



Calderón commutators and the Cauchy integral on Lipschitz curves revisited II. The Cauchy integral and its generalizations

Camil Muscalu

Dedicated to Professor Nicolae Popa, on the occasion of his 70th birthday

Abstract. This article is the second in a series of three papers, whose aim is to give new proofs to the well known theorems of Calderón, Coifman, McIntosh and Meyer [1], [4], and [5]. Here we treat the case of the Cauchy integral on Lipschitz curves and some of its generalizations.

1. Introduction

This paper is a continuation of [9] and it is the second of a three papers series. Let A be a Lipschitz function on the real line \mathbb{R} . It defines a Lipschitz curve Γ in the complex plane, with the parametrization $x \rightarrow x + iA(x)$. The Cauchy integral associated with this curve is the singular integral operator C_Γ given by

$$(1.1) \quad C_\Gamma f(x) := p.v. \int_{\mathbb{R}} \frac{f(y)}{(x-y) + i(A(x) - A(y))} dy.$$

The goal of this article is to give a new proof of the well known theorem of Coifman, McIntosh, and Meyer [5], which says that C_Γ extends naturally as a bounded linear operator from L^p into L^p for any $1 < p < \infty$. Moreover, the method of proof will also allow us to obtain several new generalizations of this important theorem, which will be described in the last section of the paper.

As is known, standard arguments reduce this problem to the problem of proving polynomial bounds for the associated Calderón commutators defined by

$$(1.2) \quad C_d f(x) := p.v. \int_{\mathbb{R}} \frac{(A(x) - A(y))^d}{(x-y)^{d+1}} f(y) dy.$$

Mathematics Subject Classification (2010): Primary 42; Secondary 35.

Keywords: Cauchy integral on Lipschitz curves, Littlewood–Paley projections, logarithmic estimates, polynomial upper bounds.

More precisely, it is enough to prove that

$$(1.3) \quad \|C_d f\|_p \leq C(d) \cdot C(p) \cdot \|f\|_p \cdot \|A'\|_\infty^d$$

for any $f \in L^p$, where $C(d)$ grows at most polynomially in d .

Simple and standard calculations, similar to the ones in [9], show that for Schwartz functions $a := A'$ and f , (1.2) exists and can be rewritten as

$$(1.4) \quad \int_{\mathbb{R}^{d+1}} \left[\int_{[0,1]^d} \operatorname{sgn}(\xi + \alpha_1 \xi_1 + \dots + \alpha_d \xi_d) d\alpha_1 \dots d\alpha_d \right] \cdot \widehat{f}(\xi) \widehat{a}(\xi_1) \dots \widehat{a}(\xi_d) e^{2\pi i x(\xi + \xi_1 + \dots + \xi_d)} d\xi d\xi_1 \dots d\xi_d.$$

As a consequence, C_d can be seen as a $(d + 1)$ -linear operator. More specifically, it is given by the map

$$(f, g_1, \dots, g_d) \rightarrow \int_{\mathbb{R}^{d+1}} \left[\int_{[0,1]^d} \operatorname{sgn}(\xi + \alpha_1 \xi_1 + \dots + \alpha_d \xi_d) d\alpha_1 \dots d\alpha_d \right] \cdot \widehat{f}(\xi) \widehat{g}_1(\xi_1) \dots \widehat{g}_d(\xi_d) e^{2\pi i x(\xi + \xi_1 + \dots + \xi_d)} d\xi d\xi_1 \dots d\xi_d.$$

However, since its symbol

$$(1.5) \quad m_d(\xi, \xi_1, \dots, \xi_d) := \int_{[0,1]^d} \operatorname{sgn}(\xi + \alpha_1 \xi_1 + \dots + \alpha_d \xi_d) d\alpha_1 \dots d\alpha_d$$

is not a classical Marcinkiewicz–Mihlin–Hörmander symbol, there are no estimates for C_d that can be easily passed to the multilinear theorem of Coifman and Meyer [4] and this is why proving (1.3) even without polynomial bounds, is a more delicate problem than estimating paraproducts. In [9], we gave a new proof of (1.3) in the particular case of the first Calderón commutator C_1 . The proof was based on the observation that even though $m_1(\xi, \xi_1)$ is not a classical symbol, when *smoothly* restricted to Whitney squares (with respect to the origin) the Fourier coefficients of the corresponding functions decay at least quadratically. This fact, together with the logarithmic bounds for the shifted Hardy–Littlewood maximal functions and Littlewood–Paley square functions (also proved in [9]), were enough to reduce the problem to a setting where the method developed in [11] and [12] could be applied. It is on the other hand not difficult to realize, that even if one were to assume that such quadratic estimates hold true in the general case of $m_d(\xi, \xi_1, \dots, \xi_d)$, these observations alone would not be enough to obtain (1.3), since then one would end up summing $O(d)$ power series, which would finally generate an exponential upper bound of type C^d . The main new idea to obtain the desired polynomial bounds in (1.3), is to realize that instead of treating m_d as a *whole multiplier* of $d + 1$ variables, one can see it as being a multiple average of various m_1 *type* multipliers. In other words, throughout this paper, we will never need to go beyond the understanding of the symbol of the first Calderón commutator, to be able to obtain polynomial bounds for all the other commutators. This may seem surprising at first glance, but could also be seen as an *explanation*

of a somewhat similar observation of Verdera [14] who showed that in a certain sense, the Cauchy integral *is dominated* by the first Calderón commutator.

Now, coming back to (1.3), we will prove the following.

Theorem 1.1. *Let $1 < p_1, \dots, p_{d+1} \leq \infty$ and $1 \leq p < \infty$ be so that $1/p_1 + \dots + 1/p_{d+1} = 1/p$. Denote by l the number of indices i for which $p_i \neq \infty$. Then, C_d extends naturally as a $(d + 1)$ -linear bounded operator from $L^{p_1} \times \dots \times L^{p_{d+1}} \rightarrow L^p$ with an operator bound of type*

$$(1.6) \quad C(d) \cdot C(l) \cdot C(p_1) \cdot \dots \cdot C(p_{d+1}),$$

where $C(d)$ grows at most polynomially in d and $C(p_i) = 1$ as long as $p_i = \infty$ for $1 \leq i \leq d + 1$.

Assuming for a moment Theorem 1.1, we see immediately that our desired estimate (1.3) follows from it by taking $p_1 = p$ and $p_2 = \dots = p_{d+1} = \infty$.

We now recall that since C_d is a $(d + 1)$ -linear operator, it has $(d + 1)$ natural adjoints. To define them, recall the definition of the associated $(d + 2)$ -linear form Λ_d given by

$$(1.7) \quad \int_{\mathbb{R}} C_d(f_1, \dots, f_{d+1})(x) f_{d+2}(x) dx = \Lambda_d(f_1, \dots, f_{d+2}),$$

where all the functions involved are Schwartz functions. Then, for every $1 \leq i \leq d + 1$, one defines C_d^{*i} by

$$(1.8) \quad \int_{\mathbb{R}} C_d^{*i}(f_1, \dots, f_{i-1}, f_{i+1}, \dots, f_{d+2})(x) f_i(x) dx = \Lambda_d(f_1, \dots, f_{d+2}).$$

For symmetry, we also use the notation $C_d = C_d^{*d+2}$.

To prove Theorem 1.1, we will show that for every $1 \leq i \leq d + 2$ and for all Schwartz functions $\phi_1, \dots, \phi_{d+1}$, one has

$$(1.9) \quad \|C_d^{*i}(\phi_1, \dots, \phi_{d+1})\|_p \leq C(d) \cdot C(l) \cdot C(p_1) \cdot \dots \cdot C(p_{d+1}) \cdot \|\phi_1\|_{p_1} \cdot \dots \cdot \|\phi_{p_{d+1}}\|_{d+1},$$

where p_j for $1 \leq j \leq d + 1$ and p are as before. From this one can immediately extend C_d^{*i} by density to an arbitrary product of L^{p_j} spaces, as long as $1 < p_j < \infty$ and to $\overline{\mathfrak{S}^\infty}$ spaces (the closure of the family of Schwartz functions in L^∞) in the case when $p_j = \infty$.

In the next section we will explain how one can then use duality arguments, to extend the definition of C_d^{*i} even further, to generic products of L^{p_j} and L^∞ spaces. These duality arguments will also clarify the necessity of proving the wider range of estimates which appears in Theorem 1.1 and (1.9) for C_d and its adjoints.

Acknowledgements. The author wishes to thank the referees for their careful corrections, which helped to improve the presentation.

2. Duality and the extension from $\overline{\mathcal{S}^\infty}$ to L^∞

For any $\# = 0, 1, \dots, d$ denote by $S(\#)$ the statement that the inequalities (1.9) for C_d and all its adjoints, extend naturally to the situation when at most $\#$ of the L^{p_j} spaces are equal to L^∞ and the rest are either \mathcal{S}^∞ or correspond to an index j for which $1 < p_j < \infty$. The goal will be to prove that $S(d)$ holds true, assuming $S(0)$ (which we shall prove later). To show that $S(\#)$ implies $S(\# + 1)$, we fix some indices $1 < p_1, \dots, p_{d+1} \leq \infty$ as in the hypothesis of Theorem 1.1. Since the argument is completely symmetric (in particular all the adjoints can be treated similarly) we can assume without loss of generality that we want to extend (1.9) for C_d , when the first $\# + 1$ functions $f_1, \dots, f_{\#+1}$ belong to L^∞ while the other $\phi_{\#+2}, \dots, \phi_{d+1}$ are Schwartz functions.

Case 1: $p > 1$. Since in this case $(L^p)^* = L^{p'}$ for $1/p + 1/p' = 1$, one can simply use duality and define $C_d(f_1, \dots, f_{\#+1}, \phi_{\#+2}, \dots, \phi_{d+1})$ to be the unique L^p function with the property that

$$\begin{aligned} \int_{\mathbb{R}} C_d(f_1, \dots, f_{\#+1}, \phi_{\#+2}, \dots, \phi_{d+1})(x) \phi_{d+2}(x) dx \\ = \int_{\mathbb{R}} C_d^{*1}(f_2, \dots, f_{\#+1}, \phi_{\#+2}, \dots, \phi_{d+1}, \phi_{d+2})(x) f_1(x) dx \end{aligned}$$

for any Schwartz function ϕ_{d+2} with $\|\phi_{d+2}\|_{p'} = 1$. This is clearly well defined as a consequence of $S(\#)$ for C_d^{*1} .

Case 2: $p = 1$. In this case one has to be a bit careful since the dual of L^1 is L^∞ and the Schwartz functions are not dense in L^∞ . However, we first observe that since $p = 1$, there must be at least two indices j_1 and j_2 for which $1 < p_{j_1}, p_{j_2} < \infty$. Again, by the symmetry of the argument, assume that these indices are precisely $\# + 2$ and $\# + 3$. To define $C_d(f_1, \dots, f_{\#+1}, \phi_{\#+2}, \dots, \phi_{d+1})$ properly as an element of L^1 , we first observe that one can define it as an element of (say) L^2 , by taking advantage of the fact that all the functions ϕ_j are Schwartz and therefore belong to all the L^q spaces, for $1 < q < \infty$. Indeed, one can for instance think of $\phi_{\#+2}, \phi_{\#+3}$ as being in L^4 while the rest of Φ_j lie in $\overline{\mathcal{S}^\infty}$ and then define $C_d(f_1, \dots, f_{\#+1}, \phi_{\#+2}, \dots, \phi_{d+1})$ as being the unique function in L^2 with the property that

$$\begin{aligned} \int_{\mathbb{R}} C_d(f_1, \dots, f_{\#+1}, \phi_{\#+2}, \dots, \phi_{d+1})(x) \phi_{d+2}(x) dx \\ = \int_{\mathbb{R}} C_d^{*1}(f_2, \dots, f_{\#+1}, \phi_{\#+2}, \dots, \phi_{d+1}, \phi_{d+2})(x) f_1(x) dx \end{aligned}$$

exactly as before, for any ϕ_{d+2} Schwartz function with $\|\phi_{d+2}\|_2 = 1$, since one can rely again on $S(\#)$ for C_d^{*1} .

Now that we know that $C_d(f_1, \dots, f_{\#+1}, \phi_{\#+2}, \dots, \phi_{d+1})$ is a well defined L^2 function, we would like to prove that it is in fact in L^1 , as desired. One can write,

for any large $M > 0$,

$$(2.1) \quad \int_{-M}^M |C_d(f_1, \dots, f_{\#+1}, \phi_{\#+2}, \dots, \phi_{d+1})(x)| dx \\ = \int_{\mathbb{R}} C_d(f_1, \dots, f_{\#+1}, \phi_{\#+2}, \dots, \phi_{d+1})(x) \tilde{\chi}_{[-M, M]}(x) dx,$$

where $|\tilde{\chi}_{[-M, M]}(x)| = \chi_{[-M, M]}(x)$ almost everywhere.

Now pick a smooth and compactly supported sequence $(f_{d+2}^n)_n$ so that $f_{d+2}^n \rightarrow \tilde{\chi}_{[-M, M]}(x)$ weakly and so that $\|f_{d+2}^n\|_\infty \leq 1$ (one can simply convolve $\tilde{\chi}_{[-M, M]}$ with a smooth approximation of the identity, to obtain such a sequence). In particular, one can then majorize (2.1) by

$$\lim_n \left| \int_{\mathbb{R}} C_d(f_1, \dots, f_{\#+1}, \phi_{\#+2}, \dots, \phi_{d+1})(x) f_{d+2}^n(x) dx \right| \\ \leq \sup_n \left| \int_{\mathbb{R}} C_d^{*1}(f_2, \dots, f_{\#+1}, \phi_{\#+2}, \dots, \phi_{d+1}, f_{d+2}^n)(x) f_1(x) dx \right|,$$

and since now $f_{d+2}^n \in \overline{\mathcal{S}^\infty}$ and $\|f_{d+2}^n\|_\infty \leq 1$, one can again use the induction hypothesis to complete the argument.

Also, a careful look at the whole duality procedure shows that if we assume (1.9) with $C(d)$ growing polynomially, then this will be preserved after replacing all the $\overline{\mathcal{S}^\infty}$ by the corresponding L^∞ .

