Séminaire BOURBAKI Juin 2006
58e année, 2005-2006, n$^{o}$ 962, p. 295 à 320

# MODULATION INVARIANT AND MULTILINEAR SINGULAR INTEGRAL OPERATORS [after Lacey and Thiele]

by **Michael CHRIST**

### Contents

1. Historical background .......... 296
2. Localized Fourier coefficients .......... 299
3. Almost-orthogonality .......... 303
4. Trees, towers, and multitrees .......... 304
5. Organizing the total sum into subsums .......... 308
6. Counting multitrees .......... 310
7. Refinement .......... 311
8. Proof of the almost-orthogonality lemma .......... 314
9. Carleson's maximal operator, according to Lacey and Thiele .......... 314
10. Open problems .......... 316
References .......... 318

## INTRODUCTION

The bilinear Hilbert transform is the operator

$$BH(f_1, f_2)(x) = \int_{\mathbb{R}} f_1(x-t) f_2(x+t)\, t^{-1} dt \tag{1}$$

where $x, t \in \mathbb{R}$ and $f_j \in L^{p_j}(\mathbb{R})$. If $t^{-1}$ were an integrable function then this integral would become absolutely convergent, for almost every $x$ for appropriate exponents $p_j$. The question of the finiteness of the conditional integral, and of inequalities in $L^p$ norms, was an open problem from roughly the mid-1960's to the late 1990's, when Michael Lacey and Christoph Thiele showed in a series of breakthrough papers that $BH$ is well-defined and bounded on appropriate $L^p$ spaces. This operator is prototypical for a class of multilinear operators with modulation symmetry, and their work has

Supported in part by NSF grant DMS-040126.

been followed by significant further developments too numerous to cite in the space available.

In this expository article I discuss the background and origins of the problem, outline the main lines of the analysis, and indicate the connection with the almost everywhere convergence of Fourier integrals. This article is not intended as an exhaustive survey, but merely as an introduction to the main ideas of the original articles [21–23]. I deliberately focus on one particular operator in order to emphasize what I consider to be the main concepts.

## 1. HISTORICAL BACKGROUND

### 1.1. Singular integrals

The most fundamental example of a Calderón-Zygmund singular integral operator is the Hilbert transform $Hf(x) = \pi^{-1}\int_{\mathbb{R}} f(x-t)\,t^{-1}\,dt$ for $x \in \mathbb{R}$. The integral fails to converge absolutely in general, and is defined as the limit as $\varepsilon \to 0$ of the integral over $|t| > \varepsilon$.

$H$ plays a fundamental role in the theory of convergence of the Fourier transform, as well as in one-dimensional complex analysis. It satisfies $\widehat{Hf}(\xi) = i\,\mathrm{sgn}(\xi)\widehat{f}(\xi)$ for all $\xi \neq 0$, where $\mathrm{sgn}(\xi) = \pm 1$ according to whether $\xi > 0$ or $< 0$. Thus $P = \frac{1}{2}(I - iH)$, where $I$ is the identity, is the projection operator onto positive frequencies: $\widehat{Pf}(\xi) = \widehat{f}(\xi)\chi_{\xi>0}$. The "partial sum" operators $\widehat{P_N f}(\xi) = \widehat{f}(\xi)\chi_{|\xi|\le N}$ can be synthesized out of $P$ together with shifts of the Fourier variable, in such a way that uniform boundedness of $P_N$ on $L^p$ is equivalent to boundedness of $H$ on $L^p$. This is the basis of the classical theorem of M. Riesz on $L^p$ norm convergence of Fourier series.

