \frac{1}{\Phi(\frac{1}{t})} w(\{y \in \mathbb{R}^n : |T_b f(y)| > t\}) \leq C \sup_{t>0} \frac{1}{\Phi(\frac{1}{t})} w(\{y \in \mathbb{R}^n : M_\Phi(\|\vec{b}\|f)(y) > t\}) \quad (1.11)$$

for all bounded functions f with compact support.

In any of the above results, if we choose $b_1 = \dots = b_m$ and $r_1 = \dots = r_m = 1$, we recover completely the corresponding results from [20] and [21].

As usual C will denote a positive constant that can change its value on each statement. If the constant depends on precise parameters it will be pointed out.

2 Preliminaries

In this section we introduce the basic tools needed for the proof of the main results.

2.1 The Fefferman-Stein inequality for A_∞ weights

A weight will always mean a positive function which is locally integrable. We say that a weight w belongs to the class A_p , $1 < p < \infty$, if there is a constant C such that

$$\left(\frac{1}{|Q|} \int_Q w(y) dy \right) \left(\frac{1}{|Q|} \int_Q w(y)^{1-p'} dy \right)^{p-1} \leq C$$

for each cube Q and where as usual $\frac{1}{p} + \frac{1}{p'} = 1$. A weight w belongs to the class A_1 if there is a constant C such that

$$\frac{1}{|Q|} \int_Q w(y) dy \leq C \inf_Q w.$$

We will denote the infimum of the constants C by $[w]_{A_p}$. Observe that $[w]_{A_p} \geq 1$ by Jensen's inequality.

Since the A_p classes are increasing with respect to p , the A_∞ class of weights is defined in a natural way by $A_\infty = \cup_{p>1} A_p$. However, the following characterization it is more interesting: there are positive constants c and ρ such that for any cube Q and any measurable set E contained in Q then

$$\frac{w(E)}{w(Q)} \leq c \left(\frac{|E|}{|Q|} \right)^\rho.$$

A simple fact that will be useful is that, for any $w \in A_p$ and any $m > 0$, $w_m = \min\{w, m\} \in A_p$ with $[w_m]_{A_p} \leq C_p[w]_{A_p}$. For more information on A_p weights we remit the reader to the references [11] and [25].

We recall now the definitions of classical maximal operators. If, as usual, M denotes the Hardy–Littlewood maximal operator, we consider for $\delta > 0$

$$M_\delta f(x) = M(|f|^\delta)^{1/\delta}(x) = \left(\sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f(y)|^\delta dy \right)^{1/\delta},$$

$$M^\#(f)(x) = \sup_{Q \ni x} \inf_c \frac{1}{|Q|} \int_Q |f(y) - c| dy \approx \sup_{Q \ni x} \frac{1}{|Q|} \int_Q |f(y) - f_Q| dy$$

and a variant of this sharp maximal operator that will become the main tool in our scheme

$$M_\delta^\# f(x) = M^\#(|f|^\delta)(x)^{1/\delta}.$$

The main inequality between these operators to be used is a version of the classical one due to C. Fefferman and E. Stein (see [24], [16]).

Lemma 2.1 *Let w be an A_∞ weight, then there exists a constant c depending upon the A_∞ condition of w such that for all $\lambda, \epsilon > 0$*

$$w(\{y \in \mathbb{R}^n : Mf(y) > \lambda, M^\#f(y) \leq \lambda\epsilon\}) \leq c\epsilon^\rho w(\{y \in \mathbb{R}^n : Mf(y) > \frac{\lambda}{2}\}).$$

As a consequence we have the following estimates for $\delta > 0$.

a) Let $\varphi : (0, \infty) \rightarrow (0, \infty)$ doubling. Then, there exists a constant c depending upon the A_∞ condition of w and the doubling condition of φ such that

$$\sup_{\lambda>0} \varphi(\lambda) w(\{y \in \mathbb{R}^n : M_\delta f(y) > \lambda\}) \leq c \sup_{\lambda>0} \varphi(\lambda) w(\{y \in \mathbb{R}^n : M_\delta^\# f(y) > \lambda\}) \quad (2.1)$$

for every function such that the left hand side is finite.

b) Let $0 < p < \infty$ there exists a positive constant C upon the A_∞ condition of w and p such that

$$\int_{\mathbb{R}^n} (M_\delta f(x))^p w(x) dx \leq C \int_{\mathbb{R}^n} (M_\delta^\# f(x))^p w(x) dx, \quad (2.2)$$

for every function f such that the left hand side is finite.

2.2 Orlicz maximal functions

By a Young function Φ we shall mean a continuous, nonnegative, strictly increasing and convex function on $[0, \infty)$ with

$$\lim_{t \rightarrow 0^+} \frac{\Phi(t)}{t} = \lim_{t \rightarrow \infty} \frac{t}{\Phi(t)} = 0.$$

We define the Φ -averages of a function f over a cube Q by

$$\|f\|_{\Phi, Q} = \|f\|_{\Phi(L), Q} = \inf\{\lambda > 0 : \frac{1}{|Q|} \int_Q \Phi\left(\frac{|f(x)|}{\lambda}\right) dx \leq 1\}.$$

For Orlicz norms we are usually only concerned about the behavior of Young functions for large values of t 's. Given two functions B and C , we write $B(t) \approx C(t)$ if $B(t)/C(t)$ is bounded and bounded below for $t \geq c > 0$. We also recall that if $B(t) \leq C(t)$ for $t \geq c > 0$, then

$$\|f\|_{B, Q} \leq C \|f\|_{C, Q}$$

with C an absolute constant. For more information on the subject see the reference [22].

Associate to this average we can define a maximal operator M_Φ given by

$$M_\Phi f(x) = M_{\Phi(L)} f(x) = \sup_{Q \ni x} \|f\|_{\Phi, Q},$$

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

For example, if $\Phi(x) = e^{x^r} - 1$, then $\|\cdot\|_{\exp L^r, Q}$ and $M_{\exp L^r}$, denote respectively the Φ -average and the maximal operator associated to Φ . Similarly we have for $\Phi(x) = x \log^r(e+x)$ $\|\cdot\|_{L(\log L)^r, Q}$ and $M_{L(\log L)^r}$. Observe that by the above remarks, $Mf \leq CM_{L(\log L)^r} f$ for any $r > 0$. These examples will be relevant in our work.

Finally, we will be using the following known pointwise inequality: if $m \in \mathbb{N}$ then

$$M_{L(\log L)^m} \sim M^{m+1} = \overbrace{M \circ \dots \circ M}^{(m+1 \text{ times})},$$

the $m+1$ iterations of the Hardy-Littlewood maximal operator. A generalization of this equivalence can be found in [8].

We begin with some technical lemmas on convex functions whose proofs are standard.