We are thus left with proving (1.9) for C_d and its adjoints. The advantage of this is that when applied to Schwartz functions, all the operators C_d^{*i} for $1 \leq i \leq d + 2$, are given by well defined expressions similar to (1.4). Later, they will be decomposed and discretized carefully, and this will allow us to reduce (1.9) even more, to some similar estimates, but for finite and well localized *model operators*.

3. Logarithmic estimates and discrete models

In this section the goal is to describe some logarithmic estimates for certain very concrete discrete model operators, which will play an important role in proving the desired (1.9).

In order to motivate them, and also to get a general idea of the strategy of the proof, let us assume for simplicity that instead of (1.9), one would like to prove $L^p \times L^\infty \times \dots \times L^\infty \rightarrow L^p$ estimates (say) for a generic paraproduct $\Pi_{d+1}(f_1, \dots, f_{d+1})$ whose $(d + 2)$ -linear form is given by

$$(3.1) \quad \int_{\mathbb{R}} \sum_k (f_1 * \Phi_k^1)(x) \cdots (f_{d+2} * \Phi_k^{d+2})(x) dx,$$

where $f_1 \in L^p$, $f_j \in L^\infty$ for $2 \leq j \leq d + 1$, while $f_{d+2} \in L^{p'}$ with $1/p + 1/p' = 1$.

As usual, all the functions $(\Phi_k^j)_k$ are smooth L^1 normalized bumps, adapted to intervals of the form $[-2^{-k}, 2^{-k}]$ for $k \in \mathbb{Z}$, and for at least two indices j_1 and j_2 one has $\int_{\mathbb{R}} \Phi_k^{j_1}(x) dx = \int_{\mathbb{R}} \Phi_k^{j_2}(x) dx = 0$ (in which case we say that their corresponding families are of Ψ type, while the others are of Φ type).

We consider several situations.

Case A: $j_1 = 1$ and $j_2 = d + 2$.

We also assume that the L^1 norms of the functions in the Φ families are not only uniformly bounded, but they are bounded by 1. In particular, for any $2 \leq j \leq d+1$ (in this case) one has

$$(3.2) \quad |f_j * \Phi_k^j(x)| \leq \|\Phi_k^j\|_1 \cdot \|f_j\|_\infty \leq \|f_j\|_\infty.$$

One can then majorize (3.1) by

$$\begin{aligned} & \prod_{j=2}^{d+1} \|f_j\|_\infty \cdot \int_{\mathbb{R}} \sum_k |f_1 * \Phi_k^1(x)| |f_{d+2} * \Phi_k^{d+2}(x)| dx \\ & \leq \prod_{j=2}^{d+1} \|f_j\|_\infty \cdot \int_{\mathbb{R}} \left(\sum_k |f_1 * \Phi_k^1(x)|^2 \right)^{1/2} \cdot \left(\sum_k |f_{d+2} * \Phi_k^{d+2}(x)|^2 \right)^{1/2} dx \\ & = \prod_{j=2}^{d+1} \|f_j\|_\infty \cdot \int_{\mathbb{R}} S(f_1)(x) S(f_{d+2})(x) dx \\ & \leq \prod_{j=2}^{d+1} \|f_j\|_\infty \cdot \|S(f_1)\|_p \cdot \|S(f_{d+2})\|_{p'} \lesssim \prod_{j=2}^{d+1} \|f_j\|_\infty \cdot \|f_1\|_p \cdot \|f_{d+2}\|_{p'} \end{aligned}$$

as desired, by using the fact that the Littlewood–Paley square function S is a bounded operator on any L^q space, for $1 < q < \infty$.

Case B: $j_1 = 1$ and $j_2 = 2$.

This case is not so simple. This time, one can only majorize (3.1) by

$$\begin{aligned} & \prod_{j=3}^{d+1} \|f_j\|_\infty \cdot \int_{\mathbb{R}} \sum_k |f_1 * \Phi_k^1(x)| |f_2 * \Phi_k^2(x)| |f_{d+2} * \Phi_k^{d+2}(x)| dx \\ & \leq \prod_{j=3}^{d+1} \|f_j\|_\infty \cdot \int_{\mathbb{R}} \left(\sum_k |f_1 * \Phi_k^1(x)|^2 \right)^{1/2} \cdot \left(\sum_k |f_2 * \Phi_k^2(x)|^2 \right)^{1/2} \\ & \quad \cdot \left(\sup_k |f_{d+2} * \Phi_k^{d+2}(x)| \right) dx \\ & = \prod_{j=3}^{d+1} \|f_j\|_\infty \cdot \int_{\mathbb{R}} S(f_1)(x) S(f_2)(x) M(f_{d+2})(x) dx \\ & \leq \prod_{j=3}^{d+1} \|f_j\|_\infty \cdot \|S(f_1)\|_{s_1} \cdot \|S(f_2)\|_{s_2} \cdot \|M(f_{d+2})\|_{s'_3} \\ & \lesssim \prod_{j=3}^{d+1} \|f_j\|_\infty \cdot \|f_1\|_{s_1} \cdot \|f_2\|_{s_2} \cdot \|f_{d+2}\|_{s'_3}, \end{aligned}$$

for any $1 < s_1, s_2, s_3 < \infty$ so that $1/s_1 + 1/s_2 = 1/s_3$, by using the fact that besides S , the Hardy–Littlewood maximal function M is also bounded on any L^q space, for $1 < q < \infty$. Clearly, the estimate we are looking for corresponds to $s_1 = s_3 = p$ and $s_2 = \infty$ and it cannot be obtained in this way, since S is unbounded on L^∞ .

If, on the other hand, one freezes the functions f_3, \dots, f_{d+1} , the expression (3.1) becomes a 3-linear form and the above estimates show that its associated bilinear operator $\Pi_2(f_1, f_2)$ is bounded from $L^{s_1} \times L^{s_2}$ into L^{s_3} . By symmetry, the same is true for both Π_2^{*1} and Π_2^{*2} . The estimate we are interested in can then be rephrased as

$$(3.3) \quad \Pi_2 : L^p \times L^\infty \rightarrow L^p.$$

To get it one needs to prove, besides the previous Banach estimates, quasi-Banach estimates, of the form $\Pi_2^{*2} : L^{r_1} \times L^{r_2} \rightarrow L^{r_3}$, for any $1 < r_1, r_2 < \infty, 0 < r_3 < \infty$ with $1/r_1 + 1/r_2 = 1/r_3$. In the case of paraproducts, there are several ways to achieve this; see for instance [4]. In the end, one can use multi-linear interpolation between the Banach and quasi-Banach estimates as in [13], to obtain the intermediate (3.3). Even more precisely, the convexity argument in [13] shows that there exist two Banach estimates and one quasi-Banach estimate (with implicit boundedness constants C_B^1, C_B^2 , and C_{q-B} , respectively) so that if one writes $C_B := \max\{C_B^1, C_B^2, C_{q-B}\}$, one has that this constant represents an upper bound for the boundedness constant of (3.3).

Case C: $j_1 = 2$ and $j_2 = 3$.

This final remaining situation can be treated similarly to the previous one. Finally, we are left with this situation which can be treated similarly to the previous one. More precisely, one can majorize (3.1) by

$$\begin{aligned} & \prod_{j=4}^{d+1} \|f_j\|_\infty \cdot \int_{\mathbb{R}} \sum_k |f_1 * \Phi_k^1(x)| |f_2 * \Phi_k^2(x)| |f_3 * \Phi_k^3(x)| |f_{d+2} * \Phi_k^{d+2}(x)| dx \\ & \leq \prod_{j=4}^{d+1} \|f_j\|_\infty \cdot \int_{\mathbb{R}} \left(\sum_k |f_2 * \Phi_k^2(x)|^2 \right)^{1/2} \left(\sum_k |f_3 * \Phi_k^3(x)|^2 \right)^{1/2} \\ & \quad \cdot \left(\sup_k |f_1 * \Phi_k^1(x)| \right) \left(\sup_k |f_{d+2} * \Phi_k^{d+2}(x)| \right) dx \\ & = \prod_{j=4}^{d+1} \|f_j\|_\infty \cdot \int_{\mathbb{R}} S(f_2)(x) S(f_3)(x) M(f_1)(x) M(f_{d+2})(x) dx \\ & \leq \prod_{j=4}^{d+1} \|f_j\|_\infty \cdot \|M(f_1)\|_{s_1} \cdot \|S(f_2)\|_{s_2} \cdot \|S(f_3)\|_{s_3} \cdot \|M(f_{d+2})\|_{s_4'} \\ & \lesssim \prod_{j=4}^{d+1} \|f_j\|_\infty \cdot \|f_1\|_{s_1} \cdot \|f_2\|_{s_2} \cdot \|f_3\|_{s_3} \|f_{d+2}\|_{s_4'}, \end{aligned}$$

for any $1 < s_1, s_2, s_3, s_4 < \infty$ so that $1/s_1 + 1/s_2 + 1/s_3 = 1/s_4$.

The estimate we are looking for corresponds to $s_1 = s_4 = p$ and $s_2 = s_3 = \infty$ and as before, it cannot be obtained in this way.

This time one freezes the functions f_4, \dots, f_{d+1} and then the expression (3.1) becomes a 4-linear form and the estimates above show that its associated 3-linear operator $\Pi_3(f_1, f_2, f_3)$ is bounded from $L^{s_1} \times L^{s_2} \times L^{s_3}$ into L^{s_4} . By symmetry, the same is true for its adjoints Π_3^{*1}, Π_3^{*2} and Π_3^{*3} . The estimate we are interested in becomes

$$(3.4) \quad \Pi_3 : L^p \times L^\infty \times L^\infty \rightarrow L^p.$$

As in the previous case, to get it, one needs to prove, besides the previous Banach estimates, quasi-Banach estimates of the form $\Pi_3^{*2}, \Pi_3^{*3} : L^{r_1} \times L^{r_2} \times L^{r_3} \rightarrow L^{r_4}$, for any $1 < r_1, r_2, r_3 < \infty, 0 < r_4 < \infty$ with $1/r_1 + 1/r_2 + 1/r_3 = 1/r_4$. In the end, one can again use multilinear interpolation between the Banach and quasi-Banach estimates, to obtain the intermediate (3.4). In particular, the same convexity argument shows that there exist two Banach and two quasi-Banach estimates (with implicit boundedness constants C_B^1, C_B^2, C_{q-B}^1 , and C_{q-B}^2 , respectively) so that if one denotes by $C_B := \max\{C_B^1, C_B^2, C_{q-B}^1, C_{q-B}^2\}$, one has that this constant is an upper bound for the boundedness constant of (3.4). This ends the discussion on the boundedness of Π_{d+1} from $L^p \times L^\infty \times \dots \times L^\infty$ into L^p since, by symmetry, it is easy to realize that any other possibility can be reduced to one of these cases.¹

There are a couple of important facts that one learns from the previous argument. First, the bounds are independent of d . The reason for this is the assumption that the L^1 norms of the Φ families are all at most 1, which implied the crucial (3.2). Then, there is the fact that after using (3.2) several times, we reduced our analysis to the study of several (Banach or quasi-Banach) corresponding estimates, for some minimal bilinear or trilinear operators.

We claim now that in spite of the fact that C_d is not a Coifman–Meyer operator, it can be studied in an analogous manner. More precisely, one can decompose it first into polynomially (in d) many *paraproduct-like pieces*, and then estimate each piece independently of d . Also as before (since the Banach estimates are easy) we will reduce the main inequality (via interpolation), to similar quasi-Banach estimates for minimal l -linear operators, for some $1 \leq l \leq d + 1$. The proof of the precise quasi-Banach estimates is in general a delicate issue, but it has already been discussed in detail in [9].

The necessity of discretizing the minimal $(l + 1)$ -linear forms justifies the introduction of the following model operators.

Fix then a positive integer l and arbitrary integers $\bar{n}_1, \dots, \bar{n}_l$ and consider families $(\Phi_{I_{\bar{n}_1}}^1)_I, \dots, (\Phi_{I_{\bar{n}_l}}^l)_I$ of L^2 normalized bumps adapted to dyadic intervals $I_{\bar{n}_j}$ (as in [9], given I , denote by $I_{\bar{n}_j}$ the interval of the same length as I , but sitting \bar{n}_j units of length $|I|$ away from I) so that at least two of them are of Ψ type (i.e., their integrals are zero). By definition, a smooth function Φ is said to be adapted

¹It should also be clear that a similar argument works in the general $\Pi_{d+1} : L^{p_1} \times \dots \times L^{p_{d+1}} \rightarrow L^p$ case. Instead of the minimal bilinear or trilinear operators which appeared before, one would have to deal with l -linear ones for some $1 \leq l \leq d + 1$, but the interpolation between the natural corresponding Banach and quasi-Banach estimates works in precisely the same way.

to an interval I , if one has

$$|\partial^\alpha \Phi(x)| \lesssim \frac{1}{|I|^{|\alpha|}} \frac{1}{(1 + \text{dist}(x, I)/|I|)^M}$$

for any derivative α so that $|\alpha| \leq 5$ and any large $M > 0$, with the implicit constants depending on it. Then, also by definition, if $1 \leq p \leq \infty$, we say that $|I|^{-1/p} \Phi$ is L^p normalized.

Define the l -linear discrete operator $T_{\mathcal{J}}$, for a finite family \mathcal{J} of dyadic intervals, by

$$(3.5) \quad T_{\mathcal{J}}(f_1, \dots, f_l) = \sum_{I \in \mathcal{J}} \frac{1}{|I|^{(l-2)/2}} \langle f_1, \Phi_{I\bar{n}_1}^1 \rangle \cdots \langle f_l, \Phi_{I\bar{n}_l}^l \rangle \Phi_I^{l+1}.$$

One has:

Theorem 3.1. *For any such a family \mathcal{J} , the l -linear operator $T_{\mathcal{J}}$ maps $L^{p_1} \times \cdots \times L^{p_l} \rightarrow L^p$ for any $1 < p_1, \dots, p_l < \infty$ with $1/p_1 + \cdots + 1/p_l = 1/p$, $0 < p < \infty$, with a bound of type*

$$O(\log \langle \bar{n}_1 \rangle \cdots \log \langle \bar{n}_l \rangle).$$

Here, as in [9], $\langle \bar{n}_j \rangle$ simply denotes $2 + |\bar{n}_j|$. Also, the implicit constants above are allowed to depend on l .