Somewhat more general CZ operators can be expressed as Fourier multiplier operators

$$\widehat{Tf}(\xi) = m(\xi)\widehat{f}(\xi) \qquad \text{where } m(r\xi) \equiv m(\xi) \text{ for all } r > 0 \tag{2}$$

and $m \in C^\infty(\mathbb{R}^d \setminus \{0\})$. General Fourier multipliers $\widehat{Tf}(\xi) = m(\xi)\widehat{f}(\xi)$ with $m \in L^\infty$ preserve $L^p(\mathbb{R})$ only for $p = 2$; there is no characterization of $L^p$ functions in terms of the absolute values of their Fourier coefficients for $p \neq 2$.

The most general Calderón-Zygmund operators in $\mathbb{R}^d$ lack convolution structure, taking the form $\int_{\mathbb{R}^d} K(x,y)f(y)\,dy$ where

$$|K(x,y)| \le C|x-y|^{-d} \quad \text{and} \quad |\nabla_{x,y}K| \le C|x-y|^{-d-1}; \tag{3}$$

again I slur over the issue of interpretation of this typically absolutely divergent integral. Roughly speaking, (3) says that the portions of $f, g$ microlocalized in phase space near $(x,\xi)$ and $(x',\xi')$ respectively interact quite weakly unless $|\xi| + |\xi'| \le C|x-x'|^{-1}$. According to the uncertainty principle, any stronger restriction of this general type is meaningless. A basic theorem [5] states that *if* such an operator is bounded on $L^2$, then it is also bounded on $L^p$ for all $p \in (1, \infty)$.

The basic symmetries of this theory are translation and dilation; if $K(x, y)$ is a Calderón-Zygmund kernel then so are $K(x - z, y - z)$ and $r^d K(rx, ry)$, *uniformly* for all $r > 0$ and $z \in \mathbb{R}^d$. The individual operators need not exhibit these symmetries, but the class as a whole does.

A third basic symmetry, with respect to *modulation*, is totally lacking in this theory. Multiplying $K(x, y)$ by $e^{i(ax+by)}$ for any nonzero $(a, b) \in \mathbb{R}^2$ destroys the bound on $\nabla K$. This lack of symmetry is perhaps even more apparent in (2), in the convolution case $K(x - y)$, where $\xi = 0$ plays a privileged role. Of course, such a modulation does not affect $L^p$ estimates, but as we will see, the bilinear Hilbert transform can be regarded as an infinite sum of modulated Calderón-Zygmund operators with different modulating frequencies, in such a way that boundedness of the sum cannot easily be inferred by summing bounds for the individual summands.

### 1.2. Calderón's commutator

Calderón had an abiding interest in partial differential equations with nonsmooth coefficients and on nonsmooth domains. He had employed algebras of singular integral operators in studying PDE, for instance in his work on uniqueness in the Cauchy problem [2]. Thus he was naturally led to investigate compositions of operators such as the canonical example $H$, the operator $M_A$ of multiplication by a function $A$ having limited smoothness, and $\frac{d}{dx}$. He showed in 1965 [3] that the commutator $[H, M_A]$ is smoothing, in the sense that $\mathcal{C}_A = \frac{d}{dx} \circ [H, M_A]$ is bounded on $L^2(\mathbb{R}^1)$, whenever $A$ is Lipschitz continuous, that is, whenever $a = dA/dx \in L^\infty$. Formally

$$\mathcal{C}_A f(x) = \int_{\mathbb{R}} f(y) \frac{A(x) - A(y)}{(x - y)^2}\, dy, \tag{4}$$

which satisfies the Calderón-Zygmund assumptions (3) when $a = \frac{dA}{dx}$ belongs to $L^\infty$. These operators possess translation and dilation invariance as a family, even though individually they lack it.

Since the commutator operator is not translation-invariant, Plancherel's theorem can not be invoked directly to establish its $L^2$ boundedness. A key realization of Calderón was that it could profitably be regarded as a *bilinear* operator, and that the full force of Fourier analysis and complex variables methods should be brought to bear on $a$.