Lemma 2.2 ([17], Lema 2.1) *If $\Phi_0, \Phi_1, \dots, \Phi_m$ are real-valued, non-negative, nondecreasing, left continuous functions defined on $[0, \infty)$ such that, with the definition*

$$\Phi_i^{-1}(x) = \inf\{y : \Phi_i(y) > x\},$$

it verify

$$\Phi_1^{-1}(x) \Phi_2^{-1}(x) \cdots \Phi_m^{-1}(x) \leq \Phi_0^{-1}(x), \tag{2.3}$$

then for all $0 \leq x_1, x_2, \dots, x_m < \infty$

$$\Phi_0(x_1 x_2 \cdots x_m) \leq \Phi_1(x_1) + \Phi_2(x_2) + \cdots + \Phi_m(x_m)$$

Lemma 2.3 Let Φ_0 be a convex function such that $\Phi_0(0) = 0$ and let Φ_1, \dots, Φ_m as in the previous lemma all verifying (2.3). If f_1, f_2, \dots, f_m are functions satisfying that $\|f_i\|_{\Phi_i, Q} < \infty$ for all $1 \leq i \leq m$ and for a given cube Q , then

$$\|f_1 \cdots f_m\|_{\Phi_0, Q} \leq m \|f_1\|_{\Phi_1, Q} \cdots \|f_m\|_{\Phi_m, Q}. \quad (2.4)$$

PROOF: Taking $\varepsilon > 0$ small enough, by Lemma 2.2 and convexity

$$\begin{aligned} & \frac{1}{|Q|} \int_Q \Phi_0 \left(\frac{f_1(x) f_2(x) \cdots f_m(x)}{m(\|f_1\|_{\Phi_1, Q} + \varepsilon)(\|f_2\|_{\Phi_2, Q} + \varepsilon) \cdots (\|f_m\|_{\Phi_m, Q} + \varepsilon)} \right) dx \leq \\ & \frac{1}{m} \frac{1}{|Q|} \int_Q \Phi_0 \left(\frac{f_1(x) f_2(x) \cdots f_m(x)}{(\|f_1\|_{\Phi_1, Q} + \varepsilon)(\|f_2\|_{\Phi_2, Q} + \varepsilon) \cdots (\|f_m\|_{\Phi_m, Q} + \varepsilon)} \right) dx \leq \\ & \frac{1}{m} \frac{1}{|Q|} \int_Q \left[\Phi_1 \left(\frac{f_1(x)}{\|f_1\|_{\Phi_1, Q} + \varepsilon} \right) + \cdots + \Phi_m \left(\frac{f_m(x)}{\|f_m\|_{\Phi_m, Q} + \varepsilon} \right) \right] dx \leq 1 \end{aligned}$$

which implies that

$$\|f_1 f_2 \cdots f_m\|_{\Phi_0, Q} \leq m(\|f_1\|_{\Phi_1, Q} + \varepsilon) \cdots (\|f_m\|_{\Phi_m, Q} + \varepsilon).$$

Now, taking letting $\varepsilon \rightarrow 0$ we get (2.4).

The main example that we will be using is the following:

$$\frac{1}{|Q|} \int_Q |f_1 \cdots f_m g| \leq C \|f_1\|_{\exp L^{r_1, Q}} \cdots \|f_m\|_{\exp L^{r_m, Q}} \|g\|_{L(\log L)^{1/r, Q}}, \quad (2.5)$$

where $r_1, \dots, r_m \geq 1$ and

$$\frac{1}{r} = \frac{1}{r_1} + \cdots + \frac{1}{r_m}.$$

Indeed, in this case we can write for any $x \geq 0$ that

$$\log^{\frac{1}{r_1}}(1+x) \cdots \log^{\frac{1}{r_m}}(1+x) \frac{x}{\log^{\frac{1}{r_1} + \cdots + \frac{1}{r_m}}(e+x)} \leq x.$$

Then (2.3) holds for

$$\begin{aligned} \Phi_0(x) &= x, \\ \Phi_i^{-1}(x) &= \log^{\frac{1}{r_i}}(1+x), \end{aligned}$$

$i = 1, \dots, m$ and

$$\Phi_{m+1}^{-1}(x) = \frac{x}{\log^{\frac{1}{r_1} + \cdots + \frac{1}{r_m}}(e+x)}.$$

The inverse functions are given by $\Phi_i(x) = e^{x^{r_i}} - 1$, $i = 1, \dots, m$, and $\Phi_{m+1}(x) \approx x \log^{\frac{1}{r_1} + \cdots + \frac{1}{r_m}}(e+x)$.

2.3 The oscillation of a function

We define the oscillation $osc_{\Phi}(f, Q)$ of a function f with respect to any Young function Φ as

$$osc_{\Phi}(f, Q) = \|f - f_Q\|_{\Phi, Q}.$$

We list the following properties that are very easy to check.

Lemma 2.4 *For any cube Q and any locally integrable functions f and g , we have*

- i) $osc_{\Phi}(f \pm g, Q) \leq osc_{\Phi}(f, Q) + osc_{\Phi}(g, Q)$
- ii) $osc_{\Phi}(\lambda f, Q) = |\lambda| osc_{\Phi}(f, Q)$ for all $\lambda \in \mathbb{R}$.
- iii) $osc_{\Phi}(|f|, Q) \leq 2 osc_{\Phi}(f, Q)$.
- iv) $osc_{\Phi}(\min\{f, g\}, Q) \leq \frac{3}{2}(osc_{\Phi}(f, Q) + osc_{\Phi}(g, Q))$.
- v) $osc_{\Phi}(\max\{f, g\}, Q) \leq \frac{3}{2}(osc_{\Phi}(f, Q) + osc_{\Phi}(g, Q))$.

Given a function g , we define for any $N \in \mathbb{N}$,

$$g^N(x) = \begin{cases} N, & g(x) > N \\ g(x), & |g(x)| \leq N \\ -N, & g(x) < -N \end{cases} \quad (2.6)$$

Now, from Lemma 2.4 and taking into account that $g^N = \max\{\min\{g, N\}, -N\}$, we infer that for any locally integrable function $osc_{\Phi}(g^N, Q) \leq C osc_{\Phi}(g, Q)$ and consequently

$$\|g^N\|_{osc_{\Phi}} \leq C \|g\|_{osc_{\Phi}} \quad (2.7)$$

for all $N \in \mathbb{N}$ and where $C > 0$ is an absolute constant.

3 Proofs of the main results

3.1 Proof of Theorem 1.1

By a kernel K in $\mathbb{R}^n \times \mathbb{R}^n$ we mean a locally integrable function defined away from the diagonal. We say that K satisfies the standard estimates if there exist positive and finite constants γ and C such that, for all distinct $x, y \in \mathbb{R}^n$ and all z with $2|x - z| < |x - y|$, it verifies

- i) $|K(x, y)| \leq C|x - y|^{-n}$,
- ii) $|K(x, y) - K(z, y)| \leq C \left| \frac{x-z}{x-y} \right|^{\gamma} |x - y|^{-n}$,
- iii) $|K(y, x) - K(y, z)| \leq C \left| \frac{x-z}{x-y} \right|^{\gamma} |x - y|^{-n}$.

We define a linear and continuous operator $T : C_0^\infty(\mathbb{R}^n) \rightarrow \mathcal{D}'(\mathbb{R}^n)$ associated to the kernel K by

$$Tf(x) = \int_{\mathbb{R}^n} K(x, y)f(y)dy$$

where $f \in C_0^\infty(\mathbb{R}^n)$ and x is not in the support of f . T is called a Calderón-Zygmund operator if K satisfies the standard estimates and if it extends to a bounded linear operator on $L^2(\mathbb{R}^n)$. These conditions imply that T is also bounded on $L^p(\mathbb{R}^n)$, $1 < p < \infty$ and is of weak type–(1,1). For more information on this subject see [4],[6] or [16].

