Forum Math. **19** (2007), 325–351
DOI 10.1515/FORUM.2007.013

Forum
Mathematicum

Paraproducts in one and several parameters

Michael Lacey and Jason Metcalfe

(Communicated by Christopher D. Sogge)

Abstract. For multiparameter bilinear paraproduct operators B we prove the estimate

$$\mathsf{B} : L^p \times L^q \mapsto L^r, \quad 1 < p, q \leq \infty.$$

Here, $1/p + 1/q = 1/r$ and special attention is paid to the case of $0 < r < 1$. (Note that the families of multiparameter paraproducts are much richer than in the one parameter case.) These estimates are the essential step in the version of the multiparameter Coifman-Meyer theorem proved by C. Muscalu, J. Pipher, T. Tao, and C. Thiele [10, 11]. We offer a different proof of these inequalities.

1991 Mathematics Subject Classification: 42B20; 42B25, 42B30.

1 Introduction

Our subject concerns the Coifman-Meyer theorem in a multiparameter setting. Namely, for bounded function $\tau : \mathbb{R}^d \to \mathbb{C}$, we set

$$\mathsf{T}(f_1, \dots, f_d)(x) \stackrel{\text{def}}{=} \int_{\mathbb{R}^d} \tau(\xi) \mathrm{e}^{i(\xi_1 + \cdots + \xi_d)x} \prod_{j=1}^{n} \hat{f}_j(\xi_j)\, d\xi_1 \cdots d\xi_d$$

in which f_j are Schwartz functions on $\mathbb{R}$ and $\hat{g}$ denotes the Fourier transform, namely

$$\hat{g}(\theta) \stackrel{\text{def}}{=} \int g(x) \mathrm{e}^{-i\theta x}\, dx.$$

One is interested in conditions under which T extends to a bounded multilinear operator on a product of L^p spaces. And the motivation for this paper is the Theorem

Research supported in part by an NSF grant and a fellowship from the Guggenheim Foundation. The hospitality of the University of British Columbia is gratefully acknowledged.
Research supported by a NSF VIGRE grant to the Georgia Institute of Technology. The author is a NSF Postdoctoral Fellow.

1.1. Theorem. *Suppose that τ obeys the estimates*

$$(1.2)\qquad \left|\frac{\partial^\alpha}{\partial\xi^\alpha}\tau(\xi)\right| \lesssim |\xi|^{-\alpha} = \prod_{j=1}^{d}|\xi_j|^{-\alpha_j}, \quad |\alpha| \le N,$$

where $\xi = (\xi_1, \dots, \xi_d)$ and likewise for the multiindices α. There is a finite choice of N so that for all $1 < p_j \le \infty$, with not all p_j being equal to infinity, the operator T extends to a bounded linear operator

$$\mathsf{T} : L^{p_1} \times \cdots \times L^{p_d} \mapsto L^r$$

where $1/r = \sum_{j=1}^{d} 1/p_j$.

In the statement of the Theorem, $A \lesssim B$ means that $A \le KB$ for some unspecified constant K.

This theorem has been the subject of wide ranging investigations since the inital results of Coifman and Meyer [3, 4]. The methods and techniques of the proof, built around the subject of paraproducts, is the main focus of this article.

The singularities permitted in τ in (1.2) invoke elements of the product theory of maximal functions, singular integrals, and related subjects. Some cases of the theorem above were found by Journé [8], following the identification of product BMO by S.-Y. Chang and R. Fefferman [1, 2, 5].

The possibility that the image space L^r can have index less than one is the primary new contribution of C. Muscalu, J. Pipher, T. Tao, and C. Thiele [10, 11]. The purpose of this article is to give a somewhat different proof, one that discusses end point issues, and is a little more leisurely than the cited articles.

The method of proving these inequalities is by way of paraproducts. And we take the the latter as the primary focus of this article. See the next section for a definition of the most familiar paraproducts. Many proofs of paraproduct results depend upon the Calderón Zygmund decomposition, which has only weak analogs in the product theory.

A very nice feature of the work of Muscalu et al. is that they find that the proof of the theorem can be understood in terms that avoid the intricacies of the product BMO theory of Chang and Fefferman. We find that some aspects of that theory enter into different endpoint estimates, such as at $p = 1$, where L^1 should be replaced by the Hardy space H^1, and $p = \infty$, where the BMO space enters in.

The main result is Theorem 4.32 below, a discrete form of the Theorem above. This theorem is contained in [10, 11], and our proof borrows elements of theirs. We offer the proof as it differs in some details. In addition, the rich family of multiparameter paraproducts is not necessarily well understood. We hope that this paper advertises [10, 11] in particular, and, more generally, the subject of multiparameter paraproducts.[1]

In the next section, we discuss the one parameter paraproducts. The proofs in this case feature initial details that can be used in the multiparameter case. We present proofs of these results in the special case of paraproducts formed from Haar functions. The subject of mutliparameter paraproducts is taken up in Section 4. Our presentation and proofs in Section 3 have been influenced by the CBMS lectures of C. Thiele [12].

2 One parameter paraproducts

For an interval I, we say that φ is *adapted to I* iff $\|\varphi\|_2 = 1$ and

$$(2.3) \qquad |\mathsf{D}^n \varphi(x)| \lesssim |I|^{-n-1/2}\left(1 + \frac{|x - c(I)|}{|I|}\right)^{-N}, \quad n = 0, 1.$$

Here, $c(I)$ denotes the center of I, and N is a large integer, whose exact value need not concern us, except to say that its value can depend upon the L^p inequalities that we are considering.[2] D denotes the derivative operator. We shall consistently work with functions which have L^2 norm at most one. Some of these functions we will also insist to have integral zero. (Terminology for this will be introduced below.)

Intervals will most typically be dyadic, and we use the notation $\mathscr{D}$ for these intervals. To be specific,

$$\mathscr{D} \stackrel{\text{def}}{=} \{[j2^k, (j+1)2^k) : j, k \in \mathbb{Z}\}.$$

With the control on the function and its derivative in the definition of adapted, elements of Littlewood-Paley theory will apply. Namely, we will have the inequalities (2.16)–(2.18) for the square function constructed from the functions $\{\varphi_I : I \in \mathscr{D}\}$.

Operators are built up from rank one operators $f \mapsto \langle f, \varphi \rangle \varphi'$. A paraproduct is, in its simplest manifestation, of the form

$$\mathsf{B}(f_1, f_2) \stackrel{\text{def}}{=} \sum_{I \in \mathscr{D}} |I|^{-1/2} \varphi_{3,I} \prod_{j=1}^{2} \langle f_j, \varphi_{j,I} \rangle.$$

Here, the functions $\varphi_{j,I}$, for $j = 1, 2, 3$ are adapted to I. Two of these three functions are assumed to be of integral zero. We should emphasize that each individual summand is of the form

[1] A substantial part of the difficulties in [7, 9] is attributable to the variety of paraproducts in the multiparameter setting.

[2] It will be clear in the sequel that $N = \min(3p_1 + 4, 3p_2 + 4)$ is sufficient for the bilinear case, for example. The main size requirement can be found in (4.67).

$$(f_1, f_2) \mapsto |I|^{-1/2} \varphi_{3,I} \prod_{j=1}^{2} \langle f_j, \varphi_{j,I} \rangle.$$

This is certainly a bounded operator from, say, $L^2 \times L^2 \to L^1$, and our desired conclusion is that the same is true for the sums above.

We will also consider higher linearities

$$\mathsf{B}(f_1, f_2, \dots, f_n) \stackrel{\text{def}}{=} \sum_{I \in \mathscr{D}} |I|^{-(n-1)/2} \varphi_{n+1,I} \prod_{j=1}^{n} \langle f_j, \varphi_{j,I} \rangle \tag{2.4}$$

where the functions $\varphi_{j,I}$ are assumed to be adapted to I and two are of integral zero. In the course of the proofs, it is convenient to consider the $n+1$ sublinear forms

$$\Lambda(f_1, f_2, \dots, f_{n+1}) \stackrel{\text{def}}{=} \sum_{I \in \mathscr{D}} |I|^{-(n-1)/2} \prod_{j=1}^{n+1} |\langle f_j, \varphi_{j,I} \rangle|. \tag{2.5}$$

Notice that this just assigns a number to the $n+1$ tuple of functions and that it dominates $\langle \mathsf{B}(f_1, f_2, \dots, f_n), f_{n+1} \rangle$. It is also of interest to consider the related sublinear operator

$$\mathsf{L}(f_1, f_2, \dots, f_n, f_{n+1}) \stackrel{\text{def}}{=} \sum_{I \in \mathscr{D}} |I|^{-(n+1)/2} \left[\prod_{j=1}^{n+1} |\langle f_j, \varphi_{j,I} \rangle| \right] \mathbf{1}_I. \tag{2.6}$$

In particular, if L maps a product of Banach spaces into L^1, then we conclude that Λ is bounded on a related product of spaces, see (2.20).