An intriguing alternative expression is obtained by writing $A(x) - A(y) = (x - y) \int_0^1 a(sx + (1 - s)y)\, ds$ to obtain a decomposition $\mathcal{C}_A(f) = \int_0^1 \mathcal{C}_s(f, a)\, ds$ where

$$\mathcal{C}_s(f, a)(x) = \int_{\mathbb{R}} f(x - t) a(x + st) t^{-1}\, dt. \tag{5}$$

Thus bounds for $\mathcal{C}_s$ from $L^2 \times L^\infty$ to $L^2$ would imply corresponding bounds for the commutator operator. The special case $\mathcal{C}_1$ is traditionally called the *bilinear Hilbert* *transform*, but all the operators $\mathcal{C}_s$ for $s \neq 0, -1$ have essentially the same intrinsic qualities and stature. Calderón asked[(1)] whether these operators do map $L^2 \times L^\infty$ to $L^2$. The problem became notorious, but was not resolved until the work of Lacey and Thiele [21, 22] in the late 1990s.

Thought of as linear operators acting on $f$, $\mathcal{C}_s$ have nonsmooth kernels $K(x,y) = (x-y)^{-1}a(xs + (1-s)y)$ which satisfy no gradient estimate. Viewed as bilinear operators, they are singular in the sense that $\mathcal{C}_s(f,a)(x)$ depends on $a(y_1)f(y_2)$ only for $(y_1,y_2)$ in a one-dimensional subset of $\mathbb{R}^2$.

It is remarkable that these building blocks $\mathcal{C}_s$ not only retain translation and dilation symmetry, but gain new *modulation* symmetries: defining $\mathcal{M}_\eta f(x) = e^{ix\eta}f(x)$,

$$\boxed{\mathcal{C}_s(\mathcal{M}_{s\eta}f, \mathcal{M}_\eta a) \equiv \mathcal{M}_{(1+s)\eta}\mathcal{C}_s(f,a).} \tag{6}$$

These are partial symmetries; there is no relation for $\mathcal{C}_s(\mathcal{M}_\eta f, \mathcal{M}_{\tilde{\eta}} a)$ unless $s\eta = \tilde{\eta}$.

In terms of the Fourier transform the operator is written

$$\mathcal{C}_s(f,a)(x) = c \iint e^{ix(\xi_1+\xi_2)} \operatorname{sgn}(s\xi_2 - \xi_1)\widehat{f}(\xi_1)\widehat{a}(\xi_2)\, d\xi_1\, d\xi_2 \tag{7}$$

for a certain constant $c$, and the modulation symmetry is reflected in the invariance of the Fourier multiplier $\operatorname{sgn}(s\xi_2 - \xi_1)$ under $\xi \mapsto \xi + (s\eta, \eta)$. This multiplier is nonsmooth along an entire line, rather than merely at the origin.

It is (perhaps) a general principle that more symmetric operators are more difficult to analyze; a featureless wall presents no cracks which can naturally be enlarged into gaps. A fundamental point to look for in the discussion below is how the symmetry is broken; see §5.

### 1.3. Carleson's maximal operator

Carleson [6] proved in 1966 that for any periodic function $f \in L^2$ of one real variable, the partial sums of the Fourier series converge to $f$ almost everywhere. The essentially equivalent statement for the real line is that $(2\pi)^{-1} \int_{|\xi|\le N} \widehat{f}(\xi)e^{ix\xi}\, d\xi$ converges to $f(x)$ as $N \to \infty$, for almost every $x \in \mathbb{R}$. The main ingredient is an estimate for Carleson's maximal operator $\mathcal{C}^\star f(x) = \sup_{N<\infty} \big| \int_{|\xi|\le N} \widehat{f}(\xi)e^{ix\xi}\, d\xi \big|$, which is essentially the same as

$$\mathcal{C}^* f(x) = \sup_{N\in\mathbb{R}} \Big| \int_{\mathbb{R}} f(x-t)e^{iNt}t^{-1}\, dt \Big|. \tag{8}$$

Carleson proved that $\mathcal{C}^*$ maps $L^2$ to weak $L^2$, that is, $|\{x : \mathcal{C}^* f(x) > \lambda\}| \le C\lambda^{-2}\|f\|_{L^2}^2$ uniformly for all $\lambda > 0$ and $f \in L^2$. Almost everywhere convergence follows immediately from this inequality since it holds trivially for functions whose Fourier transforms have compact support.