Given any positive integer m , for all $1 \leq j \leq m$, we denote by C_j^m the family of all finite subset $\sigma = \{\sigma(1), \dots, \sigma(j)\}$ of $\{1, \dots, m\}$ of j different elements. For any $\sigma \in C_j^m$ we associate the complementary sequence σ' given by $\sigma' = \{1, 2, \dots, m\} \setminus \sigma$.

Let $\vec{b} = (b_1, b_2, \dots, b_m)$ be a finite family of integrable functions. For all $1 \leq j \leq m$ and any $\sigma = \{\sigma(1), \dots, \sigma(j)\} \in C_j^m$, we will denote $\vec{b}_\sigma = (b_{\sigma(1)}, \dots, b_{\sigma(j)})$ and the product $b_\sigma = b_{\sigma(1)} \cdots b_{\sigma(j)}$. So, with this notation, we write for any m -tuple $r = (r_1, \dots, r_m)$ of positive numbers

$$\|\vec{b}_\sigma\|_{osc_{expL^{r_\sigma}}} = \|b_{\sigma(1)}\|_{osc_{expL^{r_{\sigma(1)}}}} \cdots \|b_{\sigma(j)}\|_{osc_{expL^{r_{\sigma(j)}}}}.$$

For the product of all the functions we simply write

$$\|\vec{b}\|_{osc_{expL^r}} = \|b_1\|_{osc_{expL^{r_1}}} \cdots \|b_m\|_{osc_{expL^{r_m}}}.$$

For any $\sigma \in C_j^m$, we denote

$$T_{\vec{b}_\sigma}^- f(x) = \int_{\mathbb{R}^n} (b_{\sigma(1)}(x) - b_{\sigma(1)}(y)) \cdots (b_{\sigma(j)}(x) - b_{\sigma(j)}(y))K(x, y)f(y)dy.$$

In the particular case of $\sigma = \{1, \dots, m\}$ we understand $T_{\vec{b}_\sigma}^-$ as simply $T_{\vec{b}}$.

Finally observe that $T_{\vec{b}}^-(f)$ satisfies the following homogeneity on the symbol $T_{\vec{b}}^-(\frac{f}{\|\vec{b}\|}) = T_{\vec{b}}^-(f)$ where $\tilde{b} = (b_1/\|b_1\|_{osc_{expL^{r_1}}}, \dots, b_m/\|b_m\|_{osc_{expL^{r_m}}})$ such that $\|\tilde{b}\| = 1$.

The following lemma gives a pointwise estimate of $M_\delta^\#(T_{\vec{b}}^- f)$ in terms of other maximal functions of operators of lower order. This result can be understood as the extension of Lemma 7.1 in [20] to high order commutators and for the classes of symbols considered in this paper.

Lemma 3.1 *Let $T_{\vec{b}}$ be as in Theorem 1.1 and $0 < \delta < \varepsilon < 1$. Then there exists a constant $C > 0$, depending only on δ and ε such that*

$$M_\delta^\#(T_{\vec{b}}^- f)(x) \leq C \left[\|\vec{b}\| M_{L(\log L)^{1/r}} f(x) + \sum_{j=1}^m \sum_{\sigma \in C_j^m} \|b_\sigma\|_{osc_{expL^{r_\sigma}}} M_\varepsilon(T_{\vec{b}_{\sigma'}}^- f)(x) \right] \quad (3.1)$$

for any bounded function f with compact support.

Recall that $|\vec{b}| = \prod_{j=1}^m \|b_j\|_{osc_{expL^r_j}}$.

PROOF:

Recall that the operator $T_{\vec{b}} = T_{b_1, \dots, b_m}$ is defined by

$$T_{\vec{b}}f(x) = \int_{\mathbb{R}^n} \left[\prod_{j=1}^m (b_j(x) - b_j(y)) \right] K(x, y) f(y) dy.$$

By homogeneity we may assume that $|\vec{b}| = 1$.

We first consider the case $m = 1$. Hence, we only have one symbol $b \in osc_{expL^r}$, $r \geq 1$ and (3.1) becomes

$$M_{\delta}^{\#}(T_b f)(x) \leq C \|b\|_{osc_{expL^r}} \left[M_{L(\log L)^{1/r}} f(x) + M_{\varepsilon}(Tf)(x) \right]. \quad (3.2)$$

For any $\lambda \in \mathbb{R}$ we have

$$\begin{aligned} T_b f(x) &= \int_{\mathbb{R}^n} (b(x) - b(y)) K(x, y) f(y) dy \\ &= (b(x) - \lambda) T f(x) - T((b - \lambda)f)(x). \end{aligned}$$

Now, fixed $x \in \mathbb{R}^n$, for any number c and any ball B centered at x and radius $R > 0$, since $0 < \delta < 1$ implies $|\alpha|^\delta - |\beta|^\delta \leq |\alpha - \beta|^\delta$ for any $\alpha, \beta \in \mathbb{R}$, we can estimate

$$\begin{aligned} \left(\frac{1}{|B|} \int_B \left| |T_b f(y)|^\delta - |c|^\delta \right| dy \right)^{1/\delta} &\leq \left(\frac{1}{|B|} \int_B |T_b f(y) - c|^\delta dy \right)^{1/\delta} \\ &\leq C \left[\left(\frac{1}{|B|} \int_B |(b(y) - \lambda) T f(y)|^\delta dy \right)^{1/\delta} \right. \\ &\quad \left. + \left(\frac{1}{|B|} \int_B |T((b - \lambda)f)(y) - c|^\delta dy \right)^{1/\delta} \right] \\ &= I + II. \end{aligned}$$

We analyze each term separately taking along all the proof $\lambda = (b)_{2B}$, the average of b on the ball concentric with B but with double radius. For any $1 < q < \varepsilon/\delta$ we have by Hölder and Jensen's inequality,

$$\begin{aligned} I &\leq C \left(\frac{1}{|2B|} \int_{2B} |b(y) - \lambda|^{\delta q'} dy \right)^{1/\delta q'} \left(\frac{1}{|B|} \int_B |T f(y)|^{\delta q} dy \right)^{1/\delta q} \\ &\leq C \|b\|_{osc_{expL^r}} M_{\delta q}(Tf)(x) \\ &\leq C M_{\varepsilon}(Tf)(x). \end{aligned} \quad (3.3)$$

To deal with II we split f as usual by $f = f_1 + f_2$ where $f_1 = f \chi_{2B}$ and $f_2 = f - f_1$. This yields

$$\begin{aligned}
II &\leq C \left[\left(\frac{1}{|B|} \int_B |T((b-\lambda)f_1)(y)|^\delta dy \right)^{1/\delta} \right. \\
&\quad \left. + \left(\frac{1}{|B|} \int_B |T((b-\lambda)f_2)(y) - c|^\delta dy \right)^{1/\delta} \right] \\
&= III + IV.
\end{aligned}$$