This theorem is the l -linear generalization of the bilinear Theorem 2 in [9] and since its proof is identical to the proof of that theorem, we leave it to the reader.

More precisely (as in [9]), Theorem 3.1 follows (by scale invariance and interpolation) from the more precise statement that for every $f_j \in L^{p_j}$ with $\|f_j\|_{p_j} = 1$, $1 \leq j \leq l$ and measurable set $E \subseteq \mathbb{R}$ of measure 1, there exists a subset $E' \subseteq E$ of comparable measure so that

$$(3.6) \quad \sum_{J \in \mathcal{J}} \frac{1}{|J|^{(l-1)/2}} |\langle f_1, \Phi_{J\bar{n}_1}^1 \rangle| \cdots |\langle f_l, \Phi_{J\bar{n}_l}^l \rangle| |\langle f_{l+1}, \Phi_J^{l+1} \rangle| \lesssim \log \langle \bar{n}_1 \rangle \cdots \log \langle \bar{n}_l \rangle,$$

where $f_{l+1} = \chi_{E'}$.

As in [9], the fact that one loses only logarithmic bounds in the estimates above will be important later on.

In the rest of the paper we will describe the calculations that are necessary to show how the desired (1.9) can indeed be reduced to (3.6).

4. Reduction to the discrete model

We treat the case of C_d only, since by the symmetry of the arguments, all its adjoints can be understood in a similar way. Fix indices p_j for $1 \leq j \leq d + 1$ and p as in the hypothesis of Theorem 1.1. As suggested before, the first step is to decompose C_d into polynomially (in d) many *paraproduct-like pieces* which will be analyzed afterwards.

Here, the classical Littlewood–Paley decompositions will be of great help. However, since we want to have the *perfect inequalities* (3.2) available, we need to work most of the time with *noncompact* (in frequency) approximations of the identity, which will cause several technical difficulties later on. We define them in detail in the next subsection.

4.1. Noncompact Littlewood–Paley L^1 normalized projections

Start with a Schwartz function $\Phi(x)$ which is even, positive, and such that

$$\int_{\mathbb{R}} \Phi(x) dx = 1.$$

Define also $\Psi(x)$ by

$$\Psi(x) = \Phi(x) - \frac{1}{2} \Phi(x/2)$$

and observe that $\int_{\mathbb{R}} \Psi(x) dx = 0$. Then, as usual, $\Psi_k(x)$ and $\Phi_k(x)$ denote $2^k \Psi(2^k x)$ and $2^k \Phi(2^k x)$, respectively. Notice that all the L^1 norms of Φ_k are equal to 1. Observe also that one has

$$\Psi_k(x) = \Phi_k(x) - \Phi_{k-1}(x),$$

and then it is easy to see that

$$\sum_{k \leq \bar{k}} \Psi_k = \Phi_{\bar{k}},$$

and also that

$$(4.1) \quad \sum_{k \in \mathbb{Z}} \Psi_k = \delta_0,$$

or, equivalently,

$$(4.2) \quad \sum_{k \in \mathbb{Z}} \widehat{\Psi}_k(\xi) = 1$$

for almost every $\xi \in \mathbb{R}$. On the other hand,

$$\widehat{\Psi}(0) = \int_{\mathbb{R}} \Psi(x) dx = 1 - 1 = 0.$$

Moreover, one also has that

$$\widehat{\Psi}'(0) = -2\pi i \int_{\mathbb{R}} x \Psi(x) dx = 0$$

by using the fact that Φ is an even function. As a consequence, one can write $\widehat{\Psi}(\xi)$ as

$$(4.3) \quad \widehat{\Psi}(\xi) = \xi^2 \widehat{\phi}(\xi)$$

for another smooth and rapidly decaying function ϕ .

These are the *noncompact* L^1 normalized Littlewood–Paley decompositions. The *compact* decompositions are obtained in a similar manner, but one starts instead with a Schwartz function Φ having the properties that $\text{supp } \widehat{\Phi} \subseteq [-1, 1]$ and $\widehat{\Phi}(\xi) = 1$ on the subinterval $[-1/2, 1/2]$.

4.2. Some remarks on the symbols of C_d for $d \geq 2$

Before proceeding, it is worthwhile to examine the symbol of the second commutator C_2 . It is very natural to try to see if its Fourier coefficients satisfy the same *quadratic estimates* (proved in [9]) as the symbol of C_1 . Consider for instance three Schwartz functions $\widehat{\phi}(\xi)$, $\widehat{\phi}(\xi_1)$ and $\widehat{\phi}(\xi_2)$ supported in the intervals $[-2, -1]$, $[1, 2]$ and $[-1/2, 1/2]$ respectively. Clearly, the function

$$(\xi, \xi_1, \xi_2) \rightarrow \widehat{\phi}(\xi) \widehat{\phi}(\xi_1) \widehat{\phi}(\xi_2)$$

is supported in a Whitney cube (with respect to the origin) in \mathbb{R}^3 and the goal is to understand the expression

$$(4.4) \quad \int_{\mathbb{R}^3} \left[\int_{[0,1]^2} 1_{\mathbb{R}_+}(\xi + \alpha\xi_1 + \beta\xi_2) \, d\alpha \, d\beta \right] \widehat{\phi}(\xi) \widehat{\phi}(\xi_1) \widehat{\phi}(\xi_2) \cdot e^{-2\pi i n \xi} e^{-2\pi i n_1 \xi_1} e^{-2\pi i n_2 \xi_2} \, d\xi \, d\xi_1 \, d\xi_2$$

when n, n_1 , and n_2 are arbitrary integers. Since ξ_1 can never be zero, the symbol in (4.4) can be rewritten as

$$(4.5) \quad \int_0^1 \frac{1}{\xi_1} \int_0^{\xi_1} 1_{\mathbb{R}_+}(\xi + \alpha + \beta\xi_2) \, d\alpha \, d\beta.$$

When one differentiates (4.5) with respect to ξ_1 the inner term becomes

$$\int_0^1 1_{\mathbb{R}_+}(\xi + \xi_1 + \beta\xi_2) \, d\beta,$$

which coincides with $m_1(\xi + \xi_1, \xi_2)$. The important difference now is that since $\xi + \xi_1$ lies in the interval $[-1, 1]$ and ξ_2 in $[-1/2, 1/2]$ and they both contain the origin, one can no longer apply the argument in [9].

As a consequence of this example, the Fourier coefficients in (4.4) seem to decay only *linearly*, which is clearly not enough. In any event, these comments show that passing from the analysis of the first commutator to the analysis of the second one and all the rest, is not at all a *routine* task. One should recall that ten years passed between the results of Calderón [1] and those of Coifman and Meyer [3].

On the other hand, this also shows that, from this point of view at least, the symbol of C_2 looks similar to the symbol of the bilinear Hilbert transform (given by $\text{sgn}(\xi_1 - \xi_2)$) whose Fourier coefficients also decay only *linearly*, as one can easily check. This fact might also be considered as another *possible explanation* of why Calderón suggested the study of the bilinear Hilbert transform as a step towards understanding all his commutators, in addition to the one proposed in [9].

4.3. The generic decomposition of C_d

Coming back to our goal, notice first that, because of (1.4) if f, f_1, \dots, f_{d+1} are all Schwartz functions, one can write the $(d + 2)$ -linear form $\Lambda_d(f, f_1, \dots, f_{d+1})$ associated with C_d as

$$(4.6) \quad \int_{\xi+\xi_1+\dots+\xi_{d+1}=0} \left[\int_{[0,1]^d} 1_{\mathbb{R}^+}(\xi + \alpha_1 \xi_1 + \dots + \alpha_d \xi_d) d\alpha_1 \dots d\alpha_d \right] \cdot \widehat{f}(\xi) \widehat{f}_1(\xi_1) \dots \widehat{f}_{d+1}(\xi_{d+1}) d\xi d\xi_1 \dots d\xi_{d+1}.$$

By combining several Littlewood–Paley decompositions as in (4.2), one can write

$$(4.7) \quad 1 = \sum_{k_0, k_1, \dots, k_d, k_{d+1} \in \mathbb{Z}} \widehat{\Psi}_{k_0}(\xi) \widehat{\Psi}_{k_1}(\xi_1) \dots \widehat{\Psi}_{k_d}(\xi_d) \widehat{\Psi}_{k_{d+1}}(\xi_{d+1}).$$

Now, for every $(d + 2)$ -tuple $(k_0, k_1, \dots, k_d, k_{d+1}) \in \mathbb{Z}^{d+2}$, one has one of $k_0 \geq k_1, \dots, k_d, k_{d+1}$, $k_1 \geq k_0, k_2, \dots, k_{d+1}, \dots$ or $k_{d+1} \geq k_0, k_1, \dots, k_d$. By replacing some of these inequalities with *strict inequalities*, one can also ensure that the corresponding $d + 2$ regions in \mathbb{Z}^{d+2} are all disjoint.

Fixing always the biggest parameter and summing over the rest, one can rewrite the constant 1 in (4.7) as²

$$(4.8) \quad \sum_k \widehat{\Psi}_k(\xi) \widehat{\Phi}_k(\xi_1) \dots \widehat{\Phi}_k(\xi_d) \widehat{\Phi}_k(\xi_{d+1}) + \dots + \sum_k \widehat{\Phi}_k(\xi) \widehat{\Phi}_k(\xi_1) \dots \widehat{\Phi}_k(\xi_d) \widehat{\Psi}_k(\xi_{d+1}).$$

There are $d + 2$ inner terms in the decomposition (4.8), each containing a single Ψ type of a function.

For technical reasons that will be clearer later, we assume that for the ξ and ξ_{d+1} variables we use compact Littlewood–Paley decompositions, while for the rest we use noncompact decompositions from (4.2).

We now assume that, in addition, one has $\xi + \xi_1 + \dots + \xi_{d+1} = 0$. Look at the second (for instance) sum in (4.8) and consider the $k = 0$ inner term, which we write for simplicity as

$$(4.9) \quad \widehat{\Phi}(\xi) \widehat{\Psi}(\xi_1) \dots \widehat{\Phi}(\xi_d) \widehat{\Phi}(\xi_{d+1}).$$

Since from (4.3) we know that $\widehat{\Psi}(\xi_1) = \xi_1^2 \widehat{\phi}(\xi_1)$, one can rewrite this as

$$(4.10) \quad \begin{aligned} \widehat{\Psi}(\xi_1) &= \xi_1 \widehat{\phi}(\xi_1) (-\xi - \xi_2 - \dots - \xi_{d+1}) \\ &= -\xi_1 \xi \widehat{\phi}(\xi_1) - \xi_1 \xi_2 \widehat{\phi}(\xi_1) - \dots - \xi_1 \xi_{d+1} \widehat{\phi}(\xi_1). \end{aligned}$$

Using this in (4.9), one can write it as another sum of $O(d)$ terms, containing this time two functions of Ψ type, since besides $\xi_1 \widehat{\phi}(\xi_1)$, one obtains in addition either expressions of type $\xi_j \widehat{\Phi}(\xi_j)$, for $j = 2, \dots, d + 1$, or $\xi \widehat{\Phi}(\xi)$.

If one does this for every scale $k \in \mathbb{Z}$ and every inner term in (4.8), one obtains a decomposition of $1_{\{\xi+\xi_1+\dots+\xi_{d+1}=0\}}$ as a sum of $O(d^2)$ expressions whose generic inner product terms all contain two functions of Ψ type, which are more specifically

²To be totally rigorous, some of these Φ_k functions should be Φ_{k-1} , but this is a minor issue.

of the form $\gamma\widehat{\phi}(\gamma)$.³ If one inserts this into the formula for the $(d + 2)$ -linear form (4.6), one obtains $O(d^2)$ $(d + 2)$ -linear forms which will be carefully analyzed next. This is our *generic decomposition*. To be able to go further, one needs to understand how to *unfold* the symbol of C_d .

As before in the case of paraproducts, the positions of the Ψ functions (we denote them by j_1 , and j_2 for $0 \leq j_1, j_2 \leq d + 1$) will play an important role.

There are in fact three distinct cases, depending on where these Ψ functions lie.

Case 1: $j_1 = 0$ and $j_2 = 1$

For symmetry, we change notation and rewrite the $(d + 2)$ -linear form as

$$(4.11) \quad \sum_k \int_{\xi+\xi_1+\dots+\xi_{d+1}=0} \left[\int_{[0,1]^d} 1_{\mathbb{R}^+}(\xi + \alpha_1 \xi_1 + \dots + \alpha_d \xi_d) d\alpha_1 \dots d\alpha_d \right] \cdot \widehat{\Phi}_k^0(\xi) \widehat{\Phi}_k^1(\xi_1) \dots \widehat{\Phi}_k^d(\xi_d) \widehat{\Phi}_k^{d+1}(\xi_{d+1}) \cdot \widehat{f}(\xi) \widehat{f}_1(\xi_1) \dots \widehat{f}_{d+1}(\xi_{d+1}) d\xi d\xi_1 \dots d\xi_{d+1},$$

with both families $(\widehat{\Phi}_k^0)_k$ and $(\widehat{\Phi}_k^1)_k$ being of Ψ type. Without the C_d symbol above, the expression (4.11) would be the $(d + 2)$ -linear form of a paraproduct, which could be analyzed as we described earlier. The first idea for handling it, is to try to decompose it into multiple Fourier series, on the support of the corresponding Whitney *frequency boxes*. However, it is clear that in the end this will produce this will produce an upper bound of the type of a product of $O(d)$ power series, which will grow exponentially in d even in the case of classical symbols, so one has to be very careful at this stage. The situation is in fact even worse as we pointed out earlier, since it seems that the Fourier coefficients of the symbol of C_d for $d \geq 2$, do not decay quadratically as do those of the symbol of C_1 .

The idea now is to realize that the variable ξ_1 (in this case) is in some sense *special* and the right thing to do is to look at the C_d symbol as being a multiple average of C_1 symbols (depending on ξ_1 and on a new variable $\tilde{\xi}$), which can be analysed as in [9].

