Topics in Harmonic Analysis
Lecture 7: Paraproducts, Carleson measures, and the $T(1)$ theorem

Po-Lam Yung

The Chinese University of Hong Kong

Introduction

- Last time we saw an almost orthogonality principle, due to Cotlar and Stein.
- This time we will introduce paraproducts, study Carleson measures and understand the connections of these to BMO.
- All these will come together in the proof of the celebrated $T(1)$ theorem of David and Journé, that characterizes when certain (non-convolution) singular integrals are bounded on L^2.

Outline

- $T(1)$ theorem: statement and applications
- Paraproducts
- Carleson measures and Carleson embedding
- The proof of $T(1)$ theorem

$T(1)$ theorem: statement and applications

- Let $T : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ be a continuous linear operator.
- Suppose there exists a kernel $K_0(x, y)$, defined for $x \neq y$, such that
$$Tf(x) = \int_{\mathbb{R}^n} f(y) K_0(x, y) dy$$
whenever $f \in C_c^\infty(\mathbb{R}^n)$ and x is not in the support of f.
- We assume
$$|K_0(x, y)| \lesssim |x - y|^{-n}$$
and that there exists a fixed $\delta > 0$ such that
$$|K_0(x, y) - K_0(x', y)| \lesssim \frac{|x - x'|^\delta}{|x - y|^{n+\delta}} \quad \text{if } |x - x'| \leq |x - y|/2$$
$$|K_0(x, y) - K_0(x, y')| \lesssim \frac{|y - y'|^\delta}{|x - y|^{n+\delta}} \quad \text{if } |y - y'| \leq |x - y|/2.$$

- Question: When can such an operator T be extended to be a bounded linear operator on $L^2(\mathbb{R}^n)$?
- If it were bounded on $L^2(\mathbb{R}^n)$, then it is bounded on $L^p(\mathbb{R}^n)$ for all $1 < p < \infty$, by the singular integral theorem.
- So we need $T(1) \in BMO$, in the sense that there exists a BMO function $a(x)$, such that for every $g \in C_c^\infty(\mathbb{R}^n)$ with $\int g = 0$, if g is supported on a ball $B(0,R)$ centered at the origin, then whenever $\eta \in C_c^\infty$ and is identically 1 on $B(0,2R)$, we have

$$\langle T(\eta), g \rangle + \int_{\mathbb{R}^n} T(1-\eta)(x) g(x) dx = \int_{\mathbb{R}^n} a(x) g(x) dx$$

where for $x \in B(0,R)$ we define

$$T(1-\eta)(x) := \int_{\mathbb{R}^n} (1-\eta(y))[K_0(x,y) - K_0(0,y)] dy.$$

- Moreover, if $T\colon \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ is continuous, linear, and admits a kernel representation as above, then its adjoint $T^*\colon \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ is also continuous, linear, and admits a kernel representation

$$T^* g(x) = \int_{\mathbb{R}^n} \overline{K_0(y,x)} g(y) dy$$

whenever $g \in C_c^\infty(\mathbb{R}^n)$ and x is not in the support of g.
- If T can be extended to be a bounded linear operator on $L^2(\mathbb{R}^n)$, then so can T^*, so we must have $T^*(1) \in BMO$, in the same way we had $T(1) \in BMO$.

- Finally, if T can be extended to be a bounded linear operator on $L^2(\mathbb{R}^n)$, then T must be weakly bounded, in the sense that

$$\langle T(\phi_R^{x_0}), \psi_R^{x_0} \rangle \leq AR^{-n}$$

whenever $x_0 \in \mathbb{R}^n$, $R > 0$ and $\phi_R^{x_0}, \psi_R^{x_0}$ are normalized bump functions adapted to the ball $B(x_0, R)$; here a normalized bump function adapted to $B(x_0, R)$ is a function of the form

$$x \mapsto \frac{1}{R^n} \phi \left(\frac{x - x_0}{R} \right),$$

where ϕ is a C^∞ function supported in $B(0, 1)$ with

$$\|\partial^\alpha \phi\|_{L^\infty} \leq 1$$

for all α up to some large and fixed order N (whose exact value will be irrelevant for us).
- What is remarkable is that the above 3 conditions are already sufficient.