For *III*, since $(b-\lambda)f_1$ is integrable and T is of weak type $-(1,1)$, by Kolmogorov's inequality ([25], p. 104) and (2.5) we get

$$\begin{aligned}
III &\leq \frac{C}{|2B|} \int_{2B} |b(y) - \lambda| |f(y)| dy \\
&\leq C \|b - \lambda\|_{\exp L^r, 2B} \|f\|_{L(\log L)^{1/r}, 2B} \\
&\leq C \|b\|_{\text{osc}_{\exp L^r}} M_{L(\log L)^{1/r}} f(x) \\
&= C M_{L(\log L)^{1/r}} f(x). \tag{3.4}
\end{aligned}$$

For the last term *IV* we make the election $c = (T((b-\lambda)f_2))_B$. Hence, by Jensen's inequality, it follows that

$$\begin{aligned}
IV &\leq \frac{C}{|B|} \int_B |T((b-\lambda)f_2)(y) - (T((b-\lambda)f_2))_B| dy \\
&= \frac{C}{|B|} \int_B \left| \left(\int_{\mathbb{R}^n} K(y, w)(b(w) - \lambda)f_2(w) dw \right) \right. \\
&\quad \left. - \frac{1}{|B|} \int_B \left(\int_{\mathbb{R}^n} K(z, w)(b(w) - \lambda)f_2(w) dw \right) dz \right| dy \\
&= \frac{C}{|B|} \int_B \left| \frac{1}{|B|} \int_B \left[\int_{\mathbb{R}^n \setminus 2B} (K(y, w) - K(z, w))(b(w) - \lambda)f(w) dw \right] dz \right| dy \\
&\leq \frac{C}{|B|^2} \int_B \int_B \left[\int_{\mathbb{R}^n \setminus 2B} |K(y, w) - K(z, w)| |b(w) - \lambda| |f(w)| dw \right] dz dy \\
&\leq \frac{C}{|B|^2} \int_B \int_B \left[\sum_{k=1}^{\infty} \int_{2^k R \leq |w-x| < 2^{k+1} R} \left| \frac{y-z}{y-w} \right|^\gamma \frac{1}{|y-w|^n} |b(w) - \lambda| |f(w)| dw \right] dz dy \\
&\leq C \sum_{k=1}^{\infty} \left(\frac{2R}{2^k R} \right)^\gamma \frac{1}{(2^k R)^n} \int_{2^{k+1} B} |b(w) - \lambda| |f(w)| dw \\
&\leq C \sum_{k=1}^{\infty} 2^{-k\gamma} \|b - \lambda\|_{\exp L^r, 2^{k+1} B} \|f\|_{L(\log L)^{1/r}, 2^{k+1} B}. \tag{3.5}
\end{aligned}$$

The fifth inequality follows because $y, z \in B$ and $w \in \mathbb{R}^n \setminus B$ and therefore $2|y - w| < |z - w|$, the sixth does since $|y - z| \leq 2R$ and $|y - w|^{-1} \leq ((2^k - 1)R)^{-1} \leq C(2^k R)^{-1}$, and the last one is an application of the generalized Hölder inequality (2.5).

Now, we claim that

$$\|b - \lambda\|_{expL^r, 2^{k+1}B} \leq Ck \|b\|_{osc_{expL^r}}. \quad (3.6)$$

Indeed,

$$\begin{aligned} \|b - \lambda\|_{expL^r, 2^{k+1}B} &\leq \|b - (b)_{2^{k+1}B}\|_{expL^r, 2^{k+1}B} + \|(b)_{2^{k+1}B} - \lambda\|_{expL^r, 2^{k+1}B} \\ &\leq \|b - (b)_{2^{k+1}B}\|_{expL^r, 2^{k+1}B} + |(b)_{2^{k+1}B} - \lambda| \\ &\leq Ck \|b\|_{osc_{expL^r}} \end{aligned}$$

where the last estimate follows by the standard inequality $|(b)_{2B} - b_B| \leq 2\|b\|_{osc_{expL^r}}$ (cf. [16], p. 31), and then we take supremum over the balls.

Thus, IV is estimated by

$$\begin{aligned} IV &\leq C \sum_{k=1}^{\infty} (2^{-k})^\gamma k \|b\|_{osc_{expL^r}} M_{L(\log L)^{1/r}} f(x) \\ &\leq CM_{L(\log L)^{1/r}} f(x). \end{aligned} \quad (3.7)$$

From (3.4) and (3.7) we conclude

$$II \leq CM_{L(\log L)^{1/r}} f(x)$$

which, together with (3.3), gives (3.2) and proves the lemma for $m = 1$.

Consider now the case $m \geq 2$. Then, for any $\vec{\lambda} = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^n$ we have

$$\begin{aligned} T_{\vec{b}} f(x) &= \int_{\mathbb{R}^n} (b_1(x) - b_1(y)) \cdots (b_m(x) - b_m(y)) K(x, y) f(y) dy \\ &= \int_{\mathbb{R}^n} ((b_1(x) - \lambda_1) - (b_1(y) - \lambda_1)) \cdots ((b_m(x) - \lambda_m) - (b_m(y) - \lambda_m)) K(x, y) f(y) dy \\ &= \sum_{j=0}^m \sum_{\sigma \in C_j^m} (-1)^{m-j} (b(x) - \vec{\lambda})_\sigma \int_{\mathbb{R}^n} (b(y) - \vec{\lambda})_{\sigma'} K(x, y) f(y) dy \\ &= (b_1(x) - \lambda_1) \cdots (b_m(x) - \lambda_m) T f(x) \\ &\quad + (-1)^m T((b_1 - \lambda_1) \cdots (b_m - \lambda_m) f)(x) \\ &\quad + \sum_{j=1}^{m-1} \sum_{\sigma \in C_j^m} (-1)^{m-j} (b(x) - \vec{\lambda})_\sigma \int_{\mathbb{R}^n} (b(y) - b(x))_{\sigma'} K(x, y) f(y) dy \end{aligned}$$

Now, expanding $(b(y) - \vec{\lambda})_{\sigma'} = [(b(y) - b(x)) + (b(x) - \vec{\lambda})]_{\sigma'}$ as above it is easy to see that

$$\begin{aligned} T_{\vec{b}}f(x) &= (b_1(x) - \lambda_1) \cdots (b_m(x) - \lambda_m)Tf(x) \\ &\quad + (-1)^m T((b_1 - \lambda_1) \cdots (b_m - \lambda_m)f)(x) \\ &\quad + \sum_{j=1}^{m-1} \sum_{\sigma \in C_j^m} c_{m,j} (b(x) - \vec{\lambda})_{\sigma} T_{b_{\sigma'}} f(x) \end{aligned}$$

where $c_{m,j}$ are absolute constants depending only on m and j .