2.7. Theorem. *For $n \geq 3$ and $1 < p_j \leq \infty$, define $\frac{1}{r} = \sum_{j=1}^{n-1} \frac{1}{p_j}$. Then,*

$$\mathsf{B} : \bigotimes_{j=1}^{n-1} L^{p_j} \to L^r. \tag{2.8}$$

In addition we have the endpoint estimates:

$$\mathsf{B} : \bigotimes_{j=1}^{n-1} L^1 \to L^{1/(n-1), \infty} \tag{2.9}$$

$$\mathsf{B} : \bigotimes_{j=1}^{s} L^1 \times \bigotimes_{j=1}^{n-1-s} BMO \to L^{1/s, \infty}. \tag{2.10}$$

In this last display, we require that the functions $\varphi_{j,I}$ have integral zero for any choice of j for which f_j is only assumed to be BMO.

The estimates above follow immediately from the corresponding estimates for the sublinear operator, in the case that the index of the range is between 1 and ∞. Namely for $1 < r < \infty$, this can be derived from

2.11. Theorem. *For $n \geq 3$ and $1 < p_j \leq \infty$, define $\frac{1}{s} = \sum_{j=1}^{n} \frac{1}{p_j}$. If it is the case that $0 < s < \infty$, then*

$$\|\mathsf{L}(f_1, f_2, \ldots, f_n)\|_s \lesssim \prod_{j=1}^{n} \|f\|_{p_j}. \tag{2.12}$$

In the case that any $p_j = 1$, then L^1 can be replaced by H^1, and the estimate above is true. If we do not replace L^1 by H^1, then only the weak type inequality is true. In particular, we have the estimate

$$\mathsf{L} : \bigotimes_{j=1}^{n} L^1 \to L^{1/n,\infty}. \tag{2.13}$$

In the case that any of p_j equal ∞ and the functions $\varphi_{j,I}$ have integral zero for all I, then the space L^∞ above can be replaced by BMO.

The essential case is that of $n = 3$ above, and to avoid unnecessary notations, that is the case discussed in the proof. Thus, we have three functions $\varphi_{j,I}$. Two of these are assumed to be of integral zero. Due to the symmetry of the estimates we are to prove, we can assume that these two functions occur for $j = 2, 3$.

This in particular means that we have the estimate

$$\sup_{I \in \mathscr{D}} \mathbf{1}_I(x) \frac{|\langle f_1, \varphi_{1,I} \rangle|}{\sqrt{|I|}} \lesssim \mathsf{M}(f_1)(x) \tag{2.14}$$

where M denotes the maximal function. There is another bound that applies to the second and third functions. Namely, we set

$$\mathsf{S}_j g \stackrel{\text{def}}{=} \left[\sum_{I \in \mathscr{D}} \frac{|\langle g, \varphi_{j,I} \rangle|^2}{|I|} \mathbf{1}_I \right]^{1/2}. \tag{2.15}$$

It is a consequence of the integral zero assumption placed on the functions $\varphi_{2,I}$ and $\varphi_{3,I}$ that the usual Littlewood-Paley theory applies to these square functions. Therefore, they map all L^p into themselves, for $1 < p < \infty$, and we have the usual endpoint estimates. To be explicit, these estimates are

$$\|\mathsf{S}_j f\|_p \lesssim \|f\|_p, \quad 1 < p < \infty, \tag{2.16}$$

$$\|\mathsf{S}_j f\|_1 \lesssim \|f\|_{H^1}, \tag{2.17}$$

$$\|\mathsf{S}_j f\|_{BMO} \lesssim \|f\|_\infty. \tag{2.18}$$

Note that as we are using the maximal function and square functions, we have access to the following upper bound for numerical sequence

$$\sum_n \prod_{j=1}^{3} a_{j,n} \le \|a_{1,n}\|_{\ell^\infty} \|a_{2,n}\|_{\ell^2} \|a_{3,n}\|_{\ell^2}. \tag{2.19}$$

2.1. Generalities on the proof. If $1 < p_j < \infty$ and $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 1$, we can estimate, based on (2.19) and Hölder's inequality,

$$\begin{aligned} &\int \mathsf{B}(f_1, f_2)\overline{f_3}\, dx \\ &\le \int \mathsf{L}(f_1, f_2, f_3)\, dx \le \int (\mathsf{M}f_1) \cdot (\mathsf{S}_2 f_2) \cdot (\mathsf{S}_3 f_3)\, dx \\ &\le \|\mathsf{M}f_1\|_{p_1} \|\mathsf{S}_2 f_2\|_{p_2} \|\mathsf{S}_3 f_3\|_{p_3} \lesssim \|f_1\|_{p_1} \|f_2\|_{p_2} \|f_3\|_{p_3}. \end{aligned} \tag{2.20}$$

This argument also applies when $p_1 = \infty$. When, however, any of the $p_j = 1$, one should replace the L^1 norm on f_j with the H^1 norm.

The argument must be modified when, e.g., $p_2 = \infty$. For then the square function $\|\mathsf{S}_2 f_2\|_\infty$ is no longer bounded. And indeed, the sharp estimate on the square function replaces L^∞ with BMO.

Alternate methods are required when duality cannot be applied. Here, we shall obtain inequalities of weak type. For example,

$$\lambda |\{\mathsf{B}(f_1, f_2) > \lambda\}|^{1/r} \lesssim \|f_1\|_{p_1} \|f_2\|_{p_2}, \quad \frac{1}{p_1} + \frac{1}{p_2} = \frac{1}{r}. \tag{2.21}$$

Interpolation will then supply the strong type inequalities, except for the endpoint estimates.

As the class of operators B we consider are invariant under dilations by powers of 2, this inequality follows from

$$|\{\mathsf{B}(f_1, f_2) > K\}| \le 1 \tag{2.22}$$

where K is an absolute constant, and the inequality holds for all choices of smooth compactly supported functions f_j with L^{p_j} norm 1.

The usefulness of this observation is already evident in that we have the following (obvious) estimate

$$\|fg\|_{1/2,\infty} \lesssim \|f\|_{1,\infty} \|g\|_{1,\infty}. \tag{2.23}$$

This inequality immediately generalizes to general products and indices. We use this generalization below.

One can effectively use the symmetry in the formulation of the paraproducts in passing to the sublinear function Λ, and considering weak type inequalities for it. Namely, for $1 \le p_1, p_2 \le \infty$, we define p_3 by $\frac{1}{p_1} + \frac{1}{p_2} + \frac{1}{p_3} = 1$. In particular, p_3 can be negative: For $p_1 = p_2 = 1$, we have $\frac{1}{p_3} = -1$, which we interpret as the dual index to $\frac{1}{2}$. Let $X(E)$ be the space of functions supported on a measurable set $E \subset \mathbb{R}$ and bounded by 1. We then prove the inequality

$$(2.24)\quad |\Lambda(f_1, f_2, f_3)| \lesssim |E_3|^{1/p_3} \prod_{j=1}^{2} \|f_j\|_{p_j} \quad f_3 \in X(E_3).$$

Observe that this implies (2.22). Also observe that the inequality for Λ follows from the following formulation: For all E_3, we can choose $E_3' \subset E_3$ with $|E_3'| \ge \frac{1}{2}|E_3|$, and

$$(2.25)\quad |\Lambda(f_1, f_2, f_3)| \lesssim |E_3|^{1/p_3} \prod_{j=1}^{2} \|f_j\|_{p_j} \quad f_3 \in X(E_3').$$

By dilation invariance, it suffices to prove this estimate in the case that $\|f_1\|_{p_1} = \|f_2\|_{p_2} = 1$ and $|E_3| = 1$. All of these comments apply equally well in the multi-parameter case.

2.2. H^1 and BMO. We will restrict ourselves to the dyadic versions of the real Hardy space H^1 and its dual BMO.

The Haar functions are

$$(2.26)\quad h_I = |I|^{-1/2}(\mathbf{1}_{I_-} - \mathbf{1}_{I_+}), \quad I \in \mathcal{D},$$

where I_- (I_+) is the left (right) half of I. These functions form a basis for L^2. The dyadic square function from Haar functions is formed as follows.

$$\mathsf{S}f \stackrel{\text{def}}{=} \left[\sum_{I \in \mathcal{D}} \frac{|\langle f, h_I \rangle|^2}{|I|} \mathbf{1}_I \right]^{1/2}.$$

We define the real dyadic Hardy space H^1 as those functions f with

$$\|f\|_{H^1} \stackrel{\text{def}}{=} \|f\|_1 + \|\mathsf{S}f\|_1 < \infty.$$

The dual to H^1 is BMO. This space has the equivalent norm

$$(2.27)\quad \|f\|_{BMO} \stackrel{\text{def}}{=} \sup_{J \in \mathcal{D}} \left[|J|^{-1} \sum_{I \subset J} |\langle f, h_I \rangle|^2 \right]^{1/2}.$$

These spaces are substitutes for L^1 and L^∞. In the current setting sharp endpoint estimates can be phrased in terms of these norms. And there is a rich interpolation theory between these spaces.