To be able to *execute* this plan, since the functions $\widehat{\Phi}_k^j(\xi_j)$ for $1 \leq j \leq d$ do not have compact support, one has to insert still two more compact Littlewood–Paley decompositions of unity in (4.11). More precisely, denote by

$$\tilde{\xi} = \xi + \alpha_2 \xi_2 + \dots + \alpha_d \xi_d$$

and write

$$(4.12) \quad 1 = \sum_{k_0, k_1} \widehat{\Psi}_{k_0}(\tilde{\xi}) \widehat{\Psi}_{k_1}(\xi_1) = \sum_{k_0 \ll k_1} \dots + \sum_{k_0 \sim k_1} \dots + \sum_{k_0 \gg k_1} \dots,$$

which can be rewritten as before as

$$(4.13) \quad \sum_r \widehat{\Phi}_r(\tilde{\xi}) \widehat{\Psi}_r(\xi_1) + \sum_r \widehat{\Psi}_r(\tilde{\xi}) \widehat{\Psi}_r(\xi_1) + \sum_r \widehat{\Psi}_r(\tilde{\xi}) \widehat{\Phi}_r(\xi_1).$$

³As a consequence, the Schwartz functions whose Fourier transforms are given by such expression, have integral zero.

To be totally rigorous, one should write $\widehat{\Phi}_{r-100}$ instead of $\widehat{\Phi}_r$ in (4.13) and also finitely many *middle Ψ - Ψ terms*, instead of only one. While this should be kept in mind, it is left as is for notational simplicity.

As a consequence, if we insert (4.13) into (4.11), it splits as a sum of three distinct expressions which will be analyzed separately. We will denote these cases by 1_a , 1_b , and 1_c , respectively.

Case 1_a

To see the effect of the new splitting (over r) in (4.11), we analyze, for simplicity, the term corresponding to $k = 0$. If we ignore the symbol $\int_{[0,1]^d} 1_{\mathbb{R}^+}(\xi + \alpha_1 \xi_1 + \dots + \alpha_d \xi_d) d\alpha_1 \dots d\alpha_d$, the rest of the expression becomes

$$(4.14) \quad \sum_r [\widehat{\Phi}_r(\tilde{\xi}) \widehat{\Psi}_r(\xi_1)] \widehat{\Phi}_0^0(\xi) \widehat{\Phi}_0^1(\xi_1) \dots \widehat{\Phi}_0^d(\xi_d) \widehat{\Phi}_0^{d+1}(\xi_{d+1}) = \sum_{r \leq 0} \dots + \sum_{r > 0} \dots = 1'_a + 1''_a.$$

Case $1'_a$

Using the fact that $\widehat{\Psi}_r(\xi_1)$ is compactly supported and taking into account the fact that $\widehat{\Phi}_0^1(\xi_1)$ is also of Ψ type (more precisely, as we have seen, it is of the form $\xi_1 \widehat{\phi}(\xi_1)$), one can rewrite $1'_a$ as

$$\sum_{r \leq 0} 2^r \widehat{\Phi}_r(\tilde{\xi}) \widehat{\Phi}_0^0(\xi) \widehat{\Psi}_r^1(\xi_1) \dots \widehat{\Phi}_0^d(\xi_d) \widehat{\Phi}_0^{d+1}(\xi_{d+1}) = \sum_{r \leq 0} 2^r [\widehat{\Phi}_r(\tilde{\xi}) \widehat{\Psi}_r^1(\xi_1)] \widehat{\Phi}_0^0(\xi) \widehat{\Psi}_r^1(\xi_1) \dots \widehat{\Phi}_0^d(\xi_d) \widehat{\Phi}_0^{d+1}(\xi_{d+1}) \widehat{\Phi}_r(\tilde{\xi})$$

for certain well chosen compactly supported functions $\widehat{\Phi}_r(\tilde{\xi})$, $\widehat{\Psi}_r^1(\xi_1)$ and $\widehat{\Psi}_r^1(\xi_1)$ (naturally, the first is of Φ type, while the other two are of Ψ type).

In particular, we can split the symbol

$$\left[\int_0^1 1_{\mathbb{R}^+}(\tilde{\xi} + \alpha_1 \xi_1) d\alpha_1 \right] \widehat{\Phi}_r(\tilde{\xi}) \widehat{\Psi}_r^1(\xi_1)$$

as a double Fourier series of the form

$$(4.15) \quad \sum_{n, n_1} C_{n, n_1}^r e^{2\pi i \frac{n}{2^r} \tilde{\xi}} e^{2\pi i \frac{n_1}{2^r} \xi_1},$$

where

$$\begin{aligned} C_{n, n_1}^r &= \frac{1}{2^r} \frac{1}{2^r} \int_{\mathbb{R}^2} \left[\int_0^1 1_{\mathbb{R}^+}(\tilde{\xi} + \alpha_1 \xi_1) d\alpha_1 \right] \widehat{\Phi}_r(\tilde{\xi}) \widehat{\Psi}_r^1(\xi_1) e^{-2\pi i \frac{n}{2^r} \tilde{\xi}} e^{-2\pi i \frac{n_1}{2^r} \xi_1} d\tilde{\xi} d\xi_1 \\ &= \int_{\mathbb{R}^2} \left[\int_0^1 1_{\mathbb{R}^+}(\tilde{\xi} + \alpha_1 \xi_1) d\alpha_1 \right] \widehat{\Phi}_0(\tilde{\xi}) \widehat{\Psi}_0^1(\xi_1) e^{-2\pi i n \tilde{\xi}} e^{-2\pi i n_1 \xi_1} d\tilde{\xi} d\xi_1, \end{aligned}$$

which is an expression independent of r .

Recall now from [9] the crucial fact that

$$(4.16) \quad |C_{n,n_1}^r| (= |C_{n,n_1}|) \lesssim \frac{1}{\langle n \rangle^2} \frac{1}{\langle n_1 \rangle^\#}$$

for any large number $\# > 0$.

These calculations show that the corresponding contribution of $1'_a$ in (4.11) is

$$(4.17) \quad \int_0^1 \cdots \int_0^1 \sum_{r \leq 0} 2^r \sum_{n,n_1} C_{n,n_1}^r \int_{\xi+\xi_1+\cdots+\xi_d+\xi_{d+1}=0} [\widehat{\Phi}_0^0(\xi) e^{2\pi i \frac{n}{2^r} \xi}] [\widehat{\Psi}_r^1(\xi_1) e^{2\pi i \frac{n_1}{2^r} \xi_1}] \\ \cdot [\widehat{\Phi}_0^2(\xi_2) e^{2\pi i \frac{n\alpha_2}{2^r} \xi_2}] \cdots [\widehat{\Phi}_0^d(\xi_d) e^{2\pi i \frac{n\alpha_d}{2^r} \xi_d}] \widehat{\Phi}_0^{d+1}(\xi_{d+1}) \widehat{\Phi}_r(\tilde{\xi}) \\ \cdot \widehat{f}(\xi) \widehat{f}_1(\xi_1) \cdots \widehat{f}_d(\xi_d) \widehat{f}_{d+1}(\xi_{d+1}) d\xi d\xi_1 \cdots d\xi_{d+1} d\alpha_2 \cdots d\alpha_d.$$

If now one fixes $\alpha_2, \dots, \alpha_d \in [0, 1]$, r, n and n_1 , the inner expression becomes

$$(4.18) \quad \int_{\xi+\xi_1+\cdots+\xi_d+\xi_{d+1}=0} [\widehat{f}(\xi) \widehat{\Phi}_0^0(\xi) e^{2\pi i \frac{n}{2^r} \xi}] \cdot [\widehat{f}_1(\xi_1) \widehat{\Psi}_r^1(\xi_1) e^{2\pi i \frac{n_1}{2^r} \xi_1}] \\ \cdot [\widehat{f}_2(\xi_2) \widehat{\Phi}_0^2(\xi_2) e^{2\pi i \frac{n\alpha_2}{2^r} \xi_2}] \cdots [\widehat{f}_d(\xi_d) \widehat{\Phi}_0^d(\xi_d) e^{2\pi i \frac{n\alpha_d}{2^r} \xi_d}] \\ \cdot [\widehat{f}_{d+1}(\xi_{d+1}) \widehat{\Phi}_0^{d+1}(\xi_{d+1})] \widehat{\Phi}_r(\xi + \alpha_2 \xi_2 + \cdots + \alpha_d \xi_d) d\xi d\xi_1 \cdots d\xi_{d+1}.$$

We need now the following.

Lemma 4.1. *If F, F_1, \dots, F_{d+1} and Φ are Schwartz functions, then one has*

$$(4.19) \quad \int_{\xi+\xi_1+\cdots+\xi_d+\xi_{d+1}=0} \widehat{F}(\xi) \widehat{F}_1(\xi_1) \cdots \widehat{F}_{d+1}(\xi_{d+1}) \\ \cdot \widehat{\Phi}(a\xi + a_1\xi_1 + \cdots + a_{d+1}\xi_{d+1}) d\xi d\xi_1 \cdots d\xi_{d+1} \\ = \int_{\mathbb{R}^2} F(x - at) F_1(x - a_1t) \cdots F_{d+1}(x - a_{d+1}t) \Phi(t) dt dx,$$

for all real numbers a, a_1, \dots, a_{d+1} .

Proof. The formula is based on the following fact. If Γ is a vector subspace of \mathbb{R}^{d+2} and δ_Γ represents the Dirac distribution associated to it and defined by

$$\delta_\Gamma(\phi) = \int_\Gamma \phi(\gamma) d\gamma$$

for every Schwartz function ϕ , then, $\widehat{\delta}_\Gamma = \delta_{\Gamma^\perp}$. In our case,

$$(4.20) \quad \Gamma = \{(\xi, \xi_1, \dots, \xi_{d+1}) \in \mathbb{R}^{d+2} : \xi + \xi_1 + \cdots + \xi_d + \xi_{d+1} = 0\},$$

and as a consequence Γ^\perp is the 1-dimensional subspace along the vector $(1, \dots, 1)$. Using this and the Plancherel identity, the left-hand side of (4.19) can be written as

$$\int_{\mathbb{R}} \widehat{F}(\xi) \widehat{F}_1(\xi_1) \cdots \widehat{F}_{d+1}(\xi_{d+1}) \widehat{\Phi}(a\xi + a_1\xi_1 + \cdots + a_{d+1}\xi_{d+1}) \\ \cdot e^{2\pi i x(\xi+\xi_1+\cdots+\xi_{d+1})} d\xi d\xi_1 \cdots d\xi_{d+1} dx.$$

If one adds to it

$$\widehat{\Phi}(a\xi + a_1\xi_1 + \dots + a_{d+1}\xi_{d+1}) = \int_{\mathbb{R}} \Phi(t) e^{-2\pi it(a\xi + a_1\xi_1 + \dots + a_{d+1}\xi_{d+1})} dt,$$

one immediately obtains (4.19) by using the Fourier inversion formula several times. □

We record also the following generalization of (4.19), which will be used later:

$$\begin{aligned} & \int_{\Gamma} \widehat{F}(\xi) \widehat{F}_1(\xi_1) \cdots \widehat{F}_{d+1}(\xi_{d+1}) \widehat{\Phi}_1(a\xi + a_1\xi_1 + \dots + a_{d+1}\xi_{d+1}) \\ & \quad \cdot \widehat{\Phi}_2(b\xi + b_1\xi_1 + \dots + b_{d+1}\xi_{d+1}) d\xi d\xi_1 \cdots d\xi_{d+1} \\ (4.21) \quad & = \int_{\mathbb{R}^3} F(x - at - bs) F_1(x - a_1t - b_1s) \\ & \quad \cdots F_{d+1}(x - a_{d+1}t - a_{d+1}s) \Phi_1(t) \Phi_2(s) dt ds dx. \end{aligned}$$

Now, if G is an arbitrary Schwartz function and a a real number, we denote by G^a the function defined by

$$(4.22) \quad \widehat{G}^a(\xi) = \widehat{G}(\xi) e^{2\pi ia\xi}.$$

Equivalently, one has $G^a(x) = G(x + a)$.

Using this notation and applying (4.19), our previous (4.18) becomes

$$\begin{aligned} & \int_{\mathbb{R}^2} (f * \Phi_0^{0,n/2^r})(x - t) (f_1 * \Psi_r^{1,n_1/2^r})(x) \\ & \quad \cdot \prod_{j=2}^d (f_j * \Phi_0^{j,n\alpha_j/2^r})(x - \alpha_j t) (f_{d+1} * \Phi_0^{d+1})(x) \Phi_r(t) dt dx \\ & = \int_{\mathbb{R}^2} (f * \Phi_0^{0,n/2^r})(x - t/2^r) (f_1 * \Psi_r^{1,n_1/2^r})(x) \\ & \quad \cdot \prod_{j=2}^d (f_j * \Phi_0^{j,n\alpha_j/2^r})(x - \alpha_j t/2^r) (f_{d+1} * \Phi_0^{d+1})(x) \Phi_0(t) dt dx \\ & = \int_{\mathbb{R}^2} (f * \Phi_0^{0,(n-t)/2^r})(x) (f_1 * \Psi_r^{1,n_1/2^r})(x) \\ (4.23) \quad & \cdot \prod_{j=2}^d (f_j * \Phi_0^{j,(n-t)\alpha_j/2^r})(x) (f_{d+1} * \Phi_0^{d+1})(x) \Phi_0(t) dt dx. \end{aligned}$$

If one makes a similar decomposition for an arbitrary scale $k \neq 0$, the formula analogous to (4.23) is

$$\begin{aligned} (4.24) \quad & \int_{\mathbb{R}^2} (f * \Phi_k^{0,(n-t)/2^{r+k}})(x) (f_1 * \Psi_{r+k}^{1,n_1/2^{r+k}})(x) \\ & \quad \cdot \prod_{j=2}^d (f_j * \Phi_k^{j,(n-t)\alpha_j/2^{r+k}})(x) (f_{d+1} * \Phi_k^{d+1})(x) \Phi_0(t) dt dx. \end{aligned}$$

Summarizing the preceding, if one denotes by $\vec{\alpha} = (\alpha_2, \dots, \alpha_d)$, one sees that the piece of C_d that corresponds to Case 1'_a, can be written as

$$(4.25) \quad \int_{[0,1]^{d-1}} \int_{\mathbb{R}} \left(\sum_{r \leq 0} 2^r \sum_{n, n_1} C_{n, n_1}^r C_d^{r, n, n_1, \vec{\alpha}, t} \right) \Phi_0(t) dt d\vec{\alpha},$$

where naturally $C_d^{r, n, n_1, \vec{\alpha}, t}$ is the operator whose $(d + 2)$ -linear form is given by the sum over k of the corresponding inner expressions in (4.24).