Now, fixed $x \in \mathbb{R}^n$, for any number c and any ball B centered at x and radius $R > 0$ it follows since $0 < \delta < 1$

$$\begin{aligned} \left(\frac{1}{|B|} \int_B \left| |T_{\vec{b}}f(y)|^\delta - |c|^\delta \right| dy \right)^{1/\delta} &\leq \left(\frac{1}{|B|} \int_B |T_{\vec{b}}f(y) - c|^\delta dy \right)^{1/\delta} \\ &\leq C \left[\left(\frac{1}{|B|} \int_B |(b_1(y) - \lambda_1) \cdots (b_m(y) - \lambda_m)Tf(y)|^\delta dy \right)^{1/\delta} \right. \\ &\quad \left. + \sum_{j=1}^{m-1} \sum_{\sigma \in C_j^m} \left(\frac{1}{|B|} \int_B |(b - \vec{\lambda})_{\sigma} T_{b_{\sigma'}} f(y)|^\delta dy \right)^{1/\delta} \right. \\ &\quad \left. + \left(\frac{1}{|B|} \int_B |T((b_1 - \lambda_1) \cdots (b_m - \lambda_m)f)(y) - c|^\delta dy \right)^{1/\delta} \right] \\ &= I + II + III. \end{aligned}$$

Reasoning as in (3.3) with $\lambda_i = (b_i)_{2B}$, $i = 1, \dots, m$, using this time the standard Hölder inequality for finitely many functions with $1 < q < \varepsilon/\delta$, we have

$$I \leq CM_\varepsilon(Tf)(x) \tag{3.8}$$

and

$$II \leq C \sum_{j=1}^{m-1} \sum_{\sigma \in C_j^m} M_\varepsilon(T_{b_{\sigma'}} f)(x) \tag{3.9}$$

with $C > 0$ depending only on m .

For III we split $f = f_1 + f_2$ with $f_1 = f\chi_{2B}$ and $f_2 = f - f_1$. So

$$\begin{aligned} III &\leq C \left[\left(\frac{1}{|B|} \int_B |T((b_1 - \lambda_1) \cdots (b_m - \lambda_m)f_1)(y)|^\delta dy \right)^{1/\delta} \right. \\ &\quad \left. + \left(\frac{1}{|B|} \int_B |T((b_1 - \lambda_1) \cdots (b_m - \lambda_m)f_2)(y) - c|^\delta dy \right)^{1/\delta} \right] \\ &= IV + V. \end{aligned}$$

Now, as in (3.4) and making use again of Kolmogorov's inequality and (2.5), we can estimate IV by

$$\begin{aligned}
IV &\leq \frac{C}{|2B|} \int_{2B} |b_1(y) - \lambda_1| \cdots |b_m(y) - \lambda_m| |f(y)| dy \\
&\leq C \|b_1 - \lambda_1\|_{\exp L^{r_1}, 2B} \cdots \|b_m - \lambda_m\|_{\exp L^{r_m}, 2B} \|f\|_{L(\log L)^{1/r}, 2B} \\
&\leq CM_{L(\log L)^{1/r}} f(x),
\end{aligned} \tag{3.10}$$

where recall that

$$\frac{1}{r} = \frac{1}{r_1} + \cdots + \frac{1}{r_m}.$$

Finally, for V , choosing $c = (T((b_1 - \lambda_1) \cdots (b_m - \lambda_m) f_2))_B$ and repeating the argument used to get (3.5), it follows from (2.5) and (3.6) that

$$\begin{aligned}
V &\leq C \sum_{k=1}^{\infty} 2^{-k\gamma} \frac{1}{(2^{k+1}R)^n} \int_{2^{k+1}B} \left(\prod_{j=1}^m |b_j(w) - \lambda_j| \right) |f(w)| dw \\
&\leq C \sum_{k=1}^{\infty} 2^{-k\gamma} \left(\prod_{j=1}^m \|b_j - \lambda_j\|_{\exp L^{r_j}, 2^{k+1}B} \right) \|f\|_{L(\log L)^{1/r}, 2^{k+1}B} \\
&\leq C \left[\sum_{k=1}^{\infty} 2^{-k\gamma} k^m \right] \left(\prod_{j=1}^m \|b_j\|_{\text{osc}_{\exp L^{r_j}}} \right) M_{L(\log L)^{1/r}} f(x).
\end{aligned} \tag{3.11}$$

Finally, from (3.10) and (3.11) we conclude

$$III \leq CM_{L(\log L)^{1/r}} f(x),$$

which together with (3.8) and (3.9) gives (3.1) and the proof of the lemma is finished.

We are now in a position to prove Theorem 1.1

PROOF OF THEOREM 1.1: We may assume that

$$\int_{\mathbb{R}^n} (M_{L(\log L)^{1/r}} f(x))^p w(x) dx < \infty, \tag{3.12}$$

since otherwise there is nothing to be proved.

To apply the Fefferman-Stein inequality (2.2) we first take for granted that $\|M_\delta(T_\delta f)\|_{L^p(w)}$ is finite. We will check this to the end of the proof.

We proceed by induction on m . For $m = 1$, by (2.2) and Lemma 3.1 we can estimate

$$\begin{aligned}
\|T_{b_1}f\|_{L^p(w)} &\leq \|M_\delta(T_{b_1}f)\|_{L^p(w)} \\
&\leq C \|M_\delta^\#(T_{b_1}f)\|_{L^p(w)} \\
&\leq C \|b_1\|_{\text{osc}_{\text{exp}L^{r_1}}} \left[\|M_\varepsilon(Tf)\|_{L^p(w)} + \|M_{L(\log L)^{1/r_1}}f\|_{L^p(w)} \right] \\
&\leq C \|b_1\|_{\text{osc}_{\text{exp}L^{r_1}}} \left[\|Tf\|_{L^p(w)} + \|M_{L(\log L)^{1/r_1}}f\|_{L^p(w)} \right] \\
&\leq C \|b_1\|_{\text{osc}_{\text{exp}L^{r_1}}} \left[\|Mf\|_{L^p(w)} + \|M_{L(\log L)^{1/r_1}}f\|_{L^p(w)} \right] \\
&\leq C \|b_1\|_{\text{osc}_{\text{exp}L^{r_1}}} \|M_{L(\log L)^{1/r_1}}f\|_{L^p(w)}
\end{aligned}$$

where the fourth inequality follows since $w \in A_\infty$ and therefore there exists $q > 1$ such that $w \in A_q$, then we can choose ε such that $0 < \varepsilon < p/q$ (we may take $q > p$ if necessary). The fifth holds by the classical estimate (1.3) for $m = 0$ (see [6], Chapter 1).

Suppose now that for $m - 1$ the theorem is true and let's prove it for m . So, the same argument used above and by the induction hypothesis gives

$$\begin{aligned}
\|T_{\vec{b}}f\|_{L^p(w)} &\leq \|M_\delta(T_{\vec{b}}f)\|_{L^p(w)} \\
&\leq C \|M_\delta^\#(T_{\vec{b}}f)\|_{L^p(w)} \\
&\leq C \left[\|b_1\|_{\text{osc}_{\text{exp}L^{r_1}}} \cdot \dots \cdot \|b_m\|_{\text{osc}_{\text{exp}L^{r_m}}} \|M_{L(\log L)^{1/r}}f\|_{L^p(w)} \right. \\
&\quad \left. + \sum_{j=1}^{m-1} \sum_{\sigma \in C_j^m} \|b_\sigma\|_{\text{osc}_{\text{exp}L^\sigma}} \|M_\varepsilon(T_{b_{\sigma'}})\|_{L^p(w)} \right] \\
&\leq C \left[\|b_1\|_{\text{osc}_{\text{exp}L^{r_1}}} \cdot \dots \cdot \|b_m\|_{\text{osc}_{\text{exp}L^{r_m}}} \|M_{L(\log L)^{1/r}}f\|_{L^p(w)} \right. \\
&\quad \left. + \sum_{j=1}^{m-1} \sum_{\sigma \in C_j^m} \|b_\sigma\|_{\text{osc}_{\text{exp}L^\sigma}} \|b_{\sigma'}\|_{\text{osc}_{\text{exp}L^{\sigma'}}} \|M_{L(\log L)^{1/r_{\sigma'}}}f\|_{L^p(w)} \right] \\
&\leq C \|b_1\|_{\text{osc}_{\text{exp}L^{r_1}}} \cdot \dots \cdot \|b_m\|_{\text{osc}_{\text{exp}L^{r_m}}} \|M_{L(\log L)^{1/r}}f\|_{L^p(w)},
\end{aligned}$$

