A SHORT PROOF OF THE COIFMAN-MEYER MULTILINEAR THEOREM

CAMIL MUSCALU, JILL PIPHER, TERENCE TAO, AND CHRISTOPH THIELE

Abstract. We give a short proof of the well known Coifman-Meyer theorem on multilinear operators.

1. Introduction

The main task of the present paper is to present a new proof of the classical Coifman-Meyer theorem on multilinear singular integrals, see [2], [3], [5], [6].

Let $m \in L^\infty(\mathbb{R}^n)$ be a bounded function which is smooth away from the origin and satisfies the following Marcinkiewicz-Mihlin-Hörmander type condition

$$|\partial^\alpha m(\xi)| \lesssim \frac{1}{|\xi|^{|\alpha|}}, \tag{1}$$

for sufficiently many multiindices α [1]. For $f_1, ..., f_n \in \mathcal{S}(\mathbb{R})$ Schwartz functions on the real line, we define the n-linear operator T_m by the formula

$$T_m(f_1, ..., f_n)(x) = \int_{\mathbb{R}^n} m(\xi) \widehat{f_1}(\xi_1) ... \widehat{f_n}(\xi_n) e^{2\pi i x(\xi_1 + ... + \xi_n)} \, d\xi_1 ... d\xi_n. \tag{2}$$

The following theorem holds, see [2], [3], [5], [6].

Theorem 1.1. *As defined, the multilinear operator T_m maps $L^{p_1} \times ... \times L^{p_n} \to L^p$ as long as $1 < p_i \leq \infty$, $1 \leq i \leq n$, $1/p_1 + ... + 1/p_n = 1/p$ and $0 < p < \infty$.*

When such an $n+1$-tuple $(p_1, ..., p_n, p)$, has the property that $0 < p < 1$ and $p_j = \infty$ for some $1 \leq j \leq n$ then, for some technical reasons (see [5], [7]), by L^∞ one actually means L^∞_c the space of bounded measurable functions with compact support.

The case $p \geq 1$ has been proven in [3] while the general case $p > 1/n$ has been independently settled in [6] and [5]. The interesting fact that p can be a number smaller than 1 goes back to [2]. The usual argument to prove the theorem (see [2], [3], [5], [6]) uses the celebrated $T1$ theorem of David and Journe [4] and relies on BMO theory, Carleson measures and C.Fefferman's duality theorem between the Hardy space H^1 and BMO.

As the reader will see, our proof is conceptually simpler and does not use any of the aforementioned ingredients. It is based on a careful stopping time argument involving the Hardy-Littlewood maximal function and the Littlewood-Paley square function.

[1] Throughout the paper we will write $A \lesssim B$ iff there is a universal constant $C > 0$ so that $A \leq CB$.

2. Model operators

For simplicity we treat the $n = 2$ case only. It is a standard fact by now (see for instance the papers [7], [8]) that the study of our bilinear operators T_m can be reduced to the study of finitely many discrete model operators $\Pi^j_{\mathbf{P}}$, $j = 0, 1, 2, 3$ of the form

$$\Pi^j_{\mathbf{P}}(f_1, f_2) = \sum_{P\in\mathbf{P}} \epsilon_P \frac{1}{|I_P|^{1/2}} \langle f_1, \phi_{P_1}\rangle\langle f_2, \phi_{P_2}\rangle\phi_{P_3}. \tag{3}$$

Here $\mathbf{P}$ is a collection of lacunary dyadic tiles corresponding to lattice points (k, n) in $\mathbf{Z}^2$ and $(\epsilon_P)_P$ is an arbitrary sequence of uniformly bounded constants. More precisely, P_i $i = 1, 2, 3$ are defined by $P_i = I_P \times \omega_{P_i}$ where $I_P = [n2^{-k}, (n + 1)2^{-k+1}]$, $\omega_{P_i} = [0, 2^k]$ if $i = j$ and $\omega_{P_i} = [2^k, 2^{k+1}]$ if $i \neq j$, while ϕ_{P_i} $i = 1, 2, 3$ are L^2 normalized wave packets corresponding to the Heisenberg boxes P_i $i = 1, 2, 3$. This means that the function ϕ_{P_i} is a smooth L^2 normalized bump function adapted to the interval I_P whose Fourier transform $\widehat{\phi_{P_i}}$ is supported inside the interval ω_{P_i} for $i = 1, 2, 3$. We should emphasize here that the tile P is uniquely determined by the interval I_P.