Theorem ($T(1)$ theorem)

Let $T\colon \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ be a continuous linear operator. Suppose T can be represented by a kernel K_0 as before, where K_0 satisfies

$$|K_0(x,y)| \lesssim |x-y|^{-n},$$

$$|K_0(x,y) - K_0(x',y)| \lesssim \frac{|x-x'|^\delta}{|x-y|^{n+\delta}} \quad \text{if } |x-x'| \leq |x-y|/2, \text{ and}$$

$$|K_0(x,y) - K_0(x,y')| \lesssim \frac{|y-y'|^\delta}{|x-y|^{n+\delta}} \quad \text{if } |y-y'| \leq |x-y|/2;$$

here $\delta > 0$ is some fixed constant. If

(a) $T(1) \in BMO$,

(b) $T^*(1) \in BMO$ *and*

(c) T *is weakly bounded,*

then T extends to a bounded linear operator on $L^2(\mathbb{R}^n)$.

- A classical application of the $T(1)$ theorem is to establish the L^2 boundedness of the Calderón commutators

$$C_k f(x) = \lim_{\varepsilon \to 0} \int_{|x-y|>\varepsilon} \left(\frac{A(x) - A(y)}{x - y} \right)^k \frac{f(y)}{x - y} dy$$

where A is a Lipschitz function on $\mathbb{R}$ and $k \geq 0$ is an integer. Indeed, there exists a constant C such that

$$\|C_k\|_{L^2 \to L^2} \leq C^k \|A'\|_{L^\infty}^k \quad \text{for all } k \geq 0.$$

- This in turn allows one to bound the Cauchy integral along Lipschitz curves with sufficiently small Lipschitz constants: If A is a Lipschitz function on $\mathbb{R}$ with sufficiently small Lipschitz norm, then

$$Tf(x) := -\frac{1}{2\pi i} \lim_{\varepsilon \to 0} \int_{|x-y|>\varepsilon} \frac{f(y)(1 + iA'(y))}{x - y + i(A(x) - A(y))} dy$$

is bounded on $L^2(\mathbb{R})$.
- See Stein's *Harmonic Analysis* for details and further reference (in particular, to $T(b)$ theorem that refines $T(1)$ theorem).

Paraproducts

- The proof of the $T(1)$ theorem consists of two parts: one about reduction to a special case $T(1) = T^*(1) = 0$, and another about the proof of L^2 boundedness in this special case.
- We first carry out the reduction to the special case.
- To do so, we use the following proposition:

Proposition

If $a \in BMO(\mathbb{R}^n)$, then there exists a standard Calderón-Zygmund operator $L_a \colon \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$, such that when appropriately extended as above, we have

$$L_a(1) = a, \quad L_a^*(1) = 0.$$

- Assuming this proposition, then we are led to consider

$$\tilde{T} := T - L_{T(1)} - L^*_{T^*(1)}.$$

Indeed $\tilde{T}$ satisfies all the hypothesis of T, and additionally

$$\tilde{T}(1) = 0, \quad \tilde{T}^*(1) = 0.$$

The goal is then to prove the L^2 boundedness of $\tilde{T}$; since $L_{T(1)}$ and $L^*_{T^*(1)}$ are Calderón-Zygmund operators, they are bounded on L^2. This would prove the L^2 boundedness of T.
- So let's first prove the proposition. Given $a \in BMO(\mathbb{R}^n)$, we construct L_a using paraproducts.

- Let $\psi(\xi)$ be a smooth function with compact support on the unit ball $B(0,2)$, with $\psi(\xi) \equiv 1$ on $B(0,1)$.
- Let $\varphi(\xi) = \psi(\xi) - \psi(2\xi)$ so that ψ is supported on the annulus $\{1/2 \leq |\xi| \leq 2\}$, and

$$\sum_{j \in \mathbb{Z}} \varphi(2^{-j}\xi) = 1 \quad \text{for every } \xi \neq 0.$$

- Let Ψ and Φ be the inverse Fourier transforms of ψ and φ.
- For $j \in \mathbb{Z}$, let $\Psi_j(x) = 2^{jn}\Psi(2^j x)$, $\Phi_j(x) = 2^{jn}\Phi(2^j x)$.
- For $f \in \mathcal{S}'(\mathbb{R}^n)$, let

$$S_j f = f * \Psi_j;$$

for $f \in \mathcal{S}'(\mathbb{R}^n)/\{\text{constants}\}$, let

$$\Delta_j f = f * \Phi_j.$$

Note that $S_j - S_{j-1} = \Delta_j$, and

$$\Delta_k \Delta_j = 0 \quad \text{whenever } |j-k| \geq 2.$$

- If $a, f \in \mathcal{S}(\mathbb{R}^n)$, we have

$$
\begin{aligned}
a \cdot f &= \lim_{J \to +\infty} (S_{J+3}a \cdot S_J f - S_{-J+3}a \cdot S_{-J} f) \\
&= \lim_{J \to +\infty} \sum_{j=-J+1}^{J} (S_{j+3}a \cdot S_j f - S_{j+2}a \cdot S_{j-1} f) \\
&= \sum_{j \in \mathbb{Z}} \Delta_{j+3}a \cdot S_j f + \sum_{j \in \mathbb{Z}} S_{j+2}a \cdot \Delta_j f
\end{aligned}
$$