since $M_{L(\log L)^{1/r_{\sigma'}}} \leq C M_{L(\log L)^{1/r}}$

Let us check now that for appropriate δ we have $\|M_\delta(T_{\vec{b}}f)\|_{L^p(w)} < \infty$. Indeed, as above since $w \in A_\infty$, there exists $q > 1$ such that $w \in A_q$ and we can choose δ small enough so that $p/\delta > q$. Then by Muckehaupt's theorem all is reduced to checking that $\|T_{\vec{b}}f\|_{L^p(w)} < \infty$.

Suppose that the symbols b_k 's and the weight w are all bounded functions. Since f has compact support we may assume that the support of f is contained in the ball $B_R = B(0, R)$. Then we can split the integral as

$$\int_{\mathbb{R}^n} |T_{\vec{b}}f(x)|^p w(x) dx = \int_{|x| \leq 2R} |T_{\vec{b}}f(x)|^p w(x) dx + \int_{|x| > 2R} |T_{\vec{b}}f(x)|^p w(x) dx.$$

The first integral can be easily estimated making use of the L^∞ -boundedness of the b_k 's and w and the L^q -boundedness for $q > 1$ of the Calderón–Zygmund operator T .

For the second term, by the properties of the kernel K and the boundedness of the symbols b_k 's, since $|x| > 2R$, we have the following pointwise estimate

$$\begin{aligned} |T_{\vec{b}}f(x)| &\leq C \int_{B_R} \frac{|b_1(x) - b_1(y)| \cdots |b_m(x) - b_m(y)| |f(y)|}{|x - y|^n} dy \\ &\leq \frac{C}{|x|^n} \int_{B(0, |x|)} |f(y)| dy \\ &\leq CMf(x) \\ &\leq CM_{L(\log L)^{\frac{1}{r}}} f(x). \end{aligned} \tag{3.13}$$

Thus,

$$\int_{|x| > 2R} |T_{\vec{b}}f(x)|^p w(x) dx \leq C \int_{|x| > 2R} (M_{L(\log L)^{\frac{1}{r}}} f(x))^p w(x) dx$$

which is finite by the assumption (3.12).

For the general case, we will truncate the symbols b_k 's and the weight w as follows (cf [6], p. 40). We denote by $b^{\vec{N}}$ the vector of truncated elements by N , i.e., $b^{\vec{N}} = (b_1^N, \dots, b_m^N)$ where each b_k^N is the truncation of b_k as it is defined in (2.6). Observe that in our case (2.7) becomes

$$\|b_k^N\|_{osc_{expL^{r_k}}} \leq C \|b_k\|_{osc_{expL^{r_k}}} \tag{3.14}$$

with $C > 0$ a constant independent of N . Analogously we consider the truncations of the weight w by $w_N = \inf\{w, N\}$ that satisfy

$$[w_N]_{A_\infty} \leq C [w]_{A_\infty}. \tag{3.15}$$

Then (1.7) holds for the operator $T_{b^{\vec{N}}}$ and the weight w_N . Combining (3.14) and (3.15), this estimate gives

$$\int_{\mathbb{R}^n} |T_{b^{\vec{N}}}f(x)|^p w_N(x) dx \leq C \left(\prod_{j=1}^m \|b_j\|_{osc_{expL^{r_j}}}^p \right) \int_{\mathbb{R}^n} (M_{L(\log L)^{1/r}} f(x))^p w(x) dx.$$

Next, taking into account that f has compact support, we deduce that any product $b_{i_1}^N \cdots b_{i_k}^N f$ converges in any L^q for $q > 1$ to $b_{i_1} \cdots b_{i_k} f$ as $N \rightarrow \infty$. Hence, the classical L^q -boundedness of the operator T gives, at least for a subsequence, that $|T_{b^{\vec{N}}}f(x)|^p w_N(x)$ converges pointwise almost everywhere to $|T_{\vec{b}}f(x)|^p w(x)$ and by Fatou's lemma we conclude the theorem for this general case. The theorem is proved.

3.2 Proof of Theorem 1.5

We adapt here some of the arguments from [20]. Since the proof of Theorem 1.5 is based on Theorem 1.6 we prove this first, namely we must show that

$$\sup_{t > 0} \frac{1}{\Phi(\frac{1}{t})} w(\{y \in \mathbb{R}^n : |T_{\vec{b}}f(y)| > t\}) \leq C \sup_{t > 0} \frac{1}{\Phi(\frac{1}{t})} w(\{y \in \mathbb{R}^n : M_\Phi(|\vec{b}|f)(y) > t\}) \tag{3.16}$$

for all bounded functions f with compact support. Recall that $\Phi(t) = \Phi_{\vec{b}}(t) = t \log^{1/r}(e+t)$.

In fact we are going to prove something stronger than (3.16) namely:

For every \vec{b} , $\varphi : (0, \infty) \rightarrow (0, \infty)$ doubling with $\varphi(t) \leq Ct$, $t > 0$, and for every $0 < \delta < 1$ there exists a constant C such that

$$\sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\delta(T_{\vec{b}}f)(y) > t\}) \leq C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\Phi(|\vec{b}|f)(y) > t\}) \quad (3.17)$$

for all bounded functions f with compact support.

By the Lebesgue differentiation theorem and taking $\varphi(t) = \Phi(\frac{1}{t})^{-1} = \frac{t}{\log^{1/r}(e+\frac{1}{t})}$ it is clear that (3.17) implies (3.16).

By making use of the weighted version of the Fefferman-Stein lemma 2.1, more precisely estimate (2.1), we have that

$$\sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\delta(T_{\vec{b}}f)(y) > t\}) \leq C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\delta^\#(T_{\vec{b}}f)(y) > t\}) \quad (3.18)$$

whenever the left hand side is finite. Therefore (3.17) will follow from

$$\sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\delta^\#(T_{\vec{b}}f)(y) > t\}) \leq C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\Phi(|\vec{b}|f)(y) > t\}). \quad (3.19)$$

We first check that the left hand side of (3.18) is finite for all bounded function f with compact support. By proceeding as in the proof of Theorem 1.1, we may assume that b and w are bounded. For the general case of unbounded symbols and unbounded weight we reproduce the argument used in the proof of Theorem 1.1 taking into account this time the weak-(1,1) boundedness of the operator T which gives the convergence in measure.