To explain this reduction in a few words, let $Q := I \times J$ be a dyadic rectangle in the plane, having the property that $\mathrm{diam}(Q) \sim \mathrm{dist}(Q, \{0\})$. Let also ϕ_I, ϕ_J be two L^1 normalized smooth functions such that $\mathrm{supp}(\widehat{\phi_I}) \subseteq I$ and $\mathrm{supp}(\widehat{\phi_J}) \subseteq J$. If we replace the symbol $m(\xi_1, \xi_2)$ by $\widehat{\phi_I}(\xi_1)\widehat{\phi_J}(\xi_2)$ we observe that the right hand side of (2) becomes $(f_1 * \phi_I)(x)(f_2 * \phi_J)(x)$. On the other hand, inequality (1) implies that one can think of m as being essentially constant on each such Q and so the integral in (2) when smoothly restricted to this cube, becomes roughly

$$\epsilon_Q(f_1 * \phi_I)(x)(f_2 * \phi_J)(x).$$

Then, one covers the plane by a collection of carefully selected such Whitney cubes and discretize again in the x-variable. In the end, one obtains a formula in which the general T_m is written as an average of operators of the type of the model operators above. The details can be found in [7], [8]. Now our analysis of the operators $\Pi^j_{\mathbf{P}}$ will be independent on $j = 0, 1, 2, 3$ and so we can assume without loss of generality that $j = 1$ and write from now on, for simplicity, $\Pi_{\mathbf{P}}$ instead of $\Pi^1_{\mathbf{P}}$ (the reader will observe that for $j \neq 1$, the only difference is that the roles of the Hardy Littlewood maximal function M and the Littlewood Paley square function S, get permuted). It is therefore enough to prove the theorem for the bilinear operator $\Pi_{\mathbf{P}}$.

3. The proof

First, let us observe that it is very easy to obtain the necessary L^p estimates in the particular case when all the indices are strictly between 1 and ∞. To see this, let $f \in L^p$, $g \in L^q$, $h \in L^r$ for $1 < p, q, r < \infty$ with $1/p + 1/q + 1/r = 1$. Then,

$$\left|\int_{\mathbb{R}} \Pi_{\mathbf{P}}(f, g)(x)h(x)\, dx\right| \lesssim \sum_{P\in\mathbf{P}} \frac{1}{|I_P|^{1/2}} |\langle f, \phi_{P_1}\rangle||\langle g, \phi_{P_2}\rangle||\langle h, \phi_{P_3}\rangle| =$$

$$\int_{\mathbb{R}} \sum_{P\in\mathbf{P}} \frac{|\langle f, \phi_{P_1}\rangle|}{|I_P|^{1/2}} \frac{|\langle g, \phi_{P_2}\rangle|}{|I_P|^{1/2}} \frac{|\langle h, \phi_{P_3}\rangle|}{|I_P|^{1/2}} \chi_{I_P}(x)\, dx \lesssim \tag{4}$$

$$\int_{\mathbb{R}} \left(\sup_{P\in\mathbf{P}} \frac{|\langle f, \phi_{P_1}\rangle|}{|I_P|^{1/2}} \chi_{I_P}(x)\right) \left(\sum_{P\in\mathbf{P}} \frac{|\langle g, \phi_{P_2}\rangle|^2}{|I_P|} \chi_{I_P}(x)\right)^{1/2} \left(\sum_{P\in\mathbf{P}} \frac{|\langle h, \phi_{P_3}\rangle|^2}{|I_P|} \chi_{I_P}(x)\right)^{1/2} dx \lesssim$$