Clearly, in order to prove (1.9) for (4.25), one would need to prove it for $C_d^{r, n, n_1, \vec{\alpha}, t}$ with upper bounds that are summable over r, n , and n_1 and integrable over t and $\vec{\alpha}$. These operators $C_d^{r, n, n_1, \vec{\alpha}, t}$ are essentially paraproducts and for them one can apply the argument described in the previous section. However, the presence of all of these parameters, has the consequence of *shifting the constituent functions* slightly, so this time one has to be very precise when evaluating the size of the boundedness constants. As before, the idea is to apply the corresponding (3.2) to all the indices $2 \leq j \leq d$ for which $p_j = \infty$. This is possible because the noncompact Littlewood–Paley L^1 normalized decompositions have been used. Denote by S the set of indices $2 \leq j \leq d$ for which $p_j \neq \infty$. Now, if $l = |S| + 2$ and one freezes as before the L^∞ normalized Schwartz functions corresponding to the indices in $\{2, \dots, d\} \setminus S$, one obtains a *minimal* l -linear operator denoted by $C_d^{l, r, n, n_1, \vec{\alpha}, t}$.

4.4. Banach estimates for $C_d^{l, r, n, n_1, \vec{\alpha}, t}$

Fix indices $1 < s_1, \dots, s_{l+1} < \infty$ so that $1/s_1 + \dots + 1/s_l = 1/s_{l+1}$. As in the previous section, the boundedness constants for

$$(4.26) \quad C_d^{l, r, n, n_1, \vec{\alpha}, t} : L^{s_1} \times \dots \times L^{s_l} \rightarrow L^{s_{l+1}}$$

depend on the boundedness constants of the square functions

$$\left(\sum_k |f * \Phi_k^{0, (n-t)/2^{r+k}}(x)|^2 \right)^{1/2} \quad \text{and} \quad \left(\sum_k |f_1 * \Psi_{r+k}^{1, n_1/2^{r+k}}(x)|^2 \right)^{1/2},$$

and of several maximal functions of type

$$\sup_k |f_j * \Phi_k^{j, (n-t)\alpha_j/2^{r+k}}(x)| \quad \text{for } j \in S.$$

It is not difficult to see that the square functions are the continuous analogues of the shifted discrete square functions $S^{[(n-t)/2^r]}$ and S^{n_1} of [9] and as a consequence, they are bounded on every L^q space for $1 < q < \infty$, with upper bounds of type $O(\log < [(n-t)/2^r] >)$ and $O(\log < n_1 >)$, respectively, see [9].⁴

Likewise, the maximal functions are bounded by the shifted maximal functions $M^{[(n-t)\alpha_j/2^r]}$ and therefore bounded on every L^q space for $1 < q < \infty$, with bounds of type $O(\log < [(n-t)\alpha_j/2^r] >)$; see again [9].

⁴For every real number γ , we denote its integer part by $[\gamma]$.

Using all these facts, one sees that the boundedness constants of (4.26) are no greater than

$$(4.27) \quad C \langle r \rangle^l (\log \langle n \rangle)^l (\log \langle n_1 \rangle)^l (\log \langle [t] \rangle)^l.$$

4.5. Quasi-Banach estimates for $C_d^{l,r,n,n_1,\tilde{\alpha},t}$

Fix indices $1 < r_1, \dots, r_l < \infty$ and $0 < r_{l+1} < \infty$ so that $1/r_1 + \dots + 1/r_l = 1/r_{l+1}$. We would like to estimate the boundedness constants of

$$(4.28) \quad C_d^{l,r,n,n_1,\tilde{\alpha},t} : L^{r_1} \times \dots \times L^{r_l} \rightarrow L^{r_{l+1}}$$

and its adjoint operators. To achieve this, we will have to discretize the corresponding (4.24) even further (with respect to the x variable), to be able to rewrite the operator $C_d^{l,r,n,n_1,\tilde{\alpha},t}$ in a form similar to (3.5), to which one can apply (3.6). First one has to observe that the bump functions corresponding to the index 1 in (4.24) are adapted to scales which are 2^{-r} times greater than the scales of the bump functions corresponding to the other indices. This fact suggests that the natural thing to do is to discretize using the bigger scale. On the other hand, one also observes that if a generic function Φ is a bump adapted to the dyadic interval J , and if $J \subseteq \tilde{J}$ is another dyadic interval 2^{-r} times longer than J , then $2^{5r}\Phi$ is a bump adapted to \tilde{J} as well (5 corresponds to the number of derivatives in the definition of *adaptedness*).

These facts, together with standard averaging and approximation arguments from [9] (including Fatou’s lemma, etc.) show that our problem can be reduced to estimating expressions of type

$$(4.29) \quad \frac{1}{2^{6rl}} \sum_I \frac{1}{|I|^{(l-1)/2}} |\langle f, \Phi_{I[n-t]}^0 \rangle| |\langle f_1, \Phi_{I[n_1]}^1 \rangle| \prod_{j \in S} |\langle f_j, \Phi_{I[(n-t)\alpha_j]}^j \rangle| |\langle f_{d+1}, \Phi_I^{d+1} \rangle|$$

where the functions f and $(f_j)_j$ are as in (3.6) and the $(p_j)_j$ there are the same as our $(r_j)_j$ here.⁵ Using (3.6) and interpolation, we deduce that the boundedness constants of (4.28) are no greater than

$$(4.30) \quad 2^{-6lr} (\log \langle n \rangle)^l (\log \langle n_1 \rangle)^l (\log \langle [t] \rangle)^l,$$

and the same is true for all the adjoints of the operator.

4.6. The final interpolation

Fix now indices $p, (p_j)_j$ as in (1.9). Given that the desired estimates are *on the edge* of the *Banach region*, one can first use convexity arguments and linear interpolation only to obtain many quasi-Banach estimates whose bounds do not grow too much with respect to r (at a rate of at most $2^{-\epsilon r}$, say, for some small ϵ). Then, one can use the multilinear interpolation theory from [13] and interpolate between these

⁵ The power 6 above should be read as $5 + 1$, where 5 comes from the *adaptedness* argument and 1 is a consequence of scaling; in particular all the bump functions are L^2 normalized now.

better quasi-Banach estimates and the previous Banach ones in (4.27), to realize that (1.9) for $C_d^{r,n,n_1,\vec{\alpha},t}$ comes with a bound which is acceptable by (4.25).

This completes the discussion of Case $1'_a$.

The analysis of the rest of the cases follow a similar strategy. Besides the *quadratic/logarithmic* argument, the presence of the decaying factor 2^r in (4.25) was also crucial. In the rest of the paper we shall describe the adjustments that one sometimes needs to make in the other cases in order for the above argument to work.

Case $1''_a$

The $1''_a$ part, corresponding to $r > 0$, is actually simpler, since this time the interaction between $\widehat{\Psi}_r(\xi_1)$ and $\widehat{\Phi}_0^1(\xi_1)$ gives

$$\widehat{\Psi}_r(\xi_1)\widehat{\Phi}_0^1(\xi_1) = \frac{1}{2^{rM}}\widehat{\widetilde{\Psi}}_r(\xi_1)$$

for some large constant $M > 0$, where $\widehat{\widetilde{\Psi}}_r(\xi_1)$ is another Ψ function adapted to the same scale as $\widehat{\Psi}_r(\xi_1)$. This huge decaying factor together with an argument similar to that made before leads to the resolution of this case.

Case 1_b

This is very similar to 1_a . In fact, the only difference is that this time the corresponding Fourier coefficients can be estimated by

$$|C_{n,n_1}| \lesssim \frac{1}{\langle n \rangle^2} \frac{1}{\langle n - n_1 \rangle^\#} + \frac{1}{\langle n \rangle^\#} \frac{1}{\langle n_1 \rangle^\#},$$

as shown in [9] and this still gives a contribution summable over n and n_1 .

Case 1_c

Here, one has first to realize that on the support of $\widehat{\Psi}_r(\widetilde{\xi})\widehat{\Phi}_r(\xi_1)$ the symbol $\int_0^1 1_{\mathbb{R}^+}(\widetilde{\xi} + \alpha_1\xi_1)d\alpha_1$ behaves like a classical Marcinkiewicz–Hörmander–Mihlin symbol and as a consequence, one has perfect decay for its Fourier coefficients, of type $1/(\langle n \rangle^\# \langle n_1 \rangle^\#)$.

There are two subcases $1'_c$ and $1''_c$, which correspond as before to $r < 0$ and $r \geq 0$, respectively.

Case $1'_c$

In this situation, one just has to observe that

$$\widehat{\Phi}_r(\xi_1)\widehat{\Phi}_0^1(\xi_1) = \widehat{\Phi}_r(\xi_1)\xi_1\widehat{\phi}_0^1(\xi_1) = \widehat{\Phi}_r(\xi_1)\xi_1 = 2^r\widehat{\Phi}_r(\xi_1)\frac{\xi_1}{2^r} = 2^r\widehat{\Psi}_r(\xi_1),$$

where $\widehat{\Psi}_r(\xi_1)$ is also of Ψ type. The presence of the factor 2^r above shows that this case can be treated exactly as the case $1'_a$.

Case 1''_c

This time one observes that when the two functions $\widehat{\Phi}_r(\xi_1)$ and $\widehat{\Phi}_0^1(\xi_1)$ interact, no decaying factor results and all one can say is that

$$\widehat{\Phi}_r(\xi_1) \widehat{\Phi}_0^1(\xi_1) = \widehat{\Phi}_0^1(\xi_1)$$

where $\widehat{\Phi}_0^1(\xi_1)$ is another Ψ function, since $\widehat{\Phi}_0^1(\xi_1)$ is adapted on an interval which lies inside the one corresponding to $\widehat{\Phi}_r(\xi_1)$ (recall that $r \geq 0$ now). To produce a decaying factor, one would have to argue somewhat differently.

Again, for simplicity, we will explain the changes that must be made in the $k = 0$, since as usual the argument is scale invariant. Consider the term

$$(4.31) \quad \widehat{\Phi}_0^0(\xi) \widehat{\Phi}_0^1(\xi_1) \cdots \widehat{\Phi}_0^{d+1}(\xi_{d+1}).$$

Recall that $\widehat{\Phi}_0^0(\xi)$ is of Ψ type, but it also has compact support, since the Littlewood–Paley decompositions have been chosen to be compact for the 0 and $d + 1$ positions. Pick another Φ function $\widehat{\Phi}_0^0(\xi)$, supported on a slightly larger interval and equal to 1 on the support of $\widehat{\Phi}_0^0(\xi)$. Then, split (4.31) as

$$(4.32) \quad \widehat{\Phi}_0^0(\xi) \widehat{\Phi}_0^1(\xi_1) \cdots \widehat{\Phi}_0^{d+1}(\xi_{d+1}) \widehat{\Phi}_0^0(\tilde{\xi})$$

$$(4.33) \quad + \widehat{\Phi}_0^0(\xi) \widehat{\Phi}_0^1(\xi_1) \cdots \widehat{\Phi}_0^{d+1}(\xi_{d+1}) [1 - \widehat{\Phi}_0^0(\tilde{\xi})].$$

The $(d + 2)$ -linear form determined by (4.32) (and its family of analogues for each scale) can be treated as in the $1''_a$ case and in fact it is even simpler, since this time one has that $\widehat{\Psi}_r(\tilde{\xi}) \widehat{\Phi}_0^0(\tilde{\xi}) = 0$ unless $r = 1, 2, 3$ (say).

To understand (4.33), we rewrite it as

$$(4.34) \quad \begin{aligned} &\widehat{\Phi}_0^0(\xi) \widehat{\Phi}_0^1(\xi_1) \cdots \widehat{\Phi}_0^{d+1}(\xi_{d+1}) [\widehat{\Phi}_0^0(\xi) - \widehat{\Phi}_0^0(\tilde{\xi})] \\ &= -\widehat{\Phi}_0^0(\xi) \widehat{\Phi}_0^1(\xi_1) \cdots \widehat{\Phi}_0^{d+1}(\xi_{d+1}) \left[\int_0^1 \widehat{\Phi}_0^0'((1-s)\xi + s\tilde{\xi}) ds \right] (\tilde{\xi} - \xi) \\ &= -\widehat{\Phi}_0^0(\xi) \widehat{\Phi}_0^1(\xi_1) \cdots \widehat{\Phi}_0^{d+1}(\xi_{d+1}) \left[\int_0^1 \widehat{\Phi}_0^0'(\xi + s(\alpha_2 \xi_2 + \cdots + \alpha_d \xi_d)) ds \right] \\ &\quad \cdot (\alpha_2 \xi_2 + \cdots + \alpha_d \xi_d). \end{aligned}$$

At this point we have to realize that we will lose another factor of d , because of the parenthesis $(\alpha_2 \xi_2 + \cdots + \alpha_d \xi_d)$. Each of these $O(d)$ expressions has the form

$$(4.35) \quad \widehat{\Phi}_0^0(\xi) \widehat{\Phi}_0^1(\xi_1) \cdots \widehat{\Phi}_0^{d+1}(\xi_{d+1}) \left[\int_0^1 \widehat{\Phi}_0^0'(\xi + s(\alpha_2 \xi_2 + \cdots + \alpha_d \xi_d)) ds \right],$$

where, for some $2 \leq j \leq d$ one has an extra factor of type $\alpha_j \xi_j$ in addition to the previous $\widehat{\Phi}_0^j(\xi_j)$. Clearly, this just adds another harmless Ψ function to that j term, so it is enough to analyze (4.35). The crucial observation here is to realize

that when one applies to (4.35) a factor of type $\widehat{\Psi}_r(\xi) \widehat{\Phi}_r(\xi_1)$, one has to have $0 \leq s \leq C/2^r$ in order for the corresponding term to be nonzero. This means that one can simply replace the integral

$$\int_0^1 \widehat{\Phi}_0^{0'}(\xi + s(\alpha_2 \xi_2 + \dots + \alpha_d \xi_d)) ds$$

in (4.35) by

$$\int_0^{C/2^r} \widehat{\Phi}_0^{0'}(\xi + s(\alpha_2 \xi_2 + \dots + \alpha_d \xi_d)) ds.$$

Now, for each fixed $r > 0$ and each $0 \leq s \leq C/2^r$, the corresponding form can be estimated exactly as before uniformly in s and in the end, after integration, one obtains an upper bound summable over $r > 0$. The extra factor $\widehat{\Phi}_0^{0'}(\xi + s(\alpha_2 \xi_2 + \dots + \alpha_d \xi_d))$ is of course harmless, since it just adds another average to the generic formula, as can be seen from (4.21).