3 Proofs in the one parameter case

The case of Haar paraproducts is the only case that we consider in the one parameter case. The rationale is the proof in the multiparameter case includes the one parameter case as a special instance. In addition, the Haar case is especially attractive, due to the presence of the dyadic grid.

A particular way that it enters is this. Suppose that $\mathscr{I}$ is a collection of disjoint dyadic intervals, not necessarily a partition of $\mathbb{R}$. We define the conditional expectation with respect to $\mathscr{I}$ as

$$\mathbb{E}(f|\mathscr{I})(x) \stackrel{\text{def}}{=} \begin{cases} |I|^{-1}\int_I f(y)\,dy & x \in I,\ I \in \mathscr{I}, \\ f(x) & x \notin \bigcup_{I\in\mathscr{I}} I \end{cases}$$

We leave it as an exercise that these properties of the conditional expectation are true.

(1) Integrals are preserved under conditional expectation: $\int f\,dx = \int \mathbb{E}(f|\mathscr{I})\,dx$.

(2) $f \mapsto \mathbb{E}(f|\mathscr{I})$ is a projection.

(3) $f \mapsto \mathbb{E}(f|\mathscr{I})$ is of norm one on all L^p, $1 \le p \le \infty$.

(4) $f \mapsto \mathbb{E}(f|\mathscr{I})$ is bounded as a map from dyadic H^1 into itself.

We first turn to the range of inequalities for the sublinear operator L and the proof of Theorem 2.11. Observe that by (2.19), we have

$$\mathsf{L}(f_1, f_2, f_3) \le \mathsf{M}f_1 \cdot \mathsf{S}f_2 \cdot \mathsf{S}f_3. \tag{3.28}$$

Here, we assume that we have mean zero in the second and third places, and we continue with this assumption below. To be specific, the sublinear operator is

$$\mathsf{L}(f_1, f_2, f_3) = \sum_{I\in\mathscr{D}} |I|^{-3/2}\mathbf{1}_I|\langle f_1, |h_I|\rangle\langle f_2, h_I\rangle\langle f_3, h_I\rangle|.$$

For f_1, we form the inner product with the absolute value of the Haar function. The inequalities in (2.12) then follow from Hölder's inequality, provided that all we are not discussing an endpoint estimate. When $s < 1$, one can instead apply an appropriate version of (2.23).

If any $p_j = 1$, then we only conclude that $\mathsf{M}f_j$ and $\mathsf{S}_j f_j$ are in $L^{1,\infty}$. But we can apply (2.23) to conclude the weak type estimate. If any $p_j = 1$ and $f_j \in H^1$, then we conclude that both $\mathsf{M}f_j$ and $\mathsf{S}_j f_j$ are in L^1, so that again Hölder's inequality or (2.23) will apply.

We concern ourselves with the endpoint estimates where either of p_2, p_3 is infinity and L^∞ is replaced with BMO. One class of inequalities are in fact easily available; they are

$$(3.29)\qquad \begin{aligned} &\mathsf{L} : L^\infty \otimes BMO \otimes H^1 \to L^1, \\ &\mathsf{L} : L^\infty \otimes BMO \otimes BMO \to BMO. \end{aligned}$$

Notice that these estimates can be interpolated by standard linear methods.

Since the Haar functions are an unconditional basis for both H^1 and BMO,[3] we can conclude that

$$\begin{aligned} \int \mathsf{L}(f_1, f_2, f_3)\, dx &= \sum_{I \in \mathscr{D}} \frac{|\langle f_1, |h_I| \rangle|}{\sqrt{|I|}} |\langle f_2, h_I \rangle|\, |\langle f_3, h_I \rangle| \\ &\lesssim \|f_1\|_\infty \|f_2\|_{BMO} \|f_3\|_{H^1}. \end{aligned}$$

This proves the first bound.

For the BMO estimate, for each dyadic interval J we have

$$\begin{aligned} \sum_{I \subset J} \frac{|\langle f_1, |h_I| \rangle|}{\sqrt{|I|}} |\langle f_2, h_I \rangle|\, |\langle f_3, h_I \rangle| &\lesssim \|f_1\|_\infty \prod_{j=2}^{2} \Bigl[\sum_{I \subset J} |\langle f_j, h_I \rangle|^2 \Bigr]^{1/2} \\ &\lesssim |J|\, \|f_1\|_\infty \prod_{j=1}^{2} \|f_j\|_{BMO}. \end{aligned}$$

This concludes the proof of the estimates (3.29).

The last estimates to prove are these:

$$\mathsf{L} : L^{p_1} \otimes BMO \otimes L^{p_3} \to L^s, \quad \frac{1}{s} = \frac{1}{p_1} + \frac{1}{p_3}.$$

At this point we make a more substantive reliance on the dyadic structure. The strategy is first to prove the weak type inequalities, and in particular (2.24). Namely, we will choose an exceptional set on which we will not attempt to estimate $\mathsf{L}(f_1, f_2, f_3)$ and a conditional expectation to apply to f_1, after which we will have a bounded function in the first coordinate. But then we will be in a situation for which we can appeal to the estimates in which we have duality.

[3] While we are specifically appealing to the properties of the Haar functions here, this aspect does generalize to the non-Haar functions.

We will prove that L satisfies (2.24). Thus, fix f_j functions in the appropriate spaces, of norm one. Define

$$E \overset{\text{def}}{=} \{\mathsf{M}f_1 > 1\}.$$

We do not attempt to estimate L on this set. That has the practical implication that we need only consider the sum

$$\mathsf{L}_E(f_1) \overset{\text{def}}{=} \sum_{I \not\subset E} |I|^{-1} \frac{|\langle f_1, |h_I| \rangle|}{\sqrt{|I|}} |\langle f_2, h_I \rangle| \, |\langle f_3, h_I \rangle| \mathbf{1}_I$$

(Recall that we are assuming that f_2 and f_3 are fixed.) Let $\mathscr{I}$ be the collection of maximal dyadic subintervals of E, and set $g_1 \overset{\text{def}}{=} \mathbb{E}(f_1 | \mathscr{I})$.

Note that by construction we will have the estimate $\|g_1\|_\infty \leq 2$. For otherwise, let J be the smallest dyadic interval that strictly contains I (i.e. the parent of I), and observe that $\int_J |g| \, dx \geq |J|$. That is, we contradict the maximality of I.

In addition, for each dyadic interval I not contained in E, we have $\int_I f_1 \, dx = \int_I g \, dx$. Thus, we have

$$\mathsf{L}_E(f_1) = \mathsf{L}_E(g_1). \tag{3.30}$$

Therefore, using (3.29), we can estimate

$$\begin{aligned} &|\{\mathsf{L}(f_1, f_2, f_3) > 1\}| \\ &\leq |E| + |\{\mathsf{L}(g_1, f_2, f_3) > 1\}| \lesssim 1 + [\|g_1\|_\infty \|f_2\|_{BMO} \|f_3\|_{p_3}]^{p_3} \lesssim 1. \end{aligned}$$

Our discussion of the estimates in Theorem 2.11 is complete.

Let us turn to the bilinear operator $\mathsf{B}(f_1, f_2)$ and the proof of Theorem 2.7. In the inequality (2.8), if the index r of the target space is between 1 and ∞, then we can appeal to duality, as is done explicitly in (2.20).

We discuss the proof of the weak type bounds for B, in the case that duality does not apply, namely $\frac{1}{2} \leq r < 1$. Marcinkiewicz interpolation will then deduce the strong type L^r inequalities.

In so doing, we need only prove (2.22), and we will repeat the use of conditional expectation in the argument (3.30) above. Take $f_j \in L^{p_j}$ of norm one, for $j = 1, 2$. Suppose that we have in fact $\|f_j\|_\infty \leq 1$. We conclude that in fact $\|f_j\|_q \leq 1$ for all $p_j < q < \infty$, and so for $q > 4$ large, we can use the proven bound of $L^q \times L^q$ into $L^{q/2}$ to conclude that

$$|\{\mathsf{B}(f_1, f_2) > K\}| \leq 1.$$

The general case can be reduced to this situation.