$$\int_{\mathbb{R}} Mf(x) Sg(x) Sh(x)\, dx \lesssim \|Mf\|_p \|Sg\|_q \|Sh\|_r \lesssim$$

$$\|f\|_p \|g\|_q \|h\|_r,$$

where M is the maximal function of Hardy and Littlewood and S is the discrete square function of Littlewood and Paley, see [9] and [7]. This means that theorem 1.1 is nontrivial only when one index is ∞, or less or equal than 1. To prove the general case we just need to show that the bilinear operator $\Pi_{\mathbf{P}}$ maps $L^1 \times L^1 \to L^{1/2,\infty}$ because then, by interpolation and symmetry the theorem follows as in [7]. Let $f, g \in L^1$ be such that $\|f\|_1 = \|g\|_1 = 1$. We now recall Lemma 5.4 in [1].

Lemma 3.1. *Let $0 < p < \infty$ and $A > 0$. Then the following statements are equivalent up to constants:*

(i) $\|f\|_{p,\infty} \lesssim A$.

(ii) For every set E with $0 < |E| < \infty$, there exists a subset $E' \subseteq E$ with $|E'| \sim |E|$ and $|\langle f, \chi_{E'}\rangle| \lesssim A|E|^{1/p'}$. Here p' is defined by $1/p' + 1/p = 1$ (note that p' can be a negative number!).

Proof To see that (i) implies (ii), set

$$E' := E \setminus \{x : |f(x)| \geq CA|E|^{-1/p}\}.$$

If C is a sufficiently large constant, then (i) implies $|E'| \sim |E|$ and the claim follows.

To see that (ii) implies (i), let $\lambda > 0$ be arbitrary and set $E := \{x : Re(f(x)) > \lambda\}$. Then by (ii) we have

$$\lambda|E| \sim \lambda|E'| \lesssim A|E|^{1/p'},$$

and (i) easily follows (replacing Re by $-Re$, Im, $-Im$ as necessary). ■

Using this Lemma 3.1 in the particular case $p = 1/2$ and the scale invariance, it is enough to show that given $E \subseteq \mathbb{R}$ $|E| = 1$, one can find a subset $E' \subseteq E$ with $|E'| \sim 1$ such that

$$\sum_{P\in\mathbf{P}} \frac{1}{|I_P|^{1/2}} |\langle f, \Phi_{P_1}\rangle| |\langle g, \Phi_{P_2}\rangle| |\langle h, \Phi_{P_3}\rangle| \lesssim 1 \tag{5}$$

where $h := \chi_{E'}$. Fix such a set E with $|E| = 1$. To construct the subset E', we first consider

$$\Omega_0 = \{x \in \mathbb{R} : M(f)(x) > C\} \cup \{x \in \mathbb{R} : S(g)(x) > C\} \cup \{x \in \mathbb{R} : M(g)(x) > C\}.$$

Also, define

$$\Omega = \{x \in \mathbb{R} : M(1_{\Omega_0})(x) > \frac{1}{100}\}.$$

Clearly, we have $|\Omega| < 1/2$, if C is a big enough constant which we fix from now on. Then, we define $E' := E \setminus \Omega = E \cap \Omega^c$ and observe that indeed $|E'| \sim 1$. After this, we split our sum in (5) into two parts

$$\sum_{P \in \mathbf{P}} = \sum_{I_P \cap \Omega^c \neq \emptyset} + \sum_{I_P \cap \Omega^c = \emptyset} := I + II.$$

We also assume that the set $\mathbf{P}$ is finite, since our estimates do not depend on its cardinality.

First, we estimate term I. Since $I_P \cap \Omega^c \neq \emptyset$, it follows that $\frac{|I_P \cap \Omega_0|}{|I_P|} < \frac{1}{100}$ or equivalently, $|I_P \cap \Omega_0^c| > \frac{99}{100}|I_P|$.