This ends Case 1.

Case 2: $j_1 = 0$ and $j_2 = d + 1$

The goal here is to show that after some calculations, one can in fact reduce this case to the previous Case 1. To understand this, consider again a generic $k = 0$ term, as the one in (4.31). One can split it as

$$\widehat{\Phi}_0^0(\xi) \widehat{\phi}_0^1(\xi_1) \dots \widehat{\Phi}_0^{d+1}(\xi_{d+1}) + \widehat{\Phi}_0^0(\xi) \widehat{\psi}_0^1(\xi_1) \dots \widehat{\Phi}_0^{d+1}(\xi_{d+1}) = A + B,$$

where $\widehat{\phi}_0^1(\xi_1)$ is of Φ type, compactly supported at scale one, while $\widehat{\psi}_0^1(\xi_1)$ is of Ψ type also adapted at scale one. Clearly, the B terms generate $(d + 2)$ -linear forms similar to the ones in Case 1, and so it is enough to discuss the A terms only. Here it should not be difficult to see that, by construction, at least one of the two Ψ functions $\widehat{\Phi}_0^0(\xi)$ and $\widehat{\Phi}_0^{d+1}(\xi_{d+1})$, has its support away from zero. Moreover, we claim that without loss of generality, one can assume that $\widehat{\Phi}_0^0(\xi)$ is that function. To see this, one just has to observe that the roles of the variables ξ and ξ_{d+1} are totally symmetric. Indeed, since $\xi + \xi_1 + \dots + \xi_{d+1} = 0$ a simple change of variables shows that

$$\begin{aligned} \int_{[0,1]^d} 1_{\mathbb{R}_+}(\xi + \alpha_1 \xi_1 + \dots + \alpha_d \xi_d) d\alpha_1 \dots d\alpha_d \\ = \int_{[0,1]^d} 1_{\mathbb{R}_-}(\xi_{d+1} + \beta_1 \xi_1 + \dots + \beta_d \xi_d) d\beta_1 \dots d\beta_d, \end{aligned}$$

which is obviously a similar symbol.

In particular, one can clearly rewrite A as

$$(4.36) \quad \widehat{\Phi}_0^0(\xi) \widehat{\phi}_0^1(\xi_1) \dots \widehat{\Phi}_0^{d+1}(\xi_{d+1}) \widehat{\Psi}_0^0(\xi)$$

for another well chosen compactly supported Ψ function $\widehat{\Psi}_0^0$.

Then, one rewrites (4.36) further as

$$(4.37) \quad \widehat{\Phi}_0^0(\xi) \widehat{\phi}_0^1(\xi_1) \cdots \widehat{\Phi}_0^{d+1}(\xi_{d+1}) \widehat{\Psi}_0^0(\tilde{\xi}) \\ + \widehat{\Phi}_0^0(\xi) \widehat{\phi}_0^1(\xi_1) \cdots \widehat{\Phi}_0^{d+1}(\xi_{d+1}) [\widehat{\Psi}_0^0(\xi) - \widehat{\Psi}_0^0(\tilde{\xi})].$$

Then, one can observe that the symbol

$$\int_0^1 1_{\mathbb{R}^+}(\tilde{\xi} + \alpha_1 \xi_1) d\alpha_1$$

is a classical symbol on the support of the first term in (4.37) and its analysis becomes simpler. In particular, one no longer needs to use the extra decomposition over r to study it.

We are thus left with the second term of (4.37). Modulo a minus sign, this term can be written as

$$\widehat{\Phi}_0^0(\xi) \widehat{\phi}_0^1(\xi_1) \cdots \widehat{\Phi}_0^{d+1}(\xi_{d+1}) \left[\int_0^1 \widehat{\Psi}_0^{0'}((1-s)\xi + s\tilde{\xi}) ds \right] \cdot (\alpha_2 \xi_2 + \cdots + \alpha_d \xi_d) \\ = \widehat{\Phi}_0^0(\xi) \widehat{\phi}_0^1(\xi_1) \cdots \widehat{\Phi}_0^{d+1}(\xi_{d+1}) \left[\int_0^1 \widehat{\Psi}_0^{0'}(\xi + s(\alpha_2 \xi_2 + \cdots + \alpha_d \xi_d)) ds \right] \\ \cdot (\alpha_2 \xi_2 + \cdots + \alpha_d \xi_d).$$

We decompose this last term further as

$$(4.38) \quad \widehat{\Phi}_0^0(\xi) \widehat{\phi}_0^1(\xi_1) \cdots \widehat{\Phi}_0^{d+1}(\xi_{d+1}) \left[\int_0^1 \widehat{\Psi}_0^{0'}(\xi + s(-\xi)) ds \right] \cdot (\alpha_2 \xi_2 + \cdots + \alpha_d \xi_d)$$

$$(4.39) \quad + \widehat{\Phi}_0^0(\xi) \widehat{\phi}_0^1(\xi_1) \cdots \widehat{\Phi}_0^{d+1}(\xi_{d+1}) \left[\int_0^1 \int_0^1 \widehat{\Psi}_0^{0''}((1-ts)\xi \right. \\ \left. + (1-t)s(\alpha_2 \xi_2 + \cdots + \alpha_d \xi_d)) s ds dt \right] \cdot (\alpha_2 \xi_2 + \cdots + \alpha_d \xi_d) \cdot \tilde{\xi}.$$

We claim now that the term in (4.38) can be reduced to Case 1 studied earlier. Indeed, notice first that the expression $(\alpha_2 \xi_2 + \cdots + \alpha_d \xi_d)$ contributes $d - 1$ terms and consider for instance the one corresponding to $\alpha_2 \xi_2$. The presence of ξ_2 transforms the bump function depending on this variable into one of Ψ type in an intermediate position (as in Case 1) and then, one has only to observe that the *localized Fourier coefficients* of symbols of type

$$(\tilde{\xi}, \xi_2) \rightarrow \int_0^1 \alpha_2 1_{\mathbb{R}^+}(\tilde{\xi} + \alpha_2 \xi_2) d\alpha_2$$

still satisfy the same crucial quadratic estimates that have been proved earlier in [9].

We are then left with the study of the terms in (4.39). From now on (as many times before) we think of the variables $\alpha_2, \dots, \alpha_d$ as being *frozen* and of our symbol as being of type

$$(4.40) \quad \int_0^1 1_{\mathbb{R}^+}(\tilde{\xi} + \alpha_1 \xi_1) d\alpha_1.$$

The variables ξ_1 and $\tilde{\xi}$ are of course *special*, but so is $(\alpha_2 \xi_2 + \dots + \alpha_d \xi_d)$ as appears quite explicitly in (4.39). Consider now an extra *paraproduct decomposition* of the identity, in the form of finitely many expressions of the type

$$(4.41) \quad \sum_r \widehat{\Phi}_r(\xi_1) \widehat{\Phi}_r(\tilde{\xi}) \widehat{\Phi}_r(\alpha_2 \xi_2 + \dots + \alpha_d \xi_d).$$

This can easily be obtained by combining three independent Littlewood–Paley decompositions. It is important to emphasize that at least one of the above ingredient functions must be of Ψ type. As in Case 1, the idea now is to insert this extra decomposition (4.41) into (4.39) and study the newly formed expressions. The support of the function of two variables

$$(\xi_1, \tilde{\xi}) \rightarrow \widehat{\Phi}_r(\xi_1) \widehat{\Phi}_r(\tilde{\xi})$$

will clearly play an important role, since as long as this support is a Whitney square one can decompose the symbol (4.40) on it as a double Fourier series, precisely as in Case 1. We therefore must consider two distinct situations.

The Whitney case

In this case, the above supports are all Whitney squares with respect to the origin. This means that either $\widehat{\Phi}_r(\xi_1)$ or $\widehat{\Phi}_r(\tilde{\xi})$ is of Ψ type. It is also useful to observe that since $\xi = \tilde{\xi} - (\alpha_2 \xi_2 + \dots + \alpha_d \xi_d)$, it must belong to an interval of size 2^r centered at the origin. This means that one must have $r \geq 0$ in order for (4.41) to have a nontrivial interaction with (4.39) (remember that from the beginning, we are in the case where $\widehat{\Phi}_0^0(\xi)$ has Ψ type and is also compactly supported away from the origin).

The case when $\widehat{\Phi}_r(\xi_1)$ is of Ψ type is easier since when (4.41) interacts with (4.39) the only nonzero terms are those corresponding to indices r belonging to the finite set $\{0, 1, 2\}$. After that, one just applies the method of Case 1. Notice that because of the terms $(\alpha_2 \xi_2 + \dots + \alpha_d \xi_d) \cdot \tilde{\xi}$ in (4.39), one will lose another factor of type $O(d^2)$ (after redistributing the inner terms) which is clearly acceptable.

The other case, when $\widehat{\Phi}_r(\tilde{\xi})$ is of Ψ type is more complicated, since all the scales $r \geq 0$ can contribute. However, in this case one observes that when (4.41) interacts with (4.39) then one must have either s or $1 - t$ smaller than $C/2^{r/2}$ in (4.39) (for a certain fixed but large constant $C > 0$). This then shows that this case can be treated exactly as the previously considered Case 1''.

The non-Whitney case

This case corresponds to the situation when both $\widehat{\Phi}_r(\tilde{\xi})$ and $\widehat{\Phi}_r(\xi_1)$ are of Φ type. However, as a consequence of (4.41), the support of $\widehat{\Phi}_r(\alpha_2 \xi_2 + \dots + \alpha_d \xi_d)$ must be an interval of the same size 2^r whose distance to the origin is comparable to its length. In particular, $\xi = \tilde{\xi} - (\alpha_2 \xi_2 + \dots + \alpha_d \xi_d)$ must belong to an interval of a similar kind. Then, because of the presence of $\widehat{\Phi}_0^0(\xi)$, one must have $r \sim 0$ in

order for (4.41) to have a nontrivial interaction with (4.39). Hence the new term that gets multiplied with (4.39) in this case, is of the form

$$(4.42) \quad \widehat{\Phi}_0(\xi_1) \widehat{\Phi}_0(\widetilde{\xi}) \widehat{\Phi}_0(\alpha_2 \xi_2 + \cdots + \alpha_d \xi_d).$$

At this point it is important to remember the factor $\widetilde{\xi}$ in (4.39). When it is multiplied with the above $\widehat{\Phi}_0(\widetilde{\xi})$ it transforms this function into one of Ψ type, which is clearly very good news. We denote it by $\widehat{\Psi}_0(\widetilde{\xi})$ for the rest of the discussion. We are still not done yet, since this new Ψ type function does not have support away from the origin. However, we can apply to this situation a treatment similar to the one used in the previous Case 1. Before doing so, to summarize, the expression that we now face consists of a product of a term of the type

$$(4.43) \quad \widehat{\Phi}_0(\xi_1) \widehat{\Psi}_0(\widetilde{\xi}) \widehat{\Phi}_0(\alpha_2 \xi_2 + \cdots + \alpha_d \xi_d)$$

with the previous (4.39) which no longer contains the original factor $\widetilde{\xi}$. At this point, insert another decomposition of the identity, of the type

$$\sum_r \widehat{\Phi}_r(\xi_1) \widehat{\Phi}_r(\widetilde{\xi}),$$

where as before either $\widehat{\Phi}_r(\xi_1)$ or $\widehat{\Phi}_r(\widetilde{\xi})$ is of Ψ type. And finally, exactly as in Case 1, observe that when this new decomposition is multiplied with the above (4.43), the index r must be smaller than zero to produce nontrivial terms, and also a small factor of type 2^r arises naturally, from the interaction between $\widehat{\Phi}_r(\widetilde{\xi})$ and $\widehat{\Psi}_0(\widetilde{\xi})$. After that, the argument is identical to the one used for Case 1.

Case 3: $j_1 = 2$ and $j_2 = 3$

Finally, it is not difficult to see that Case 3 can be analyzed in the same way as was Case 1, since there are now two Ψ type functions in intermediate positions.

This ends our proof since, by symmetry, any other case can be reduced to one of these three.

5. Generalizations

To be able to describe and motivate the generalizations we mentioned at the beginning of the paper, we would first like to recall the classical calculations of Calderón, which gave rise to his commutators.

5.1. Calculus with functions of linear growth

Let $A(x)$ be a complex valued function of one real variable having *linear growth*, more precisely satisfying $A' \in L^\infty$. We denote by $Hf(x)$ the classical Hilbert transform given by

$$(5.1) \quad Hf(x) := p.v. \int_{\mathbb{R}} f(x - y) \frac{dy}{y},$$

and by $Af(x)$ the operator of multiplication by $A(x)$. The question that started the whole theory, was whether the commutator $[H, A]$ was smoothing of order one, or equivalently if $[H, A] \circ D$ maps L^p into L^p boundedly for $1 < p < \infty$, where $Df(x) := f'(x)$.⁶

For f smooth and compactly supported, one can write

$$\begin{aligned} (HA - AH) \circ D(f)(x) &= (HA - AH)(f')(x) \\ &= p.v. \int_{\mathbb{R}} A(x - y)f'(x - y) \frac{dy}{y} - p.v. \int_{\mathbb{R}} A(x)f'(x - y) \frac{dy}{y} \\ &= p.v. \int_{\mathbb{R}} A(y)f'(y) \frac{1}{x - y} dy - p.v. \int_{\mathbb{R}} A(x)f'(y) \frac{1}{x - y} dy \\ &= -p.v. \int_{\mathbb{R}} \frac{A(x) - A(y)}{x - y} f'(y) dy = p.v. \int_{\mathbb{R}} \left(\frac{A(x) - A(y)}{x - y} \right)' f(y) dy \\ &= -p.v. \int_{\mathbb{R}} \frac{A'(y)}{x - y} f(y) dy + p.v. \int_{\mathbb{R}} \frac{A(x) - A(y)}{(x - y)^2} f(y) dy \\ &= -H(A'f)(x) + C_1 f(x), \end{aligned}$$

where C_1 is precisely the first Calderón commutator. Since both $f \rightarrow H(A'f)$ and C_1 are bounded on L^p for $1 < p < \infty$, these show that indeed $[H, A]$ is smoothing of order one.