Suppose that $\text{supp } f \subset B_R = B(0, R)$. Hence, since $0 < \delta < 1$, it follows

$$\begin{aligned} \varphi(t) w(\{y \in \mathbb{R}^n : M_\delta(T_{\vec{b}}f)(y) > t\}) &\leq C \varphi(t) |\{y \in \mathbb{R}^n : M_\delta(\chi_{B_{2R}} T_{\vec{b}}f)(y) > t/2\}| \\ &\quad + C \varphi(t) |\{y \in \mathbb{R}^n : M_\delta(\chi_{\mathbb{R}^n \setminus B_{2R}} T_{\vec{b}}f)(y) > t/2\}| \\ &= I + II. \end{aligned}$$

For I we use that M is of weak type-(1,1) and the fact that $\varphi(t) \leq Ct$, then

$$\begin{aligned} I &\leq Ct |\{y \in \mathbb{R}^n : M(\chi_{B_{2R}} T_{\vec{b}}f)(y) > t/2\}| \\ &\leq C \int_{B_{2R}} |T_{\vec{b}}f(y)| dy \\ &\leq CR^{n/2} \left(\int_{\mathbb{R}^n} |Tf(y)|^2 dy \right)^{1/2}, \end{aligned}$$

which is finite since T is a Calderón-Zygmund operator and using that the symbols b_k 's are bounded.

For II we take into account the pointwise estimate (3.13) and the well known fact that $(Mf)^\delta \in A_1$, then we have

$$\begin{aligned} II &\leq Ct |\{y \in \mathbb{R}^n : M_\delta(Mf)(y) > Ct\}| \\ &\leq Ct |\{y \in \mathbb{R}^n : Mf(y) > Ct\}| \\ &\leq C \int_{\mathbb{R}^n} |f(y)| dy < \infty. \end{aligned}$$

Combining the homogeneity and the linearity of $T_{\vec{b}}$, it is easy to see that we may assume that $\|\vec{b}\| = 1$ in both (1.10) and (3.16).

To prove (3.19) we proceed by induction on m .

3.3 The case $m = 1$

This case is essentially taken from [20] and we repeat it, with minor modifications, for the sake of completeness. In this case the operator T_b is simply defined by one single function b

$$T_b f = [b, T]f = bT(f) - T(bf),$$

where T is any Calderón–Zygmund operator. Recall that by homogeneity we may assume that $\|b\| = \|b\|_{osc_{expL^r}} = 1$ and therefore what we must prove is

$$\sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\delta^\#([b, T]f)(y) > t\}) \leq C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_{L(\log L)^{1/r}}(f)(y) > t\}) \quad (3.20)$$

for all bounded functions f with compact support. Now applying Lemma 3.1 with any α such that $\delta < \alpha < 1$ we have that the left hand side of (3.20) is estimated by

$$\begin{aligned} &C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : C [M_{L(\log L)^{1/r}}(f)(y) + M_\alpha(Tf)(y)] > t\}) \\ &\leq C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_{L(\log L)^{1/r}}(f)(y) > t\}) + C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\alpha(Tf)(y) > t\}) \end{aligned}$$

where we also have used the doubling condition of φ .

Next, considering the estimate

$$M_\alpha^\#(Tf)(y) \leq C_\alpha Mf(y), \quad (3.21)$$

which holds for all $0 < \alpha < 1$ (see [1], Theorem 2.1), if we further select α such that $0 < \delta < \alpha < 1$, the Fefferman-Stein's lemma yields

$$\begin{aligned} &\sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\delta^\#([b, T]f)(y) > t\}) \leq \\ &\leq C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_{L(\log L)^{1/r}}(f)(y) > t\}) + C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\alpha^\#(Tf)(y) > t\}) \\ &\leq C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_{L(\log L)^{1/r}}(f)(y) > t\}) + C \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M(f)(y) > t\}) \\ &\leq C \sup_{t>0} C \varphi(t) w(\{y \in \mathbb{R}^n : M_{L(\log L)^{1/r}}(f)(y) > t\}), \end{aligned}$$

since trivially $M(f) = M_L(f) \leq M_{L(\log L)^{1/r}}(f)$. This finishes the proof of (3.20).

3.4 The general case

Suppose now that (3.19) holds for $m - 1$ and let's prove it for m . Recall that $\Phi(t) = \Phi_b(t) = t \log^{1/r}(e + t)$ with $\frac{1}{r} = \frac{1}{r_1} + \dots + \frac{1}{r_m}$.

Then, by Lemma 3.1,

$$\begin{aligned} & \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\delta^\#(T_b^- f)(y) > t\}) \\ & \leq \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : C \left[M_\Phi(f)(y) + \sum_{j=1}^m \sum_{\sigma \in C_j^m} \|b_\sigma\|_{osc_{expL^r\sigma}} M_{\varepsilon'}(T_{b_{\sigma'}^-} f)(y) \right] > t\}) \\ & \leq C_m \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\Phi(f)(y) > t\}) \\ & \quad + C_m \sum_{j=1}^m \sum_{\sigma \in C_j^m} \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_{\varepsilon'}(T_{b_{\sigma'}^-}(\|b_\sigma\|_{osc_{expL^r\sigma}} f))(y) > t\}) \end{aligned}$$

Since $\varepsilon' < 1$ and we already checked that the distribution set on the left hand side is finite, combining the Fefferman–Stein's lemma together with the induction hypothesis on (3.19) to $T_{b_{\sigma'}^-}$ we can estimate the last expression by

$$\begin{aligned} & C_m \sum_{j=1}^m \sum_{\sigma \in C_j^m} \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_{\varepsilon'}^\#(T_{b_{\sigma'}^-}(\|b_\sigma\|_{osc_{expL^r\sigma}} f))(y) > t\}) \\ & \leq C_m \sum_{j=1}^m \sum_{\sigma \in C_j^m} \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_{\Phi_{b_{\sigma'}^-}}(\|b_{\sigma'}\|_{osc_{expL^r\sigma'}} \|b_\sigma\|_{osc_{expL^r\sigma}} f)(y) > t\}) \\ & \leq C_m \sum_{j=1}^m \sum_{\sigma \in C_j^m} \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_{\Phi_{b_{\sigma'}^-}}(f)(y) > t\}) \end{aligned}$$

since $\|\vec{b}_{\sigma'}\|_{osc_{expL^r\sigma'}} \|\vec{b}_\sigma\|_{osc_{expL^r\sigma}} = \|\vec{b}\| = 1$. Finally, using the trivial observation that $M_{\Phi_{b_{\sigma'}^-}}(f) \leq M_{\Phi_{b_\sigma^-}}(f) = M_\Phi(f)$ we have

$$\sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\delta^\#(T_b^- f)(y) > t\}) \leq C_m \sup_{t>0} \varphi(t) w(\{y \in \mathbb{R}^n : M_\Phi(f)(y) > t\})$$

and the claim (3.19) is proved.

We need the following lemma concerning estimates of the maximal operator M_Φ , which is a more general version than the one given in Lemma 8.3 [20]. The proof is standard and we shall omit it.