We are now going to describe three decomposition procedures, one for each function f, g, h. Later on, we will combine them, in order to handle our sum. First, define

$$\Omega_1 = \{x \in \mathbb{R} : M(f)(x) > \frac{C}{2^1}\}$$

and set

$$\mathbf{T}_1 = \{P \in \mathbf{P} : |I_P \cap \Omega_1| > \frac{1}{100}|I_P|\},$$

then define

$$\Omega_2 = \{x \in \mathbb{R} : M(f)(x) > \frac{C}{2^2}\}$$

and set

$$\mathbf{T}_2 = \{P \in \mathbf{P} \setminus \mathbf{T}_1 : |I_P \cap \Omega_2| > \frac{1}{100}|I_P|\},$$

and so on. (The constant $C > 0$ is the one which we fixed before). Since there are finitely many tiles, this algorithm ends after a while, producing the sets $\{\Omega_n\}$ and $\{\mathbf{T}_n\}$ such that $\mathbf{P} = \cup_n \mathbf{T}_n$. Independently, define

$$\Omega'_1 = \{x \in \mathbb{R} : S(g)(x) > \frac{C}{2^1}\}$$

and set

$$\mathbf{T}'_1 = \{P \in \mathbf{P} : |I_P \cap \Omega'_1| > \frac{1}{100}|I_P|\},$$

then define

$$\Omega'_2 = \{x \in \mathbb{R} : S(g)(x) > \frac{C}{2^2}\}$$

and set

$$\mathbf{T}'_2 = \{P \in \mathbf{P} \setminus \mathbf{T}'_1 : |I_P \cap \Omega'_2| > \frac{1}{100}|I_P|\},$$

and so on, producing the sets $\{\Omega'_n\}$ and $\{\mathbf{T}'_n\}$ such that $\mathbf{P} = \cup_n \mathbf{T}'_n$. We would like to have such a decomposition available for the function h also. To do this, we first need to construct the analogue of the set Ω_0, for it. We will therefore pick $N > 0$ a big enough integer such that for every $P \in \mathbf{P}$ we have $|I_P \cap \Omega''^{c}_{-N}| > \frac{99}{100}|I_P|$ where we defined

$$\Omega''_{-N} = \{x \in \mathbb{R} : S(h)(x) > C2^N\}.$$

Then, similarly to the previous algorithms, we define

$$\Omega''_{-N+1} = \{x \in \mathbb{R} : S(h)(x) > \frac{C2^N}{2^1}\}$$

and set

$$\mathbf{T}''_{-N+1} = \{P \in \mathbf{P} : |I_P \cap \Omega''_{-N+1}| > \frac{1}{100}|I_P|\},$$

then define

$$\Omega''_{-N+2} = \{x \in \mathbb{R} : S(h)(x) > \frac{C2^N}{2^2}\}$$

and set

$$\mathbf{T}''_{-N+2} = \{P \in \mathbf{P} \setminus \mathbf{T}''_{-N+1} : |I_P \cap \Omega''_{-N+2}| > \frac{1}{100}|I_P|\},$$

and so on, constructing the sets $\{\Omega''_n\}$ and $\{\mathbf{T}''_n\}$ such that $\mathbf{P} = \cup_n \mathbf{T}''_n$. Then we write the term I as

$$\sum_{n_1,n_2>0,n_3>-N} \sum_{P \in \mathbf{T}_{n_1,n_2,n_3}} \frac{1}{|I_P|^{3/2}} |\langle f, \Phi_{P_1}\rangle||\langle g, \Phi_{P_2}\rangle||\langle h, \Phi_{P_3}\rangle||I_P|, \tag{6}$$