[(1)] The question is widely attributed to Calderón, though I know of no reference.

It is equivalent to establish bounds for the linear operators $\int_{\mathbb{R}} f(x-t)e^{iN(x)t}\, t^{-1}\, dt$ which are *uniform* over all measurable real-valued selection functions $N$. Once again these operators enjoy forms of translation, dilation, and modulation invariance. For instance, $\mathcal{L}(\mathcal{M}_\eta f)(x) = \mathcal{M}_\eta \mathcal{L}' f(x)$, where $\mathcal{L}'$ is obtained from $\mathcal{L}$ by replacing the function $N(x)$ by $N(x) - \eta$.

Fefferman [17] later gave a second proof of Carleson's theorem. Lacey and Thiele used elements of both of these analyses to prove[(2)]

THEOREM 1.1. — *Let $p_1, p_2, q \in (1, \infty]$ satisfy $q^{-1} = p_1^{-1} + p_2^{-1}$, and assume that no more than one of these exponents is infinite. Then there exists $C < \infty$ such that $\|BH(f_1, f_2)\|_{L^q} \le C\|f_1\|_{L^{p_1}}\|f_2\|_{L^{p_2}}$ for all Schwartz class functions.*

### 1.4. Two roads diverge

Calderón proved the bound he sought for the commutator operator without understanding the bilinear Hilbert transform, and went on to analyze [4] the Cauchy integral associated to Lipschitz curves with small Lipschitz constant by an extension of those ideas. Further developments have included a vast literature on elliptic boundary problems on Lipschitz domains, analytic capacity in one complex variable [27], the work of Coifman-Meyer-Mcintosh [9] on the Cauchy integral, and the $T(1)$ theorem of David and Journé [13]. A theory of multilinear Calderón-Zygmund singular operators was developed [10, 12], which however does not include $\mathcal{C}_s$; it encompasses operators which have a Fourier representation like (7) with $\operatorname{sgn}(s\xi_2 - \xi_1)$ replaced by functions smooth away from $\xi = 0$ and satisfying $m(r\xi) \equiv m(\xi)$ for $r > 0$. These operators lack modulation invariance, and are less singular. Some of that theory provides essential building blocks for the analysis outlined here.

## 2. LOCALIZED FOURIER COEFFICIENTS

### 2.1. A frame with a preferred scale

Let $\psi : \mathbb{R}^1 \to \mathbb{C}$ be an infinitely differentiable function supported in $(0, 2)$ such that $\sum_{n\in\mathbb{Z}} \psi(t-n) \equiv 1$ for all $t \in \mathbb{R}$. Then the set of all functions $\{\psi_{k,n} = e^{ikt}\psi(t-n) : k, n \in \mathbb{Z}\}$ is a *frame* for $L^2(\mathbb{R}^1)$; for any $f \in L^2$,

$$f = c\sum_{k,n} \psi_{k,n}\langle f, \psi_{k,n}\rangle \tag{9}$$

for a certain constant $c$ whose precise value is of no consequence for the type of inequality in question here. The inverse Fourier transform of $\psi$ is a Schwartz function, and multiplying it by $(2\pi)^{-1/2}$ yields a function $\varphi$ such that $\{\varphi_{k,n}(x) = e^{inx}\varphi(x - k)\}$ is likewise a frame for $L^2$. It is good intuition to think of $\varphi_{k,n}(x)$ as being

(2) Their theorem actually applies for all $q > \frac{2}{3}$.