Define

$$E \stackrel{\text{def}}{=} \bigcup_{j=1}^{2} \{\mathsf{M} f_j > 1\}, \quad F \stackrel{\text{def}}{=} \left\{\mathsf{M}\mathbf{1}_E > \frac{1}{2}\right\}.$$

Clearly, the set F has measure bounded by an absolute constant. We will now estimate $\mathsf{B}(f_1, f_2)$ on the set F. Define

$$\mathsf{B}_F(f_1, f_2) = \sum_{I \not\subset F} |I|^{-1/2} h_I \langle f_1, |h_I| \rangle \langle f_2, h_I \rangle,$$

and set $\mathscr{I}$ to the collection of maximal dyadic intervals contained in F. We set $g_j \stackrel{\text{def}}{=} \mathbb{E}(f_j | \mathscr{I})$. Then, certainly we have $\|g_j\|_\infty \leq 1$. We claim that

$$\mathsf{B}_F(g_1, g_2)\mathbf{1}_{F^c} = \mathsf{B}(f_1, f_2)\mathbf{1}_{F^c}. \tag{3.31}$$

And this will complete our proof.

Suppose that I is a dyadic interval that is not contained in F. The Haar function associated to I is constant on the two sub halves of I, which we denote as $I_\pm$. By our definition of F, neither $I_\pm$ can be contained in E, hence we have

$$\int_{I_\pm} f_j \, dx = \int_{I_\pm} g_j \, dx.$$

This proves (3.31), and so we have completed the proof of the norm bounds for B.

4 Multiparameter paraproducts

We now consider paraproducts formed over sums of dyadic rectangles in $\mathbb{R}^d$. The class of paraproducts is then invariant under a d parameter family of dilations, a situation that we refer to as "multi-parameter."[4]

Let us say that a function φ *is adapted to a rectangle* $R = \bigotimes_{j=1}^{d} R_j$ iff $\varphi(x_1, \ldots, x_d) = \prod_{j=1}^{d} \varphi_j(x_j)$, with each φ_j adapted to the interval R_j in the sense of (2.3).

Our paraproducts are of the same general form

$$\mathsf{B}(f_1, f_2, \ldots, f_n) \stackrel{\text{def}}{=} \sum_{R \in \mathscr{R}} \frac{\varphi_{n+1, R}}{|R|^{(n-1)/2}} \prod_{v=1}^{n} \langle f_v, \varphi_{v, R} \rangle.$$

[4] In this paper, the number of dimensions will be the number of parameters. In general, the two are however distinct. Consider $\mathbb{R}^{d_1} \otimes \mathbb{R}^{d_2}$ and rectangles in this space formed from a cube in each space $\mathbb{R}^{d_j}$.

Here, we let $\mathscr{R} \stackrel{\text{def}}{=} \mathscr{D}^d$ be the class of dyadic rectangles. With the obvious changes, we will also use the notations for the sublinear forms and operators given in (2.5) and (2.6).

The Theorem in this setting is

4.32. Theorem. *Let $n \geq 2$ and $1 < p_v \leq \infty$ for $1 \leq v \leq n$, and define $\frac{1}{r} = \sum_{v=1}^{n} \frac{1}{p_v}$. Assume that for each choice of coordinate $1 \leq j \leq d$, there are two choices of $1 \leq v \leq n+1$ for which we have*

$$(4.33) \qquad \int_{\mathbb{R}} \varphi_{v,R}(x_1, x_2, \ldots, x_n)\, dx_j = 0, \quad \textit{for all } x_k \textit{ with } k \neq j \textit{ and all } R.$$

Then, we have the inequality

$$(4.34) \qquad \mathsf{B} : \bigotimes_{v=1}^{n} L^{p_v} \to L^r.$$

Assume that the functions $\varphi_{v,R}$ satisfy (4.33) for all j. In the instance that $p_v = 1$, the inequality remains true if we replace L^1 by H^1 defined below. In the instance that $p_v = \infty$, we can replace L^∞ by the larger space $BMO = (H^1)^*$ defined below.

The critical distinction comes from the assumption about the zeros, (4.33). Let us say that *there are x_j zeros in the vth position* iff

$$(4.35) \qquad \int \varphi_{v,R}(x_1, \ldots, x_d)\, dx_j = 0 \quad \text{for all } x_k \text{ with } k \neq j.$$

And so our assumption is that for each $1 \leq j \leq d$ there are two choices of v for which we have zeros in the vth position.

Again, the critical case is $n = 2$, so that B is bilinear. There are essentially d distinct cases. The first case, with the greatest similarity to the one parameter case, is where we have, for example, x_j zeros in first and second positions for all $1 \leq j \leq d$. The other cases do not have a proper analog in the one parameter case.

4.1. H^1 and BMO. We turn to the product Hardy space theory, as developed by S.-Y. Chang and R. Fefferman [1, 2]. This section is not strictly speaking needed, but does inform the modes of proof below.

$H^1(\mathbb{C}_+^d)$ will denote the d-fold product real valued Hardy space. This space consists of functions $f : \mathbb{R}^d \to \mathbb{R}$ where $\mathbb{R}^d$ is viewed as the boundary of

$$\mathbb{C}_+^d = \bigotimes_{j=1}^{d} \{z \in \mathbb{C} : \operatorname{Re}(z) > 0\}.$$

We require that there is a function $F : \mathbb{C}_+^d \to \mathbb{C}$ that is holomorphic in each variable separately and

$$f(x) = \lim_{\|y\| \to 0} \operatorname{Re} F(x_1 + iy_1, \ldots, x_d + iy_d).$$

The norm of f is taken to be

$$\|f\|_{H^1} = \lim_{y_1 \downarrow 0} \cdots \lim_{y_d \downarrow 0} \|F(x_1 + iy_1, \ldots, x_d + iy_d)\|_{L^1(\mathbb{R}^d)}.$$

Product $H^1(\mathbb{C}_+^d)$ has the equivalent norm

$$\|f\|_{H^1} \stackrel{\text{def}}{=} \|f\|_1 + \|\mathsf{S}f\|_1 \tag{4.36}$$

where S is the square function formed over the product Haar basis

$$\mathsf{S}f \stackrel{\text{def}}{=} \left[\sum_{R \in \mathscr{R}} \frac{|\langle f, h_R \rangle|^2}{|R|} \mathbf{1}_R \right]^{1/2}.$$

For a rectangle $R = \prod_{j=1}^d R_{(j)} \in \mathscr{D}^d$, we have set

$$h_R(x_1, \ldots, x_d) = \prod_{j=1}^d h_{R_{(j)}}(x_j).$$

The last product is over one dimensional Haar functions as in (2.26). The basis $\{h_R : R \in \mathscr{D}^d\}$ is the d-fold tensor product of the Haar basis.

The dual of $H^1(\mathbb{C}_+^d)$ is $BMO(\mathbb{C}_+^d)$, the d-fold product BMO space. It is a Theorem of S.-Y. Chang and R. Fefferman [2] that this space has an explicit characterization in terms of the product Haar basis. In particular, Chang and Fefferman showed that the product BMO space has the equivalent norm

$$\|b\|_{BMO} = \sup_{U \subset \mathbb{R}^d} \left[|U|^{-1} \sum_{R \subset U} |\langle b, h_R \rangle|^2 \right]^{1/2}$$

where it essential that the supremum be formed over all subsets $U \subset \mathbb{R}^d$ of finite measure.

4.2. The governing operators. We describe a range of operators, which encompass the d parameter maximal function at one end and the d parameter square function at the other. These operators, as we shall see, govern the behaviors of these paraproducts.

To be explicit, in the two parameter setting, these operators are as follows. First we have the maximal function,

$$\mathsf{MM}f \overset{\mathrm{def}}{=} \sup_{R \in \mathcal{R}} \frac{|\langle f, \varphi_R \rangle|}{\sqrt{|R|}} \mathbf{1}_R,$$

which is a variant of the strong maximal function.

The reason for the iterated style notation becomes clearer with the second type of governing operator. It is

$$\mathsf{S}_1\mathsf{M}_2 f \overset{\mathrm{def}}{=} \Bigg[\sum_{R_{(1)} \in \mathcal{D}} \sup_{R_{(2)} \in \mathcal{D}} \frac{|\langle f, \varphi_{R_{(1)} \times R_{(2)}} \rangle|^2}{|R|} \mathbf{1}_R \Bigg]^{1/2}, \quad R = R_{(1)} \times R_{(2)}.$$

In order for this to be a bounded operator, the functions $\{\varphi_R\}$ must have zeros in the first coordinate. But then, the operator will be bounded on all L^p's for $1 < p < \infty$. There is also the operator $\mathsf{S}_2\mathsf{M}_1$ in which the role of the coordinates is changed.

A third type of operator is

$$\mathsf{M}_1\mathsf{S}_2 f \overset{\mathrm{def}}{=} \sup_{R_{(1)} \in \mathcal{D}} \Bigg[\sum_{R_{(2)} \in \mathcal{D}} \frac{|\langle f, \varphi_{R_{(1)} \times R_{(2)}} \rangle|^2}{|R|} \mathbf{1}_R \Bigg]^{1/2}, \quad R = R_{(1)} \times R_{(2)}.$$

The functions $\{\varphi_R\}$ must now have zeros in the second coordinate. And there is a corresponding operator $\mathsf{M}_2\mathsf{S}_1$ in which the roles of the coordinates are reversed.