where $\mathbf{T}_{n_1,n_2,n_3} := \mathbf{T}_{n_1} \cap \mathbf{T}'_{n_2} \cap \mathbf{T}''_{n_3}$. Now, if P belongs to $\mathbf{T}_{n_1,n_2,n_3}$ this means in particular that P has not been selected at the previous $n_1 - 1$, $n_2 - 1$ and $n_3 - 1$ steps respectively, which means that $|I_P \cap \Omega_{n_1-1}| < \frac{1}{100}|I_P|$, $|I_P \cap \Omega'_{n_2-1}| < \frac{1}{100}|I_P|$ and $|I_P \cap \Omega''_{n_3-1}| < \frac{1}{100}|I_P|$ or equivalently, $|I_P \cap \Omega^c_{n_1-1}| > \frac{99}{100}|I_P|$, $|I_P \cap \Omega'^{c}_{n_2-1}| > \frac{99}{100}|I_P|$ and $|I_P \cap \Omega''^{c}_{n_3-1}| > \frac{99}{100}|I_P|$. But this implies that

$$|I_P \cap \Omega^c_{n_1-1} \cap \Omega'^{c}_{n_2-1} \cap \Omega''^{c}_{n_3-1}| > \frac{97}{100}|I_P|. \tag{7}$$

In particular, using (7), the term in (6) is smaller than

$$\sum_{n_1,n_2>0,n_3>-N} \sum_{P \in \mathbf{T}_{n_1,n_2,n_3}} \frac{1}{|I_P|^{3/2}} |\langle f, \Phi_{P_1}\rangle||\langle g, \Phi_{P_2}\rangle||\langle h, \Phi_{P_3}\rangle||I_P \cap \Omega^c_{n_1-1} \cap \Omega'^{c}_{n_2-1} \cap \Omega''^{c}_{n_3-1}| =$$

$$\sum_{n_1,n_2>0,n_3>-N} \int_{\Omega^c_{n_1-1} \cap \Omega'^{c}_{n_2-1} \cap \Omega''^{c}_{n_3-1}} \sum_{P \in \mathbf{T}_{n_1,n_2,n_3}} \frac{1}{|I_P|^{3/2}} |\langle f, \Phi_{P_1}\rangle||\langle g, \Phi_{P_2}\rangle||\langle h, \Phi_{P_3}\rangle|\chi_{I_P}(x)\, dx$$

$$\lesssim \sum_{n_1,n_2>0,n_3>-N} \int_{\Omega^c_{n_1-1} \cap \Omega'^{c}_{n_2-1} \cap \Omega''^{c}_{n_3-1} \cap \Omega_{\mathbf{T}_{n_1,n_2,n_3}}} M(f)(x) S(g)(x) S(h)(x)\, dx$$

$$\lesssim \sum_{n_1,n_2>0,n_3>-N} 2^{-n_1}2^{-n_2}2^{-n_3}|\Omega_{\mathbf{T}_{n_1,n_2,n_3}}|, \tag{8}$$

where

$$\Omega_{\mathbf{T}_{n_1,n_2,n_3}} := \bigcup_{P\in\mathbf{T}_{n_1,n_2,n_3}} I_P.$$

On the other hand we can write

$$|\Omega_{\mathbf{T}_{n_1,n_2,n_3}}| \leq |\Omega_{\mathbf{T}_{n_1}}| \leq |\{x\in\mathbb{R} : M(\chi_{\Omega_{n_1}})(x) > \frac{1}{100}\}|$$

$$\lesssim |\Omega_{n_1}| = |\{x\in\mathbb{R} : M(f)(x) > \frac{C}{2^{n_1}}\}| \lesssim 2^{n_1}.$$

Similarly, we have $|\Omega_{\mathbf{T}_{n_1,n_2,n_3}}| \lesssim 2^{n_2}$ and also $|\Omega_{\mathbf{T}_{n_1,n_2,n_3}}| \lesssim 2^{n_2\alpha}$, for every $\alpha \geq 1$, since $|E'| \sim 1$. In particular, it follows that

$$|\Omega_{\mathbf{T}_{n_1,n_2,n_3}}| \lesssim 2^{n_1\theta_1}2^{n_2\theta_2}2^{n_3\alpha\theta_3} \tag{9}$$

for any $0 \leq \theta_1, \theta_2, \theta_3 < 1$, such that $\theta_1+\theta_2+\theta_3 = 1$. Now we split the sum in (8) into

$$\sum_{n_1,n_2>0,n_3>0} 2^{-n_1}2^{-n_2}2^{-n_3}|\Omega_{\mathbf{T}_{n_1,n_2,n_3}}| + \sum_{n_1,n_2>0,0>n_3>-N} 2^{-n_1}2^{-n_2}2^{-n_3}|\Omega_{\mathbf{T}_{n_1,n_2,n_3}}|. \tag{10}$$