Lemma 3.2 *Let $w \in A_1$, then there exists a positive constant C such that for any $t > 0$ and any locally integrable function f*

$$w(\{y \in \mathbb{R}^n : M_\Phi f(y) > t\}) \leq C \int_{\mathbb{R}^n} \Phi\left(\frac{|f(y)|}{t}\right) w(y) dy.$$

We are now in position to prove Theorem 1.5

PROOF OF THE THEOREM 1.5: By homogeneity it is enough to assume $t = \|b\| = 1$ and hence we must prove

$$w(\{y \in \mathbb{R}^n : |T_b f(y)| > 1\}) \leq C \int_{\mathbb{R}^n} \Phi(|f(y)|)w(y)dy.$$

Now, since Φ is submultiplicative, namely $\Phi(ab) \leq 2\Phi(a)\Phi(b)$, $a, b \geq 0$ we have by Theorem 1.6 and Lemma 3.2

$$\begin{aligned} w(\{y \in \mathbb{R}^n : |T_b f(y)| > 1\}) &\leq C \sup_{t>0} \frac{1}{\Phi(\frac{1}{t})} w(\{y \in \mathbb{R}^n : |T_b f(y)| > t\}) \\ &\leq C \sup_{t>0} \frac{1}{\Phi(\frac{1}{t})} w(\{y \in \mathbb{R}^n : M_\Phi f(y) > t\}) \\ &\leq C \sup_{t>0} \frac{1}{\Phi(\frac{1}{t})} \int_{\mathbb{R}^n} \Phi\left(\frac{|f(y)|}{t}\right)w(y)dy \\ &\leq C \sup_{t>0} \frac{1}{\Phi(\frac{1}{t})} \int_{\mathbb{R}^n} \Phi(|f(y)|)\Phi\left(\frac{1}{t}\right)w(y)dy \\ &\leq C \int_{\mathbb{R}^n} \Phi(|f(y)|)w(y)dy \end{aligned}$$

and the proof is concluded.

References

- [1] ALVAREZ, J., and PÉREZ, C., *Estimates with A_∞ weights for various singular integral operators*, Boll. Un. Mat. Ital. A (7) **8** (1994), no. 1, 123–133.
- [2] CHIARENZA, F., FRASCA, M., and LONGO, P., *Interior $W^{2,p}$ estimates for non divergence elliptic equations with discontinuous coefficients*, Recherche Mat. **40** (1991), 149–168.
- [3] CHIARENZA, F., FRASCA, M., and LONGO, P., *$W^{2,p}$ -solvability of the Dirichlet problem for nondivergence elliptic equations with VMO coefficients*, Trans. Amer. Math. Soc. **334** (1993), 841–853.
- [4] CHRIST, M., *Lectures on Singular Integral Operators*, Reg. Conferences Series in Math. **77**. Amer. Math. Soc., Providence, 1990.
- [5] COIFMAN, R., *Distribution function inequalities for singular integrals*, Proc. Acad. Sci. U.S.A. **69**, 2838–2839.
- [6] COIFMAN, R., and MEYER, Y., *Wavelets. Calderón-Zygmund and multilinear operators*. Cambridge Studies in Advanced Mathematics, 48. Cambridge University Press, Cambridge, 1997.
- [7] COIFMAN, R., ROCHBERG, R., and WEISS, G., *Factorization theorems for Hardy spaces in several variables*, Ann. of Math. **103** (1976), 611–635.

- [8] CAROZZA, M., and PASSARELLI, A., *Composition of maximal operators*, Pub. Mat. **40** (1996), 397–409.
- [9] DI FAZIO, G., and RAGUSA, M. A., *Interior estimates in Morrey spaces for strong solutions to nondivergence form equations with discontinuous coefficients*, J. of Functional Analysis **112** (1993), 241–256.
- [10] GARCÍA-CUERVA, J., HARBOURE, E., SEGOVIA, C., and TORREA, J. L., *Weighted norm inequalities for commutators of strongly singular integrals*, Indiana Univ. Math. J. **40** (1991), 1397–1420.
- [11] GARCÍA-CUERVA, J. and RUBIO DE FRANCIA, J. L., *Weighted Norm Inequalities and Related Topics*, North Holland Math. Studies **116**, North Holland, Amsterdam, 1985.
- [12] GATTO, A. E., and VÁGI, S., *On the exponential integrability of fractional integrals on spaces of homogeneous type*, Colloq. Math. **64** (1993), 121–127.
- [13] GRECO, L., and IWANIEC, T., *New inequalities for the Jacobian*, Ann. Inst. Henri Poincaré, **11** (1994), 17–35.
- [14] IWANIEC, T., and SBORDONE, C., *Weak minima of variational integrals*, J. Reine Angew. Math. **454**, 143–161.
- [15] JANSON, S., *Mean oscillation and commutators of singular integral operators*, Ark. Mat. **16** (1978), 263–270.
- [16] JOURNÉ, J. L., *Calderón-Zygmund Operators, Pseudo-Differential Operator, and the Cauchy Integral of Calderón*, Lecture Notes in Math., Vol. 994, Springer-Verlag, New York, 1983.
- [17] O’NEIL, R., *Fractional integration in Orlicz spaces*, Trans. Amer. Math. Soc. **115** (1963), 300–328.
- [18] PÉREZ, C., *Weighted norm inequalities for singular integral operators*, J. London Math. Soc. **49** (1994), 296–308.
- [19] PÉREZ, C., *On sufficient conditions for the boundedness of the Hardy–Littlewood maximal operator between weighted L^p -spaces with different weights*, Proc. of the London Math. Soc. (3) **71** (1995), 135–157.
- [20] PÉREZ, C., *Endpoint estimates for commutators of singular integral operators*, J. Funct. Anal. **128** (1995), 163–185.
- [21] PÉREZ, C., *Sharp estimates for commutators of singular integrals via iterations of the Hardy-Littlewood maximal function*, J. Fourier Anal. Appl. **3** (6) (1997), 743–756.
- [22] RAO, M.M., and REN, Z.D., *Theory of Orlicz Spaces*, Marcel Dekker, New York, 1991.
- [23] ROCHBERG, R. and WEISS, G., *Derivatives of analytic families of Banach spaces*, Ann. of Math. **118** (1983), 315–347.

- [24] STEIN, E. M., *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series, **43**, Princeton University Press, Princeton , NJ, (1993).
- [25] TORCHINSKY, A., *Real Variable Methods in Harmonic Analysis*, Academic Press, New York, 1988.
- [26] WILSON, J. M., *Weighted norm inequalities for the continuous square functions*, Trans. Amer. Math. Soc. **314** (1989), 661–692.
- [27] ZIEMER, W. P., *Weakly Differentiable Functions*, Graduate Texts in Mathematics, Vol. 120, Springer-Verlag, New York/Berlin, 1989.

CARLOS PÉREZ
DEPARTAMENTO DE ANÁLISIS MATEMÁTICO
FACULTAD DE MATEMÁTICAS
UNIVERSIDAD DE SEVILLA
41080 SEVILLA
SPAIN
E-mail address: CARLOSPEREZ@US.ES

RODRIGO TRUJILLO-GONZÁLEZ
DEPARTAMENTO DE ANÁLISIS MATEMÁTICO
UNIVERSIDAD DE LA LAGUNA
38271 LA LAGUNA - S/C DE TENERIFE
SPAIN
E-mail address: ROTRUJIL@ULL.ES