  (with convergence in say $\mathcal{S}'$ or L^2).
- So each sum on the right hand side is like half of the product of a and f; these are called paraproducts.
- We focus on the first term, and let

$$
L_a(f) = \sum_{j \in \mathbb{Z}} \Delta_{j+3}a \cdot S_j f.
$$

$$L_a(f) = \sum_{j \in \mathbb{Z}} \Delta_{j+3} a \cdot S_j f$$

- The Fourier support of $\Delta_{j+3} a$ is in $\{2^{j+2} \leq |\xi| \leq 2^{j+4}\}$, and that of $S_j f$ is in $\{|\xi| \leq 2^{j+1}\}$.
- Thus the Fourier support of $\Delta_{j+3} a \cdot S_j f$ is in

  $$\{2^{j+1} \leq |\xi| \leq 2^{j+5}\},$$

  and the sum defining $L_a(f)$ is an almost orthogonal sum. (This is why we had chosen to write $j+3$ in place of j in the first place.)
- We will now extend the domain of definition of $L_a(f)$: we define $L_a(f)$ as an element of $\mathcal{S}'(\mathbb{R}^n)$ by the above formula whenever $a \in BMO$ and $f \in \mathcal{S}$.

$$L_a(f) = \sum_{j \in \mathbb{Z}} \Delta_{j+3} a \cdot S_j f$$

- First note that if $a \in BMO$, then

$$\|\Delta_{j+3} a\|_{L^\infty} \lesssim \|a\|_{BMO} \quad \text{uniformly in } j.$$

Hence if $f \in \mathcal{S}$, then

$$\|\Delta_{j+3} a \cdot S_j f\|_{L^2} \lesssim \|a\|_{BMO} \|f\|_{L^2} \quad \text{uniformly in } j.$$

- Next note that $\Delta_{j+3} a \cdot S_j f$ has frequency support in $|\xi| \simeq 2^j$.
- Also, if $g \in \mathcal{S}$, then $\|\Delta_j g\|_{L^2} \lesssim 2^{-|j|n}$ uniformly in j.
- Thus if $g \in \mathcal{S}$, then

$$\sum_{M < |j| < M'} |\langle \Delta_{j+3} a \cdot S_j f, g \rangle| \lesssim 2^{-Mn} \to 0$$

as $M, M' \to \infty$. Thus the sum defining $L_a(f)$ converges in $\mathcal{S}'$.

- From now on, let $a \in BMO(\mathbb{R}^n)$.
- We have just defined $L_a \colon \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$, and it is easy to check that this map is continuous.
- Also, one can check that L_a has a kernel representation

$$L_a f(x) = \int_{\mathbb{R}^n} f(y) K_0(x,y) dy$$

whenever $f \in C_c^\infty(\mathbb{R}^n)$ and x is not in the support of f, where

$$K_0(x,y) := \sum_{j \in \mathbb{Z}} (\Delta_{j+3} a)(x) \Psi_j(x-y) \quad \text{for } x \neq y.$$

- Since $\|\Delta_{j+3} a\|_{L^\infty} \lesssim \|a\|_{BMO}$ uniformly in j, it is easy to check that

$$|K_0(x,y)| \lesssim |x-y|^{-n}.$$

Similarly, $|\partial_{x,y}^\lambda K_0(x,y)| \lesssim |x-y|^{-n-|\lambda|}$ for all multiindices λ.

- Chasing through the definitions, we see that $L_a(1) = a$ and $L_a^*(1) = 0$.
- For instance, suppose $g \in C_c^\infty(B(0,R))$ with $\int g = 0$, and $\eta \in C_c^\infty$ is identically 1 on $B(0, 2R)$. Then

$$\langle L_a \eta, g\rangle + \int_{\mathbb{R}^n} L_a(1-\eta)(x)g(x)dx$$
$$= \lim_{J\to\infty} \sum_{|j|\leq J} \langle \Delta_{j+3} a \cdot S_j \eta, g\rangle$$
$$+ \lim_{J\to\infty} \sum_{|j|\leq J} \int\int (1-\eta)(y)\Delta_{j+3}a(x)\Psi_j(x-y)g(x)dydx,$$

where we used Fubini to evaluate $\int_{\mathbb{R}^n} L_a(1-\eta)(x)g(x)dx$; this is possible because the supports of $(1-\eta)$ and g are disjoint.
- The sum of the above two limits is equal to

$$\lim_{J\to\infty} \sum_{|j|\leq J} \int \Delta_{j+3}a(x)g(x)dx,$$

and we want to show that it is equal to $\int a(x)g(x)dx$.