Besides the commutators, Calderón pointed out that more general operators such as

$$(5.2) \quad f \rightarrow p.v. \int_{\mathbb{R}} F\left(\frac{A(x) - A(y)}{x - y}\right) f(y) \frac{dy}{x - y},$$

or even

$$(5.3) \quad f \rightarrow p.v. \int_{\mathbb{R}} F\left(\frac{A(x) - A(y)}{x - y}\right) G\left(\frac{B(x) - B(y)}{x - y}\right) f(y) \frac{dy}{x - y}$$

merit study as well, since they appear naturally in complex analysis or boundary value problems in PDE.⁷

The work of Coifman, McIntosh, and Meyer [5], proved the desired estimates for all these operators, by reducing them to the previous (1.3). Finally, we also remark that the above calculations show that, if one writes $a := A'$, then for $g, Dg \in L^p$ with $1 < p < \infty$, one has

$$(5.4) \quad H(ADg) = AH(Dg) - H(aDg) + C_1 g.$$

This is remarkable since a priori there is no direct and obvious way to even define $H(ADg)$. Notice that on the right-hand side, all the compositions make sense.

⁶As a consequence of this, HA could be written as $HA = AH + [H, A]$ and therefore belonged to Calderón's algebra.

⁷Here the functions F and G are for instance analytic in a certain disc around the origin, while $\|A'\|_\infty$ and $\|B'\|_\infty$ are supposed to be strictly smaller than the radii of convergence of F and G respectively. Of course, one can consider similar operators with more than two factors.

5.2. Calculus with functions of polynomial growth, part 1

Suppose now that $A(x)$ is a function having *quadratic growth*, more precisely satisfying $A'' \in L^\infty$. A natural question is the following. *Is it true that the commutator $[H, A]$ is still a smoothing operator?* We will see that this time $[H, A]$ is smoothing of order two.

Indeed, for f a smooth and compactly supported function, given also the previous calculations, one has

$$\begin{aligned} (HA - AH) \circ D^2(f)(x) &= (HA - AH) \circ D(f')(x) \\ &= p.v. \int_{\mathbb{R}} \left(\frac{A(x) - A(y)}{(x - y)^2} - \frac{A'(y)}{x - y} \right) f'(y) dy \\ &= p.v. \int_{\mathbb{R}} \frac{A(x) - A(y) - A'(y)(x - y)}{(x - y)^2} f'(y) dy \\ &= -p.v. \int_{\mathbb{R}} \frac{-A'(y) - A''(y)(x - y) + A'(y)}{(x - y)^2} f(y) dy \\ &\quad + 2 p.v. \int_{\mathbb{R}} \frac{A(x) - A(y) - A'(y)(x - y)}{(x - y)^3} f(y) dy \\ &= H(A''f)(x) + 2 p.v. \int_{\mathbb{R}} \left(\frac{A(x) - T_y^1 A(x)}{(x - y)^2} \right) \frac{f(y)}{x - y} dy, \end{aligned}$$

where

$$T_y^1 A(x) := A(y) + \frac{A'(y)}{1!}(x - y)$$

is the Taylor polynomial of order 1 of the function A about the point y . Since $f \rightarrow H(A''f)$ is clearly a bounded operator, the problem reduces to proving L^p bounds for the linear operator

$$(5.5) \quad f \rightarrow p.v. \int_{\mathbb{R}} \left(\frac{A(x) - T_y^1 A(x)}{(x - y)^2} \right) \frac{f(y)}{x - y} dy.$$

For functions of arbitrary *polynomial growth* satisfying $A^{(d)} \in L^\infty$ for some $d \geq 1$, one can similarly show that $[H, A]$ is smoothing of order d , if the operator

$$(5.6) \quad f \rightarrow p.v. \int_{\mathbb{R}} \left(\frac{A(x) - T_y^{d-1} A(x)}{(x - y)^d} \right) \frac{f(y)}{x - y} dy$$

satisfies the usual estimates, where $T_y^{d-1} A(x)$ is the Taylor polynomial of order $d-1$ of the function A about the point y .

More generally, as before, one can ask if the operators

$$(5.7) \quad f \rightarrow p.v. \int_{\mathbb{R}} F \left(\frac{A(x) - T_y^{d-1} A(x)}{(x - y)^d} \right) \frac{f(y)}{x - y} dy,$$

or even

$$(5.8) \quad f \rightarrow p.v. \int_{\mathbb{R}} G \left(\frac{B(x) - T_y^{d_1-1} B(x)}{(x - y)^{d_1}} \right) H \left(\frac{C(x) - T_y^{d_2-1} C(x)}{(x - y)^{d_2}} \right) \frac{f(y)}{x - y} dy$$

are bounded on L^p for $1 < p < \infty$, assuming as before, that F, G , and H are analytic while $A^{(d)}$, $B^{(d_1)}$, and $C^{(d_2)}$ are all in L^∞ .⁸ As a consequence of the method in this paper, one can answer all these questions affirmatively. Indeed, it is not difficult to see, thanks to the general averaging formula for the remainder term of a Taylor series, that as before, the problem reduces to proving $L^p \times L^\infty \times \dots \times L^\infty \rightarrow L^p$ estimates for the $(k + 1)$ -linear operator with symbol

$$\frac{1}{(d!)^k} \int_{[0,1]^k} \text{sgn}(\xi + \alpha_1 \xi_1 + \dots + \alpha_k \xi_k) (1 - \alpha_1)^d \dots (1 - \alpha_k)^d d\alpha_1 \dots d\alpha_k,$$

which grow at most polynomially in k . However this symbol is very similar to the symbol of the k th Calderón commutator and because of this, one can prove the desired estimates similarly. In fact, for $k = 1$ one can actually see that

$$\int_0^1 \text{sgn}(\xi + \alpha_1 \xi_1) (1 - \alpha_1)^d d\alpha_1$$

is even better than the symbol of the first commutator, since at least along the line $\xi + \xi_1 = 0$ it becomes smoother, due to the presence of the extra factor $(1 - \alpha_1)^d$. In particular, the *quadratic estimates* for its Fourier coefficients are still available.

It is quite likely that the $T1$ theorem of David and Journé, [6], can be used to handle the cases when F, G , and H are of the form x^n for some positive integer n . In fact, for $F(x) = x$ this has been verified in [4].⁹ The more general case $G(x) = H(x) = x$ has also been treated directly in [2].

5.3. Calculus with functions of polynomial growth, part 2

There is an alternative calculation that one can perform to understand the previous question. Let us first observe that $[H, A]$ is smoothing of order two if and only if $[H, A] \circ H$ is smoothing of order two. Then, using the fact that $H^2 = -I$, one can write

$$\begin{aligned} (AH - HA) \circ H(f) &= AH^2 f - H(AHf) \\ &= -Af - H(AHf) = \frac{1}{2} (H^2(Af) + AH^2 f - 2H(AHf)). \end{aligned}$$

If one ignores the coefficient $1/2$, the expression in parentheses (calculated at an arbitrary point x) becomes

$$\begin{aligned} &p.v. \int_{\mathbb{R}^2} f(x + t + s)A(x + t + s) \frac{dt}{t} \frac{ds}{s} + p.v. \int_{\mathbb{R}^2} f(x + t + s)A(x) \frac{dt}{t} \frac{ds}{s} \\ &- p.v. \int_{\mathbb{R}^2} f(x + t + s)A(x + t) \frac{dt}{t} \frac{ds}{s} - p.v. \int_{\mathbb{R}^2} f(x + t + s)A(x + s) \frac{dt}{t} \frac{ds}{s} \\ &= p.v. \int_{\mathbb{R}^2} f(x + t + s) (A(x + t + s) - A(x + t) - A(x + s) + A(x)) \frac{dt}{t} \frac{ds}{s}. \end{aligned}$$

⁸In fact, even more generally, one can consider operators having an arbitrary number of similar factors.

⁹It is also interesting to see on page 94 of [4], another instance where these operators appear naturally.

To see if the last expression is smoothing of order two, one has (after integrating by parts with respect to the t variable)

$$\begin{aligned}
 & p.v. \int_{\mathbb{R}^2} f''(x+t+s) \left(\frac{A(x+t+s) - A(x+t) - A(x+s) + A(x)}{ts} \right) dt ds \\
 &= -p.v. \int_{\mathbb{R}^2} f'(x+t+s) \left(\frac{A'(x+t+s) - A'(x+t)}{ts} \right) dt ds \\
 & \quad + p.v. \int_{\mathbb{R}^2} f'(x+t+s) \left(\frac{A(x+t+s) - A(x+t) - A(x+s) + A(x)}{t^2s} \right) dt ds \\
 & := I + II.
 \end{aligned}$$

Integrating by parts with respect to the s variable now, one can rewrite I as

$$\begin{aligned}
 & p.v. \int_{\mathbb{R}^2} f(x+t+s) A''(x+t+s) \frac{dt ds}{t s} \\
 & \quad - p.v. \int_{\mathbb{R}^2} f(x+t+s) \frac{A'(x+t+s) - A'(x+s)}{s^2} ds \frac{dt}{t} \\
 &= -A''(x) f(x) - H(C_{1,A'} f)(x),
 \end{aligned}$$

where $C_{1,A'}$ is the Calderón first commutator associated to the Lipschitz function A' .

Similarly, one can rewrite II as

$$\begin{aligned}
 & -H(C_{1,A'} f)(x) + p.v. \int_{\mathbb{R}^2} f(x+t+s) \frac{A(x+t+s) - A(x+t) - A(x+s) + A(x)}{t^2s^2} dt ds \\
 (5.9) \quad &= -H(C_{1,A'} f)(x) + p.v. \int_{\mathbb{R}^2} f(x+t+s) \left(\frac{\Delta_t}{t} \circ \frac{\Delta_s}{s} A(x) \right) \frac{dt ds}{t s},
 \end{aligned}$$

where in general, $\Delta_h g(x)$ denotes the usual finite difference at scale h given by

$$\Delta_h g(x) = g(x+h) - g(x).$$

The problem can therefore be reduced to that of proving L^p estimates for the linear operator in (5.9). The reader might remember it from our previous paper [9]. If A has arbitrary polynomial growth (i.e., $A^{(d)} \in L^\infty$ for some $d \geq 1$), then an analogous calculation shows that $[H, A]$ is smoothing of order d , if the operator

$$(5.10) \quad f \rightarrow p.v. \int_{\mathbb{R}^d} f(x+t_1+\dots+t_d) \left(\frac{\Delta_{t_1}}{t_1} \circ \dots \circ \frac{\Delta_{t_d}}{t_d} A(x) \right) \frac{dt_1}{t_1} \dots \frac{dt_d}{t_d}$$

is L^p bounded.

Then, more generally, one can ask the same question about the operator given by the expression

$$(5.11) \quad p.v. \int_{\mathbb{R}^d} f(x+t_1+\dots+t_d) F \left(\frac{\Delta_{t_1}}{t_1} \circ \dots \circ \frac{\Delta_{t_d}}{t_d} A(x) \right) \frac{dt_1}{t_1} \dots \frac{dt_d}{t_d},$$

or by (in the case of two factors)

$$(5.12) \quad p.v. \int_{\mathbb{R}^d} f(x + t_1 + \dots + t_d) G\left(\frac{\Delta_{a_1 t_1}}{t_1} \circ \dots \circ \frac{\Delta_{a_d t_d}}{t_d} B(x)\right) \cdot H\left(\frac{\Delta_{b_1 t_1}}{t_1} \circ \dots \circ \frac{\Delta_{b_d t_d}}{t_d} C(x)\right) \frac{dt_1}{t_1} \dots \frac{dt_d}{t_d},$$

where $(a_j)_j$ and $(b_j)_j$ are all real numbers different from zero and F, G , and H are analytic functions as before. It is also interesting to compare these formulas with the classical expression (5.2), which can be rewritten as

$$p.v. \int_{\mathbb{R}} f(x + t) F\left(\frac{\Delta_t}{t} A(x)\right) \frac{dt}{t}.$$

We now claim that essentially without any extra effort, one can prove the desired estimates for all these operators. Indeed, using that

$$\frac{\Delta_{t_1}}{t_1} \circ \dots \circ \frac{\Delta_{t_d}}{t_d} A(x) = \int_0^1 \dots \int_0^1 A^{(d)}(x + \alpha_1 t_1 + \dots + \alpha_d t_d) d\alpha_1 \dots d\alpha_d,$$

it is not difficult to see that (5.11) can be reduced to the study of the $(k + 1)$ -linear operator with symbol¹⁰

$$(5.13) \quad \left(\int_{[0,1]^k} \text{sgn}(\xi + \alpha_1 \xi_1 + \dots + \alpha_k \xi_k) d\alpha_1 \dots d\alpha_k \right)^d,$$

while the more general (5.12) to $(k + 1)$ -linear operators whose symbols are products of type

$$(5.14) \quad \prod_{i=1}^d \left(\int_{[0,1]^k} \text{sgn}(\xi + c_1^i \alpha_1 \xi_1 + \dots + c_k^i \alpha_k \xi_k) d\alpha_1 \dots d\alpha_k \right)$$

for various nonzero real numbers $(c_j^i)_{i,j}$. The only thing that needs to be realized at this point is that the method extends naturally to cover *product symbols* of type (5.13) and (5.14) as well, since each individual factor can be decomposed as before, as a Fourier series with Fourier coefficients that decay at least quadratically. More specifically, the only difference in the argument is that instead of the Littlewood–Paley decomposition in (4.12), which works very well in the $d = 1$ case, one has to consider a product of d such similar decompositions, each naturally corresponding to the factors of (5.14).¹¹

¹⁰Notice that this is precisely the symbol of the k th Calderón commutator raised to the power d !

¹¹For instance, for $d = 2$, one takes a product of two 2-dimensional Littlewood–Paley decompositions of type (4.12), one for the pair of variables (ξ_1, ξ) corresponding to the first factor and another for the pair of variables $(\xi_1, \tilde{\xi})$ corresponding to the second factor.