A fourth type of operator is the familiar two parameter square function

$$\mathsf{SS}f \overset{\mathrm{def}}{=} \Bigg[\sum_{R \in \mathcal{R}} \frac{|\langle f, \varphi_R \rangle|^2}{|R|} \mathbf{1}_R \Bigg]^{1/2}.$$

Here, we require that the functions φ_R have zeros in both coordinates. As with the maximal function MM, the subscripts are not needed in this case.

In general, we set T_j to be either the square function S or the maximal function M, both formed over a set of functions $\{\varphi_I : I \in \mathcal{D}\}$ acting on the jth coordinate. For a permutation of the coordinates $\pi : \{1, \dots, d\} \mapsto \{1, \dots, d\}$, set

$$\mathsf{T} \overset{\mathrm{def}}{=} \mathsf{T}_{\pi(1)} \cdots \mathsf{T}_{\pi(d)}. \tag{4.37}$$

The subscript $\pi(j)$ indicates in which coordinate the operator $\mathsf{T}_{\pi(j)}$ operates. In each position in which coordinate $\mathsf{T}_{\pi(j)}$ is the square function, we require that the functions $\{\varphi_R : R \in \mathscr{R}\}$ have zeros in that coordinate. Note that these operators can be viewed as

$$(4.38)\quad \mathsf{T}f(x) = \left\| \cdots \left\| \left\{ \frac{|\langle f, \varphi_R \rangle|}{\sqrt{|R|}} \mathbf{1}_R : R \right. \right. \right.$$

$$\left. \left. \left. = R_{(1)} \times \cdots \times R_{(d)} \in \mathscr{R} \right\} \right\|_{\ell^{\sigma(\pi(d))}(R_{(\pi(d))})} \cdots \right\|_{\ell^{\sigma(\pi(1))}(R_{(\pi(1))})}$$

where $\sigma(j)$ is either 2 or ∞ for all j.

The point of these definitions is that for all paraproducts B of d parameters, there are three choices of T_k, $k = 1, 2, 3$, operators as in in (4.37) for which we have

$$(4.39)\quad \langle \mathsf{B}(f_1, f_2), f_3 \rangle \le \int \prod_{k=1}^{3} \mathsf{T}_k(f_k)\, dx.$$

This is a consequence of the essential hypothesis on there being two zeros in each coordinate. In those two positions, one uses the square function. In every other position, the maximal function is used.

***L^p* bounds for the operators T.** Let us discuss the mapping properties of these operators, beginning with the maximal operator. We have been careful to insist that the functions φ_R are products of adapted functions. Thus, in the two parameter case, we can appeal to the one parameter maximal function twice, as follows.

$$\|\mathsf{M}_1\mathsf{M}_2 f\|_p \lesssim \|\mathsf{M}_1\{\mathsf{M}_2 f\}\|_p \lesssim \|\mathsf{M}_2 f\|_p \lesssim \|f\|_p.$$

Likewise, by a d fold iteration of this argument, it follows that

$$\|\mathsf{M} \cdots \mathsf{M} f\|_p \lesssim \|f\|_p, \quad 1 < p \le \infty.$$

The same estimates are true for the square function, but are not as straightforward to deduce.

4.40. Lemma. *Assume that the functions $\{\varphi_R\}$ have zeros in every coordinate. Then we have the inequalities below*

$$(4.41)\quad \|\mathsf{S}\cdots\mathsf{S}f\|_p \lesssim \|f\|_p, \quad 1<p<\infty,$$

$$(4.42)\quad \|\mathsf{S}\cdots\mathsf{S}f\|_1 \lesssim \|f\|_{H^1}.$$

At the L^∞ endpoint, the correct estimate is

$$(4.43)\quad \sup_U |U|^{-1} \sum_{R\subset U} |\langle f, \varphi_R\rangle|^2 \lesssim \|f\|_{L^\infty}^2.$$

The supremum is formed over all subsets $U \subset \mathbb{R}^d$ of finite measure.

Proof. We should consider the one parameter inequality

$$(4.44)\quad \|\mathsf{S}f\|_2 \lesssim \|f\|_2,$$

as this will explain in part the assumptions used in the definition of adapted (2.3).

Consider first the inner product $\langle \varphi_I, \varphi_J\rangle$ for two dyadic intervals $|I| \le |J|$. Using the fact that φ_I has integral zero and that φ_J admits a control on it's first derivative, we estimate

$$\rho(I,J) \overset{\text{def}}{=} |\langle \varphi_I, \varphi_J\rangle|$$
$$\le \int |\varphi_I(x)[\varphi_J(x) - \varphi_J(c(I))]|\,dx \le \left(\frac{|I|}{|J|}\right)^{3/2}\left(1 + \frac{|c(I)-c(J)|}{|J|}\right)^{-2}.$$

Here recall that $c(I)$ is the center of I. For this particular argument we only need $N = 3$, say, though other parts of the argument require higher values.

Now, observe that we have

$$\rho_1 \overset{\text{def}}{=} \sup_I \sum_{J:|I|\le|J|} \rho(I,J) < \infty,$$

$$\rho_2 \overset{\text{def}}{=} \sup_J \sum_{I:|I|\le|J|} \rho(I,J) < \infty.$$

To prove (4.44), observe that by duality, it suffices to prove the estimate

$$\left\|\sum_I a_I\varphi_I\right\|_2 \lesssim \left[\sum_I |a_I|^2\right]^{1/2}.$$

Assume that the right hand side is one, and estimate the square of the left hand side as follows, with a generous use of the Cauchy-Schwarz inequality.

$$\begin{aligned}&\Big\|\sum_I a_I\varphi_I\Big\|_2\\&\le 2\sum_{|I|\le|J|}|a_Ia_J|\rho(I,J)\le\left[\sum_I\Big|\sum_{J:|I|\le|J|}|a_J|\rho(I,J)\Big|^2\right]^{1/2}\\&\le 2\left[\sum_I\Big\{\sum_{J:|I|\le|J|}|a_J|^2\rho(I,J)\Big\}\Big\{\sum_{J:|I|\le|J|}\rho(I,J)\Big\}\right]^{1/2}\le 2\rho_1\rho_2.\end{aligned}$$

This completes the proof of (4.44).

The proof of (4.41) in the case of $p=2$ follows from an iteration of the one parameter result, just like the argument for the maximal function.

The H^1 to L^1 estimate is an easy consequence of the definition of the H^1 norm in (4.36).

The last estimate (4.43) is not as accessible, as it relies upon a "localization lemma" Lemma 4.48. Fix a set $U\subset\mathbb{R}^d$ of finite measure and a function f bounded uniformly by one.

Set T_0 to be the d parameter maximal function, and define a sequence of functions f_k by taking

$$f_0\stackrel{\text{def}}{=}f\mathbf{1}_{\{\mathsf{T}_0\mathbf{1}_U>2^{-1}\}},$$

and then inductively define f_k, $k\ge 1$ by

$$f_0+\cdots+f_k=f\mathbf{1}_{\{\mathsf{T}_0\mathbf{1}_U>2^{-1-k}\}}. \tag{4.45}$$

Of course we have

$$\sum_{R\subset U}|\langle f_0,\varphi_R\rangle|^2\lesssim|U|$$

by the L^2 boundedness of the square function. By Lemma 4.48, for integers $k\ge 1$ we have

$$\sum_{R\subset U}|\langle f_k,\varphi_R\rangle|^2\lesssim 2^{-k}|U| \tag{4.46}$$

for an appropriate choice of integer N in the definition of adapted, (2.3). As this last estimate is summable in k, the proof is complete. □

The case of general operators T is treated in this proposition.

4.47. Proposition. *All possible operators* T *as in* (4.37) *map* L^p *into itself for all* $1 < p < \infty$. *This holds provided the functions* $\{\varphi_R : R \in \mathscr{R}\}$ *have zeros in each coordinate in which* T *is equal to a square function. The norm depends only on the constants that enter into the definition of adapted in* (2.3).