To estimate the first term in (10) we use the inequality (9) in the particular case $\theta_1 = \theta_2 = 1/2$, $\theta_3 = 0$, while to estimate the second term we use (9) for θ_j, $j = 1,2,3$ such that $\theta_1 < 1$, $\theta_2 < 1$ and $\alpha\theta_3 - 1 > 0$. With these choices, the geometric sums in (10) are finite. This ends the discussion on I.

Now term II is much simpler, being just an error term. We split

$$\mathbf{P} := \bigcup_{d>0} \mathbf{P}_d$$

where

$$\mathbf{P}_d := \{P\in\mathbf{P} : \frac{\mathrm{dist}(I_P,\Omega^c)}{|I_P|} \sim 2^d\}$$

and easily observe that

$$\sum_{P\in\mathbf{P}_d; I_P\subseteq\Omega} |I_P| \lesssim |\Omega| \sim 1. \tag{11}$$

Then, term II is smaller than

$$\sum_{d>0}\sum_{P\in\mathbf{P}_d;I_P\subseteq\Omega} |I_P|\left|\frac{\langle f,\Phi_{P_1}\rangle}{|I_P|^{1/2}}\right|\left|\frac{\langle g,\Phi_{P_2}\rangle}{|I_P|^{1/2}}\right|\left|\frac{\langle h,\Phi_{P_3}\rangle}{|I_P|^{1/2}}\right| \lesssim \sum_{d>0}\sum_{P\in\mathbf{P}_d;I_P\subseteq\Omega} |I_P|2^d2^d2^{-Kd} \lesssim 1,$$

for any big number $K > 0$, and this ends the proof.

References

[1] Auscher P., Hofmann S., Muscalu C., Tao T. and Thiele C. *Carleson measures, trees, extrapolation and Tb theorems*, Publ. Mat., vol. 46, 257-325, [2002].

[2] Coifman R. and Meyer Y., *On commutators of singular integrals and bilinear singular integrals*, Trans. Amer. Math. Soc. vol. 212, 315-331, [1975].

[3] Coifman R. and Meyer Y., *Ondelettes et operateurs III. Operateurs multilineaires*, Hermann, Paris, [1991].

[4] David G. and Journe J-L., A boundedness criterion for generalized Calderon-Zygmund operators, Ann. of Math., vol. 120, 371-397, [1984].

[5] Grafakos L. and Torres R., *Multilinear Calderon-Zygmund theory*, Adv. Math., vol. 165, 124-164, [2002].

[6] Kenig C. and Stein E., *Multilinear estimates and fractional integration*, Math. Res. Lett., vol. 6, 1-15, [1999].

[7] Muscalu C., Tao T. and Thiele C., *Multilinear operators given by singular multipliers*, J. Amer. Math. Soc. 15, [2002], 469-496.

[8] Muscalu C., Pipher J., Tao T. and Thiele, C., *Bi-parameter paraproducts*, Preprint, [2003].

[9] Stein, E., *Harmonic Analysis: Real Variable Methods, Orthogonality, and Oscillatory Integrals.* Princeton University Press, Princeton, [1993].

Department of Mathematics, Cornell University, Ithaca, NY 14853
E-mail address: camil@math.cornell.edu

Department of Mathematics, Brown University, Providence, RI 02912
E-mail address: jpipher@math.brown.edu

Department of Mathematics, UCLA, Los Angeles, CA 90095
E-mail address: tao@math.ucla.edu

Department of Mathematics, UCLA, Los Angeles, CA 90095
E-mail address: thiele@math.ucla.edu