5.4. Extended Calderón algebra

Given the previous discussion, it is also natural to consider operators of type

$$(5.15) \quad f \rightarrow p.v. \int_{\mathbb{R}} f(x+t) F\left(\frac{\Delta_t}{t} \circ \frac{\Delta_t}{t} A(x)\right) \frac{dt}{t},$$

or even

$$(5.16) \quad f \rightarrow p.v. \int_{\mathbb{R}} f(x+t) G\left(\frac{\Delta_t}{t} B(x)\right) H\left(\frac{\Delta_t}{t} \circ \frac{\Delta_t}{t} C(x)\right) \frac{dt}{t},$$

and so on, and ask if they are L^p bounded if $F, G,$ and H are analytic and $A'', B', C'' \in L^\infty$. They can also be treated by the method of this paper. For instance, the study of (5.15) can be reduced to the study of multilinear operators with symbols of type

$$\int_{[0,1]^{2k}} \text{sgn}(\xi + (\alpha_1 + \beta_1) \xi_1 + \dots + (\alpha_k + \beta_k) \xi_k) \bar{\alpha} d\beta,$$

which clearly can be analyzed in a similar manner. One might ask: *Are they related to Calderón's algebra in any way?*¹² To be able to answer this, we need to recall the bilinear Hilbert transform ([7], [8]).

If $\alpha \in \mathbb{R} \setminus \{0, 1\}$ the bilinear Hilbert transform with parameter α denoted by BHT_α , is the bilinear operator defined by

$$(5.17) \quad BHT_\alpha(f, g)(x) := p.v. \int_{\mathbb{R}} f(x+t) g(x+\alpha t) \frac{dt}{t}.$$

It is known that these operators satisfy many L^p estimates of Hölder type ([7], [8]). As a consequence of the results in [7] and [8] and of the L^p theorem for (5.15) in the particular case $F(x) = x$, one obtains by a straightforward calculation that

$$BHT_2(f, A)(x) - 2HAf(x) + AHf(x)$$

is a smoothing term of order two, if $A'' \in L^\infty$. In particular, given also the earlier calculations for functions of quadratic growth, one has that

$$BHT_2(f, A) = AHf + \text{smoothing term of order 2},$$

which shows that the operator $f \rightarrow BHT_2(f, A)$ belongs to Calderón's algebra.¹³ Also, the parameter 2 above can be replaced by any other α , as a consequence of the fact that the theorem for (5.15) holds true if one replaces $\Delta_t/t \circ \Delta_t/t$ by $\Delta_{at}/t \circ \Delta_{bt}/t$ for some appropriate constants a and b .

¹²Generally speaking, we say that an operator belongs to Calderón's algebra, if it can be written as a sum of a classical Fourier integral operator with a well defined symbol $a(x, \xi)$ (such as AH for example) and a smoothing operator of a certain order.

¹³An even simpler calculation shows that if $A' \in L^\infty$, then $BHT_2(f, A) = AHf$ plus first order smoothing term.

Finally, let us also remark that one can *add derivatives* freely to the general operators in (5.11) and (5.12) as well and still obtain operators bounded on L^p . A typical example would be

$$(5.18) \quad f \rightarrow p.v. \int_{\mathbb{R}^2} f(x+t+s) F\left(\frac{\Delta_t}{t} \circ \frac{\Delta_t}{t} \circ \frac{\Delta_s}{s} A(x)\right) \cdot G\left(\frac{\Delta_t}{t} \circ \frac{\Delta_s}{s} \circ \frac{\Delta_s}{s} \circ \frac{\Delta_s}{s} B(x)\right) \frac{dt ds}{t s}$$

for $A''', B'''' \in L^\infty$. The straightforward (by now) details are left to the reader.¹⁴

5.5. Circular commutators

In [9] we noticed a certain *symmetric* bilinear operator which we called a *circular commutator*.

We would like to describe its natural trilinear generalization here.¹⁵ Thus, we record the following theorem.

Theorem 5.1. *Let $a \neq 0, b \neq 0,$ and $c \neq 0$ be three fixed real numbers. Consider also three Lipschitz functions $A, B,$ and C . Then, the expression*

$$p.v. \int_{\mathbb{R}^3} \left(\frac{\Delta_{at_1}}{t_1} A(x+t_2)\right) \left(\frac{\Delta_{bt_2}}{t_2} B(x+t_3)\right) \left(\frac{\Delta_{ct_3}}{t_3} C(x+t_1)\right) \frac{dt_1}{t_1} \frac{dt_2}{t_2} \frac{dt_3}{t_3}$$

viewed as a trilinear operator in $A', B',$ and C' maps $L^{p_1} \times L^{p_2} \times L^{p_3}$ into L^p boundedly, for every $1 < p_1, p_2, p_3 \leq \infty$ with $1/p_1 + 1/p_2 + 1/p_3 = 1/p$ and $1/2 < p < \infty$.

The proof of this theorem uses again the same method. If we assume for simplicity that $a = b = c = 1,$ then the symbol of the corresponding trilinear operator is given by the following *circular product*

$$m_2(\xi_1, \xi_2, \xi_3) \cdot m_2(\xi_2, \xi_3, \xi_1) \cdot m_2(\xi_3, \xi_1, \xi_2)$$

and such symbols can clearly be treated in the same manner as before.

5.6. $T1$ calculations

It is well known that the $T1$ theorem [6] can handle Calderón commutators quite successfully. More precisely, it allows one to reduce the study of C_k to the study of C_{k-1} . Let us briefly recall the details, in the particular case of the first two commutators.

Assume that A is smooth and compactly supported, and that $A' \in L^\infty$. Rewrite $C_{1,A}$ as

$$C_{1,A}f(x) = p.v. \int_{\mathbb{R}} f(x+t) \frac{A(x+t) - A(x)}{t^2} dt.$$

¹⁴The same is true if one considers their natural generalizations, when one may face more than two kernels and more than two factors in the corresponding expressions.

¹⁵There are of course many other multilinear generalizations of this, as the reader can imagine.

Then, integrating by parts, one gets

$$\begin{aligned}
 C_{1,A}1(x) &= p.v. \int_{\mathbb{R}} \frac{A(x+t) - A(x)}{t^2} dt = -p.v. \int_{\mathbb{R}} (A(x+t) - A(x)) \left(\frac{1}{t}\right)' dt \\
 &= p.v. \int_{\mathbb{R}} A'(x+t) \frac{dt}{t} = H(A')(x),
 \end{aligned}$$

which is a BMO function.

Suppose now that A and B are smooth, compactly supported and $A', B' \in L^\infty$. As before, rewrite the *second commutator* associated to A and B as

$$C_{2,A,B}f(x) = p.v. \int_{\mathbb{R}} f(x+t) \frac{(A(x+t) - A(x))(B(x+t) - B(x))}{t^3} dt.$$

Then, one has

$$\begin{aligned}
 C_{2,A,B}1(x) &= p.v. \int_{\mathbb{R}} \frac{(A(x+t) - A(x))(B(x+t) - B(x))}{t^3} dt \\
 &= -\frac{1}{2} p.v. \int_{\mathbb{R}} (A(x+t) - A(x))(B(x+t) - B(x)) \left(\frac{1}{t^2}\right)' dt \\
 &= \frac{1}{2} p.v. \int_{\mathbb{R}} A'(x+t) \frac{B(x+t) - B(x)}{t^2} dt + \frac{1}{2} p.v. \int_{\mathbb{R}} B'(x+t) \frac{A(x+t) - A(x)}{t^2} dt \\
 &= \frac{1}{2} C_{1,B}(A')(x) + \frac{1}{2} C_{1,A}(B')(x),
 \end{aligned}$$

in which both terms of the last sum are BMO functions. Similar calculations can be made for to more general *commutators* of type

$$(5.19) \quad f \rightarrow \int_{\mathbb{R}} f(x+t) \left(\frac{\Delta_{a_1 t}}{t} A_1(x)\right) \cdots \left(\frac{\Delta_{a_d t}}{t} A_d(x)\right) \frac{dt}{t}$$

for every sequence $(a_j)_j$ of nonzero real numbers.

Even though the operators of the previous sections do not come in a *standard form* (in particular, their kernels have a product structure) it is still tempting to check if there is an analogous *T1 type reduction of complexity* available for them.

Consider smooth and compactly supported functions A and B satisfying $A'', B'' \in L^\infty$. The simplest analogue of the first commutator, is the operator $T_{1,A}$ given by

$$T_{1,A}f(x) = p.v. \int_{\mathbb{R}^2} f(x+t+s) \left(\frac{\Delta_{at}}{t} \circ \frac{\Delta_{bt}}{t} A(x)\right) \frac{dt}{t} \frac{ds}{s}$$

for nonzero real numbers a and b . One can write

$$\begin{aligned}
 T_{1,A}1(x) &= p.v. \int_{\mathbb{R}^2} \frac{A(x+at+bs) - A(x+at) - A(x+bs) + A(x)}{t^2 s^2} dt ds \\
 &= p.v. \int_{\mathbb{R}^2} (A(x+at+bs) - A(x+at) - A(x+bs) + A(x)) \left(\frac{1}{t}\right)' \left(\frac{1}{s}\right)' dt ds \\
 &= p.v. \int_{\mathbb{R}^2} A''(x+at+bs) \frac{1}{t} \frac{1}{s} dt ds = H \circ H(A'')(x) = -A''(x),
 \end{aligned}$$

which is in L^∞ and therefore still in BMO.

However, as we will see, the problem becomes more complicated and the symmetry gets broken at the next step, when one considers the analogue of the second commutator $T_{2,A,B}$ given by

$$(5.20) \quad T_{2,A,B}f(x) = p.v. \int_{\mathbb{R}^2} f(x+t+s) \left(\frac{\Delta_{at}}{t} \circ \frac{\Delta_{bs}}{s} A(x) \right) \left(\frac{\Delta_{ct}}{t} \circ \frac{\Delta_{ds}}{s} B(x) \right) \frac{dt}{t} \frac{ds}{s}$$

for nonzero real numbers a, b, c , and d . This time, one has

$$T_{2,A,B}1(x) = \frac{1}{4} p.v. \int_{\mathbb{R}^2} (\Delta_{at} \circ \Delta_{bs} A(x)) (\Delta_{ct} \circ \Delta_{ds} B(x)) \left(\frac{1}{t} \right)'' \left(\frac{1}{s} \right)'' dt ds.$$

One particular term that appears after integrating by parts is the one which corresponds to the situation when a pair of s and t derivatives arises in the first term and another similar pair arises in the second term. The operator obtained in this way is

$$p.v. \int_{\mathbb{R}^2} A''(x+at+bs) B''(x+ct+ds) \frac{dt}{t} \frac{ds}{s}$$

is clearly of $BHT_{\alpha,\beta}$ type and it is unlikely that such operators map $L^\infty \times L^\infty$ into BMO.

Even worse, if instead of (5.20) one considers its natural generalization with four factors T_{4,A_1,A_2,A_3,A_4} , a similar calculation generates the expression

$$p.v. \int_{\mathbb{R}^2} A_1''(x+a_1t+b_1s) A_2''(x+a_2t+b_2s) A_3''(x+a_3t+b_3s) A_4''(x+a_4t+b_4s) \frac{dt}{t} \frac{ds}{s}$$

and it is known that for generic choices of $(a_j)_j$ and $(b_j)_j$ this 4-linear operator does not satisfy any L^p estimates of Hölder type, [10].

References

- [1] CALDERÓN, A. P.: Commutators of singular integral operators. *Proc. Nat. Acad. Sci. USA* **53** (1965), 1092–1099.
- [2] COHEN, J. AND GOSSELIN, J.: A BMO estimate for multilinear singular integrals. *Illinois J. Math.* **30** (1986), 445–465.
- [3] COIFMAN, R. R. AND MEYER, Y.: On commutators of singular integrals and bilinear singular integrals. *Trans. Amer. Math. Soc.* **212** (1975), 315–331.
- [4] COIFMAN, R. R. AND MEYER, Y.: *Wavelets. Calderón–Zygmund and multilinear operators*. Cambridge Studies in Advanced Mathematics 48, Cambridge University Press, Cambridge, 1997.
- [5] COIFMAN, R., MCINTOSH, A. AND MEYER, Y.: L’intégral de Cauchy définit un opérateur borné sur L^2 pour les courbes Lipschitziennes. *Ann. of Math. (2)* **116** (1982), 361–387.
- [6] DAVID, G. AND JOURNÉ, J.-L.: A boundedness criterion for generalized Calderón–Zygmund operators. *Ann. of Math. (2)* **120** (1984), 371–397.
- [7] LACEY, M. AND THIELE, C.: L^p estimates on the bilinear Hilbert transform for $2 < p < \infty$. *Ann. of Math. (2)* **146** (1997), 693–724.

- [8] LACEY, M. AND THIELE, C.: On Calderón's conjecture. *Ann. of Math. (2)* **149** (1999), 475–496.
- [9] MUSCALU, C.: Calderón commutators and the Cauchy integral on Lipschitz curves revisited I. First commutator and generalizations. *Rev. Mat. Iberoam.* **30** (2014), no. 2, 727–750.
- [10] MUSCALU, C.: Some remarks on the n -linear Hilbert transform, for $n \geq 4$. Preprint, arXiv: 1209.6391v3, 2013.
- [11] MUSCALU, C., PIPHER, J., TAO, T. AND THIELE, C.: Bi-parameter paraproducts. *Acta Math.* **193** (2004), 269–296.
- [12] MUSCALU, C., PIPHER, J., TAO, T. AND THIELE, C.: Multi-parameter paraproducts. *Rev. Mat. Iberoam.* **22** (2006), no. 3, 963–976.
- [13] MUSCALU, C., TAO, T. AND THIELE, C.: Multi-linear operators given by singular multipliers. *J. Amer. Math. Soc.* **15** (2002), 469–496.
- [14] VERDERA, J.: L^2 boundedness of the Cauchy integral and Menger curvature. In *Harmonic analysis and boundary value problems (Fayetteville, AR, 2000)*, 139–158. Contemp. Math. 277, Amer. Math. Soc., Providence, 2001.

Received November 12, 2012.

CAMIL MUSCALU: Department of Mathematics, Cornell University, Ithaca, NY, 14850, USA.

E-mail: camil@math.cornell.edu