Proof. In the two parameter case, observe that

$$\mathsf{M}_1\mathsf{S}_2 \le \mathsf{S}_2\mathsf{M}_1,$$

the inequality holding pointwise. More generally, for any operator T as in (4.37), we can make the operator larger by moving all maximal functions to the right of all square functions. Thus, it suffices to bound operators of the form

$$\mathsf{T} = \mathsf{S}_1 \cdots \mathsf{S}_v \mathsf{M}_{v+1} \cdots \mathsf{M}_d.$$

Recall that the maximal function is bounded as a vector valued map from $L^p(\ell^2)$ into itself for all $1 < p < \infty$. This is a well known result of C. Fefferman and Stein [6]. Namely, we have the estimate

$$\left\| \left[\sum_n |\mathsf{M} \cdots \mathsf{M} g_n|^2 \right]^{1/2} \right\|_p \lesssim \left\| \left[\sum_n |g_n|^2 \right]^{1/2} \right\|_p.$$

For a function $f \in L^p(\mathbb{R}^d)$ and a dyadic rectangle $R_v \in \mathscr{D}^v$, set

$$f_{R_v}(x_1, \ldots, x_v, x_{v+1}, \ldots, x_d)$$

$$\stackrel{\text{def}}{=} \varphi_{R_v}(x_1, \ldots, x_v) \int f(x_1, \ldots, x_v, x_{v+1}, \ldots, x_d) \overline{\varphi_{R_v}(x_1, \ldots, x_v)}\, dx_1 \ldots dx_v.$$

It is a consequence of the one variable Littlewood-Paley inequalities that we have

$$\left\| \left[\sum_{R_v \in \mathscr{D}^v} |f_{R_v}|^2 \right]^{1/2} \right\|_p \lesssim \|f\|_p.$$

To conclude, observe that we have

$$\mathsf{T}f \lesssim \left[\sum_{R_v \in \mathscr{D}^v} |\mathsf{M}_{v+1} \cdots \mathsf{M}_d f_{R_v}|^2 \right]^{1/2}.$$

The L^p norm of the last quantity is clearly bounded by $\|f\|_p$. □

The proof of Theorem 4.32 for a particular range of indices can now be given. Suppose that we are seeking to bound the paraproduct B from $L^{p_1} \otimes L^{p_2}$ into L^r where $1 < r < \infty$. Then, by (4.39) and the previous lemma, we have

$$\langle \mathsf{B}(f_1, f_2), f_3 \rangle \le \prod_{k=1}^{3} \|\mathsf{T}_k f_k\|_{p_k}$$

where by abuse of notation we take p_3 to be the conjugate index to r.

The remainder of the theory that we develop is needed to address the case in which the paraproduct does not obey a duality estimate.

4.3. A Localization Lemma. We will need an estimate which refines the L^2 estimates for the operators T proved in Proposition 4.47.

We will make a further definition in which these operators, defined as as mixture of sums and supremums over rectangles, are restricted to a subset of rectangles. Thus, if $\mathcal{O} \subset \mathcal{R}$, and $\mathsf{T} = \mathsf{SSS}$, for instance in the three parameter case, we set

$$\mathsf{T}_{\mathcal{O}} f = \left[\sum_{R \in \mathcal{O}} \frac{|\langle f, \varphi_R \rangle|^2}{|R|} \mathbf{1}_R \right]^{1/2}.$$

Of course here we insist that the function φ_R have zeros in each coordinate. More generally, to define $\mathsf{T}_{\mathcal{O}}$, in the expression (4.38), we restrict the rectangles to be in the collection $\mathcal{O}$ rather than all possible rectangles.

4.48. Lemma. *Suppose that $\mathcal{O} \subset \mathcal{R}$ and that there is a constant $\mu > 1$ so that for a function f,*

(4.49) $\quad \operatorname{supp}(f) \cap \mu R = \emptyset, \quad R \in \mathcal{O}.$

Then it is the case that we have

(4.50) $\quad \|\mathsf{T}_{\mathcal{O}} f\|_2 \lesssim \mu^{-N'} \|f\|_2.$

The exponent N' is a function of only the exponent N in the definition of adapted, (2.3).

Proof. This lemma is a corollary to the proof of L^2 boundedness of the operators T_j, and to deduce it, we will rely upon a degree of flexibility built into the definition of adapted.

If we knew that the functions φ_R were supported on, e.g., $\frac{\mu}{2} R$, then the conclusion of the Lemma would be obvious. Thus, the problem at hand is one of Schwartz tails.

We make a further specification of the definition of adapted, (2.3), which is applied to functions on the real line. Fix a constant K and an integer N say that φ is (K, N)-*adapted* to an interval I iff

$$|\mathsf{D}^n\varphi(x)| \le K|I|^{-n-1/2}\left(\frac{1}{2}+\frac{|x-c(I)|}{|I|}\right)^{-N}, \quad n=0,1.$$

In addition, the L^2 norm of φ is at most one.

Say that φ is (K,N)-*adapted* to a rectangle $R=\bigotimes_{j=1}^{d} R_{(j)} \subset \mathbb{R}^d$ iff φ is a product

$$\varphi(x_1,\ldots,x_d)=\prod_{j=1}^{d}\varphi_j(x_j)$$

with each φ_j (K,N)-*adapted* to $R_{(j)}$. This definition naturally extends to collections of rectangles.

Now, fix K_0 and N_0 so that Proposition 4.47 holds for all functions $\{\varphi_R : R \in \mathscr{R}\}$ which are (K_0,N_0)-adapted to $\mathscr{R}$ (and have zeros in the right coordinates). Here, we take $K_0 \ge 1$ so that there will be no difficulties with having the functions φ_R be of L^2 norm one.

For N' as in the conclusion of the Lemma, set $N_1 = N_0 + N'$. Consider functions $\{\varphi_R : R \in \mathscr{R}\}$ that are (K_0,N_1)-adapted to $\mathscr{R}$. And let T be the corresponding operator constructed from these functions.

Observe that we can define a new set of functions $\{\tilde\varphi_R : R \in \mathscr{R}\}$ that are $(2K_0,N_0)$ adapted to $\mathscr{R}$, and satisfy

$$\tilde\varphi_R(x)=\mu^{N'}\varphi_R(x), \quad x \notin \mu R.$$

In those coordinates $1 \le j \le d$ where there is no zero, this is accomplished by multiplying by a smooth function that is zero on a large neighborhood of R, and identically $\mu^{N'/d}$ in $\mathbb{R}-\mu R_{(j)}$. If the coordinate has a zero, observe that

$$\left|\int_{\mathbb{R}-\mu R_{(j)}} \varphi_{R_{(j)}}(x_j)\,dx_j\right| \le 2K_0|R_{(j)}|^{1/2}(\mu)^{-N_1+1}$$

provided $N_1 > 2$. And so we can set $\tilde\varphi_{R_{(j)}}$ in a neighborhood of R to cancel out this integral.

The operator $\tilde{\mathsf{T}}$ constructed from the functions $\{\tilde\varphi_R : R \in \mathscr{R}\}$ will satisfy an L^2 bound that is independent of μ. For a function f as in (4.49), we have

$$\mu^{N'}\mathsf{T}_{\mathcal{O}}f=\tilde{\mathsf{T}}_{\mathcal{O}}f.$$

And the right hand side admits an L^2 bound independent of μ and N', so the proof is complete. □

We will also need the following corollary to the previous Lemma.

4.51. Corollary. *Let $\mathcal{O}$ be a collection of rectangles whose shadow has finite measure. If f is a bounded function, we have the estimate*

$$(4.52)\quad \|\mathsf{T}_{\mathcal{O}} f\|_2 \lesssim |\mathrm{sh}(\mathcal{O})|^{1/2} \|f\|_\infty .$$

Proof. Let $f \in L^\infty$ be bounded by one, set $U = \mathrm{sh}(\mathcal{O})$, and define f_k, for $k \geq 0$ as in (4.45). We shall see that

$$(4.53)\quad \sum_k \|\mathsf{T}_{\mathcal{O}} f_k\|_2 \lesssim |\mathrm{sh}(\mathcal{O})|^{1/2}.$$

Indeed, applying Lemma 4.48, we see that

$$\|\mathsf{T}_{\mathcal{O}} f_k\|_2 \lesssim 2^{-N'k} |\mathrm{sh}(\mathcal{O})|^{1/2}$$

where we can assume that $N' > 4$ say. □

4.4. The Proof of Theorem 4.32. We only treat the bilinear case of the theorem, as the higher order linearities are easy to accommodate into this proof. We also restrict our attention to the two parameter setting. Straightforward modifications adapt the argument to an arbitrary number of parameters. The first cases that we consider are those in which B is to be mapped into a space L^r with $\frac{1}{2} \leq r \leq 1$.

Some of the generalities of the proof of the one dimensional case remain in force in the current setting, in particular, it will suffice for us to establish (2.25). That is, we shall demonstrate this: For all $f_j \in L^{p_j}$ of norm 1 and set $E_3 \subset \mathbb{R}^2$ of measure one, there is an open subset $E_3' \subset E_3$ of measure at least $\frac{1}{2}$, so that for f_3 a smooth function compactly supported in E_3' and with L^∞ norm at most one, we have

$$(4.54)\quad \sum_R |R|^{-1/2} \prod_{j=1}^{3} |\langle f_j, \varphi_{j,R} \rangle| \lesssim 1.$$

Moreover, it suffices to take f_j in a dense class of functions, and so we take f_1 and f_2 to be smooth and compactly supported.

Observe that as all f_j smooth and compactly supported, the sum above is at most 1 if the sum is restricted to rectangles that have at least one side length either small or large, as in these cases the inner products above decay rapidly.[5] Thus, we can assume that the sum is restricted to a finite number of rectangles, and we should provide an estimate for the sum that is independent of the exact number or nature of the rectangles.

This last sum is over positive summands. It will be useful to us to organize the sum over appropriate subcollections of $\mathcal{R}$. For a collection of dyadic rectangles $\mathcal{O}$ set

[5] With the precise definition of small and large depending upon the functions f_j.

$$\text{(4.55)}\quad \operatorname{Sum}(\mathcal{O}) \stackrel{\text{def}}{=} \sum_{R \in \mathcal{O}} |R| \prod_{j=1}^{3} \frac{|\langle f_j, \varphi_{j,R} \rangle|}{\sqrt{|R|}}$$

where the functions $\varphi_{j,R}$ are associated with the bilinear paraproduct that we considering.

We will be working with different collections of rectangles $\mathcal{O}$. The *shadow* of $\mathcal{O}$ is defined to be

$$\operatorname{sh}(\mathcal{O}) \stackrel{\text{def}}{=} \bigcup_{R \in \mathcal{O}} R.$$

The Definition of E_3'. Let T_j, for $j = 1, 2, 3$, be the three operators as in (4.39). (Though at this point the function f_3 is not yet specified.) For the sake of symmetry, set T_0 to be the maximal operator in 2 parameters.

Define $4\nu \stackrel{\text{def}}{=} \min(p_1, p_2)$, and set

$$\text{(4.56)}\quad \Omega_{j,l} \stackrel{\text{def}}{=} \{\mathsf{T}_j f_j > \kappa 2^l\}, \quad l \in \mathbb{Z},\ j = 1, 2$$

$$\text{(4.57)}\quad \Omega_l \stackrel{\text{def}}{=} \bigcup_{j=1}^{2} \Omega_{j,l},$$

$$\text{(4.58)}\quad \Omega \stackrel{\text{def}}{=} \bigcup_{l \in \mathbb{N}} \left\{ \mathsf{T}_0 \mathbf{1}_{\Omega_l} > \frac{1}{100} 2^{-\nu l} \right\},$$

$$\text{(4.59)}\quad \tilde{\Omega} \stackrel{\text{def}}{=} \left\{ \mathsf{T}_0 \mathbf{1}_{\Omega} > \frac{1}{2} \right\}.$$

In these definitions, we fix a value of $\kappa \simeq 1$ so that $|\tilde{\Omega}| < \frac{1}{2}$, and then take $E_3' = E_3 \cap \tilde{\Omega}^c$, so that the measure of this set is at least $\frac{1}{2}$. This is possible, since we can estimate, using the L^2 bound for the maximal function and the L^{p_j} bounds for the T_j,

$$|\Omega| \le \sum_{l \in \mathbb{N}} |\{\mathsf{T}_0 \mathbf{1}_{\Omega_l} > 2^{-\nu l}\}| \le K_1 \sum_{l \in \mathbb{N}} 2^{2\nu l} |\Omega_l| \le K_2 \sum_{l \in \mathbb{N}} \sum_{j=1}^{2} \kappa^{-p_j} 2^{(2\nu - p_j)l}.$$

And the last sum is less than $\frac{1}{8}$ for a fixed $\kappa \simeq 1$.

The Decomposition of $\mathcal{R}$. We decompose the collection of all rectangles. A rectangle R is in $\mathcal{O}_{j,l}$ iff l is the greatest integer such that

$$\text{(4.60)}\quad |R \cap \Omega_{j,l}| = |R \cap \{\mathsf{T}_j f_j > \kappa 2^l\}| \ge \frac{1}{100} |R|.$$

Here, we are extending the definition in (4.56) to $j = 1, 2, 3$ and to $l \in \mathbb{Z}$.

Observe that as each f_j is smooth, it is necessarily the case that $\mathsf{T}_j f_j$ is a bounded function. Thus, the definition above makes sense, and for each j, every rectangle $R \in \mathscr{R}$ will be a member of some $\mathcal{O}_{j,l}$, for $l \in \mathbb{Z}$.

For integers $\vec{l} = (l_1, l_2, l_3) \in \mathbb{O} \stackrel{\text{def}}{=} (-\mathbb{N}) \times (-\mathbb{N}) \times \mathbb{Z}$, we define

$$(4.61) \quad \mathcal{O}_{\vec{l}} \stackrel{\text{def}}{=} \bigcap_{j=1}^{3} \mathcal{O}_{j,l_j}.$$

This is not a complete decomposition of the collection of rectangles, a point we return to below.

We appeal to the principal technical estimate, proved below. Observe that, by Lemma 4.70 and using the notation of (4.55), it is the case that

$$(4.62) \quad \mathrm{Sum}(\mathcal{O}_{\vec{l}}) \lesssim 2^{l_1+l_2+l_3} |\mathrm{sh}(\mathcal{O}_{\vec{l}})|.$$

We therefore need effective estimates for the shadow. Since each T_j operator is bounded on all L^p spaces, we of course have the estimate

$$(4.63) \quad |\mathrm{sh}(\mathcal{O}_{\vec{l}})| \lesssim \min_j 2^{-p_j l_j} \lesssim 2^{-\theta_1 p_1 l_1 - \theta_2 p_2 l_2 - \theta_3 p_3 l_3}, \quad \vec{l} \in \mathbb{O}.$$

Here, $l_1, l_2 < 0$ while $l_3 \in \mathbb{Z}$, and $\theta_j \geq 0$ with $\theta_1 + \theta_2 + \theta_3 = 1$. Recall that p_1 and p_2 are specified to us in advance, but as f_3 is a bounded function on a set of finite measure, we are free to take any value of $1 < p_3 < \infty$ that we wish. In particular, it is effective to take p_3 to be relatively close to one for $l_3 > 0$ while we take p_3 large for $l_3 \leq 0$.

The sums are treated separately based on the sign of the last coordinate of $\vec{l} \in \mathbb{O}$. Combining (4.62) and (4.63), we see that

$$(4.64) \quad \sum_{\vec{l} \in (-\mathbb{N})^3} \mathrm{Sum}(\mathcal{O}_{\vec{l}}) \lesssim \sum_{\vec{l} \in (-\mathbb{N})^3} 2^{l_1(1-p_1\theta_1) + l_2(1-p_2\theta_2) + l_3(1-p_3\theta_3)}.$$

We should choose $0 < \theta_1 < \frac{1}{p_1}$, and $0 < \theta_2 < \frac{1}{p_2}$ so that $\theta_1 + \theta_2 < 1$. We are then still free to chose $p_3 > 1$, but close enough to one so that $p_3\theta_3 = p_3(1 - \theta_1 - \theta_2) < 1$. Thus, this last sum is no more than a constant.

Let us consider the case of $l_1, l_2 \leq 0$ while $l_3 > 0$. The minimum in (4.63) occurs for $j = 3$, and we have the estimate

$$(4.65) \quad \sum_{\vec{l} \in (-\mathbb{N})^2 \otimes \mathbb{N}} \mathrm{Sum}(\mathcal{O}_{\vec{l}}) \lesssim 2^{l_1 + l_2 - l_3(p_3 - 1)}.$$

This clearly sums to a constant for p_3 sufficiently large.

Some rectangles are not in the classes defined above. To treat the remaining cases, set

$$\begin{aligned} &\mathscr{P}_{\vec{l}} \stackrel{\text{def}}{=} \mathcal{O}_{1,l_1} \cap \mathcal{O}_{2,l_2}, \\ &\vec{l} = (l_1, l_2) \in \mathbb{P} \stackrel{\text{def}}{=} \{\mathbb{Z} - (-\mathbb{N})\} \times \{\mathbb{Z} - (-\mathbb{N})\}. \end{aligned} \tag{4.66}$$

This decomposition does not take the role of f_3 into account, and so our next steps are to deduce information about this function.

Suppose that $R \in \mathscr{P}_{\vec{l}}$. Then either l_1 or l_2 must be positive. Suppose that l_1 is. By (4.58), it is then the case that, by definition, $R \subset \Omega$, but moreover

$$2^{v' l_1} R \cap E_3' = \emptyset$$

where $2v' = v$.

Now, the function f_3 satisfies the conditions of (4.49), with $\mathcal{O} = \mathscr{P}_{\vec{l}}$ and $\mu = 2^{v' l_1 / 2}$. Then by Lemma 4.48, we have

$$\|\mathsf{T}_{3, \mathscr{P}_{\vec{l}}} f_3\|_2 \lesssim 2^{-vN' l_1/2} \lesssim 2^{-10 l_1} \tag{4.67}$$

for appropriate choice of N in (2.3).

Note that we have proved the inequality

$$\|\mathsf{T}_{3, \mathscr{P}_{\vec{l}}} f_3\|_2 \lesssim \min(2^{-10 l_1}, 2^{-10 l_2}).$$

We can then apply (4.73) to see that

$$\mathrm{Sum}(\mathscr{P}_{\vec{l}}) \lesssim 2^{l_1 + l_2} \min(2^{-10 l_1}, 2^{-10 l_2}).$$

This is clearly summable to a constant over the indices $\mathbb{P}$, as either l_1 or l_2 must be positive.

This proof will permit e.g. $p_1 = 1$, with the additional hypothesis that $f_1 \in H^1$, and that all functions $\varphi_{1,R}$ satisfy (4.33) for all coordinates $1 \le j \le d$. By duality, this implies the *BMO* estimate of our Theorem.

The endpoint estimates. The endpoint estimates concern the case when, say, $p_2 = \infty$, which is a case not handled in the discussion above. (Note that assuming that $f_2 \in L^\infty$, we do not need to make additional assumptions about the zeros of the functions $\varphi_{2,R}$.)

We again prove (4.54). And the method of proof is quite close to the argument above. Use the same notation as in (4.56), but now define

(4.68) $\Omega \stackrel{\text{def}}{=} \bigcup_{l \in \mathbb{N}} \left\{ \mathsf{T}_0 \mathbf{1}_{\Omega_{1,l}} > \frac{1}{100} 2^{-\nu l} \right\},$

(4.69) $\tilde{\Omega} \stackrel{\text{def}}{=} \left\{ \mathsf{T}_0 \mathbf{1}_{\Omega} > \frac{1}{2} \right\}.$

We take $E_3' = E_3 \cap \Omega$, so that again we have $|E_3'| \geq \frac{1}{2}$.

We define the sets $\mathcal{O}_{j,l}$ as in (4.60) but we shall only use this for $j = 1, 3$. Set (in contrast to (4.61)),

$$\mathcal{O}_{\vec{l}} = \mathcal{O}_{1,l_1} \cap \mathcal{O}_{3,l_3}, \quad \vec{l} = (l_1, l_3).$$

We then have the estimate below, as a consequence of Corollary 4.51 and (4.73),

$$\mathrm{Sum}(\mathcal{O}_{\vec{l}}) \leq 2^{-l_1-l_3} |\mathrm{sh}(\mathcal{O}_{\vec{l}})|^{1/2} \|\mathsf{T}_{\mathcal{O}_{\vec{l}}} f_2\|_2 \lesssim 2^{l_1+l_3} |\mathrm{sh}(\mathcal{O}_{\vec{l}})|.$$

We estimate the shadow

$$|\mathrm{sh}(\mathcal{O}_{\vec{l}})| \lesssim \min(2^{-p_1 l_1}, 2^{-p_3 l_3}).$$

This follows on the one hand from the assumption that $f_1 \in L^{p_1}$. But recall that we can choose $1 < p_3 < \infty$ in an arbitrary fashion, as f_3 is bounded by one and supported on a set of measure at most one.

Pulling these estimate together, we see that

$$\mathrm{Sum}(\mathcal{O}_{\vec{l}}) \lesssim 2^{l_1(1-\theta_1 p_1) + l_3(1-\theta_3 p_3)},$$

where the θ_1, θ_3 are non-negative and sum to one. The index p_1 is specified to us, but p_3 can be taken arbitrarily. For $\vec{l} \in (-\mathbb{N}) \times \mathbb{Z}$, we should take p_3 close to one for $l_3 \leq 0$, but $p_3 = 4$, say, for $l_3 > 0$. Doing so we see that

$$\sum_{\vec{l} \in (-\mathbb{N}) \times \mathbb{Z}} \mathrm{Sum}(\mathcal{O}_{\vec{l}}) \lesssim 1.$$

We turn to the case where $l_1 > 0$. As before, we should now gain additional information about the function f_3. But the reasoning of the previous section, and in particular (4.50) and (4.67), leads us immediately to

$$\|\mathsf{T}_{3,\mathcal{O}_{1,l_1}} f_3\|_2 \lesssim 2^{-10 p_1 l_1}.$$

On the other hand, Corollary 4.51 implies that

$$\|\mathsf{T}_{2,\mathcal{O}_{1,l_1}} f_2\|_2 \lesssim |\mathrm{sh}(\mathcal{O}_{1,l})|^{1/2} \lesssim 2^{-(1/2) p_1 l_1}.$$

Appealing to (4.74), we see that

$$\mathrm{Sum}(\mathcal{O}_{l_1}) \lesssim 2^{l_1(1-10-p_1/2)},$$

which is clearly summable over $l_1 > 0$.

The principal technical estimate. In this section we isolate the principal technical estimate in proof of Theorem 4.32.

4.70. Lemma. *Suppose that for three constants* $0 < \lambda_j < \infty$, $j = 1, 2, 3$ *and a collection of rectangles* $\mathcal{O}$ *we have*

$$(4.71) \quad |R \cap \{\mathsf{T}_j f_j > \lambda_j\}| \le \frac{1}{100}|R|, \quad R \in \mathcal{O}, \ j = 1, 2, 3.$$

Then we have the estimate

$$(4.72) \quad \mathrm{Sum}(\mathcal{O}) \lesssim |\mathrm{sh}(\mathcal{O})| \prod_{j=1}^{3} \lambda_j.$$

Suppose that (4.71) *does not hold for* $j = 3$. *Then we have the estimate*

$$(4.73) \quad \mathrm{Sum}(\mathcal{O}) \lesssim \lambda_1 \lambda_2 |\mathrm{sh}(\mathcal{O})|^{1/2} \|\mathsf{T}_{3,\mathcal{O}} f_3\|_2.$$

Suppose that (4.71) *does not hold for* $j = 2$ *and* $j = 3$. *Then we have the estimate*

$$(4.74) \quad \mathrm{Sum}(\mathcal{O}) \lesssim \lambda_1 \|\mathsf{T}_{2,\mathcal{O}} f_2\|_2 \|\mathsf{T}_{3,\mathcal{O}} f_3\|_2.$$

We will apply this in settings in which we have a good estimate for the shadow of $\mathcal{O}$ in terms of the λ_j.

Proof. Set

$$W = \mathrm{sh}(\mathcal{O}) \cap \bigcap_{j=1}^{3} \{\mathsf{T}_j f_j < \lambda_j\}.$$

Then $R \cap W$ has measure at least $\frac{97}{100}|R|$. This permits us to restrict the range of integration below to W.

$$\mathrm{Sum}(\mathcal{O}) \lesssim \int_W \sum_{R \in \mathcal{O}} \prod_{j=1}^{3} \frac{|\langle f_j, \varphi_{j,R} \rangle|}{\sqrt{|R|}} \mathbf{1}_R \, dx \lesssim \int_W \prod_{j=1}^{3} \mathsf{T}_j f_j \, dx \lesssim |\mathrm{sh}(\mathcal{O})| \prod_{j=1}^{3} \lambda_j.$$

This proves our first conclusion. The remaining conclusions follow from the same reasoning, with the use of the Cauchy-Schwarz inequality. □

References

[1] Chang Sun-Yung A. and Fefferman Robert: Some recent developments in Fourier analysis and H^p-theory on product domains. Bull. Amer. Math. Soc. (N.S.) **12** (1985), 1–43

[2] ———: A continuous version of duality of H^1 with BMO on the bidisc. Ann. of Math. (2) **112** (1980), 179–201

[3] Coifman R. and Meyer Y.: Commutateurs d'intégrales singulières et opérateurs multilinéaires. Ann. Inst. Fourier (Grenoble) **28** (1978), 177–202 (French, with English summary)

[4] Coifman Ronald R. and Meyer Yves: Au delà des opérateurs pseudo-différentiels. Astérisque 57. Société Mathématique de France, Paris, 1978 (French)

[5] Fefferman Robert: Harmonic analysis on product spaces. Ann. of Math. (2) **126** (1987), 109–130

[6] Fefferman C. and Stein E. M.: Some maximal inequalities. Amer. J. Math. **93** (1971), 107–115

[7] Ferguson Sarah H. and Lacey Michael T.: A characterization of product BMO by commutators. Acta Math. **189** (2002), 143–160

[8] Journé Jean-Lin: Calderón-Zygmund operators on product spaces. Rev. Mat. Iberoamericana **1** (1985), 55–91

[9] Lacey Michael T. and Terwilleger Erin: Little Hankel Operators and Product BMO (2004)

[10] Mucalu Camil, Pipher Jill, Tao Terrance, and Thiele Christoph: Bi-parameter paraproducts. Acta Math. **193** (2004), 269–296

[11] ———: Multi-parameter paraproducts. arxiv:math.CA/0411607

[12] Thiele Christoph: Wave Packets, Multilinear Operators and Carleson Theorems. CBMS Lectures, May 2004

Received September 12, 2005

Michael Lacey, School of Mathematics, Georgia Institute of Technology, Atlanta, GA 30332 USA
lacey@math.gatech.edu

Jason Metcalfe, Department of Mathematics, University of California, Berkeley, CA 94720-3840 USA
metcalfe@math.berkeley.edu