- But

$$\sum_{|j|\leq J}\int \Delta_{j+3}a(x)g(x)dx = \int a(x)[S_{J+3}g(x) - S_{-J+2}g(x)]dx,$$

and $S_{J+3}g(x) \to g(x)$ in $\mathcal{S}(\mathbb{R}^n)$ as $J \to +\infty$. So it remains to show that

$$\int a(x)S_{-J+2}g(x)dx \to 0 \quad \text{as } J \to \infty.$$

- Now since $\int_{\mathbb{R}^n} g(y)dy = 0$, we have

$$\begin{aligned}
&\int a(x)S_{-J+2}g(x)dx \\
&= \int [a(x) - a_{B(0,2^J)}]S_{-J+2}g(x)dx \\
&= \iint [a(x) - a_{B(0,2^J)}][\Psi_{-J+2}(x-y) - \Psi_{-J+2}(x)]g(y)dydx
\end{aligned}$$

- Since g has compact support and Ψ is Schwartz, using the mean-value theorem, we have

$$|\Psi_{-J+2}(x-y)-\Psi_{-J+2}(x)|\lesssim 2^{-J}\frac{2^{-Jn}}{(1+2^{-J}|x|)^{n+1}}$$

for any y in the support of g and any $x\in\mathbb{R}^n$, where the implicit constant depends on the support of g.

- Thus

$$\begin{aligned}&\left|\int a(x)S_{-J+2}g(x)dx\right|\\ \leq&2^{-J}\iint\left|a(x)-a_{B(0,2^J)}\right|\frac{2^{-Jn}}{(1+2^{-J}|x|)^{n+1}}\left|g(y)\right|dydx\\ \lesssim&2^{-J}\|a\|_{BMO}\|g\|_{L^1}\to 0\end{aligned}$$

as $J\to+\infty$. This proves $L_a(1)=a$. Similarly $L_a^*(1)=0$.

- To finish the proof of the proposition, we just need to show that L_a extends to a bounded linear operator on L^2.
- By almost orthogonality between the summands defining $L_a(f)$, it suffices to prove the following claim:

$$\sum_{j\in\mathbb{Z}} \|\Delta_{j+3}a \cdot S_j f\|_{L^2}^2 \lesssim \|f\|_{L^2}^2 \quad \text{whenever } f \in \mathcal{S}.$$

- One may be tempted to prove the above claim by bounding $\|\Delta_{j+3}a\|_{L^\infty} \lesssim \|a\|_{BMO}$, and summing $\sum_{j\in\mathbb{Z}} \|S_j f\|_{L^2}^2$; unfortunately this does not work, for the latter sum is usually divergent.
- So the proof of the claim must proceed differently. It will rely on the notion of Carleson measures.

Carleson measures and Carleson embedding

- A measure $d\mu$ on the upper half space $\mathbb{R}^{n+1}_+$ is said to be a Carleson measure, if there exists a constant C, such that

$$d\mu(B(x,r)\times(0,r)) \leq C|B(x,r)|$$

for every ball $B(x,r) \subset \mathbb{R}^n$.

- We think of the smallest such C as the norm of the Carleson measure, written $\|d\mu\|_{\mathcal{C}}$.
- We will need two lemmas, one connecting BMO functions to Carleson measures, and another for estimating integrals involving Carleson measures.

Lemma (Carleson embedding)

If $a \in BMO(\mathbb{R}^n)$, then

$$d\mu := \sum_{j\in\mathbb{Z}} \delta_{2^{-j}}(y)|\Delta_{j+3}a(x)|^2 dx$$

is a Carleson measure on $\mathbb{R}^{n+1}_+$, with $\|d\mu\|_{\mathcal{C}} \lesssim \|a\|^2_{BMO}$.

Lemma

If $d\mu$ is a Carleson measure on $\mathbb{R}^{n+1}_+$, and $F(x,y)$ is a measurable function on $\mathbb{R}^{n+1}_+$, then

$$\int_{\mathbb{R}^{n+1}_+} |F(x,y)|^p d\mu \le 3^n \|F^*\|^p_{L^p(\mathbb{R}^n)} \|d\mu\|_{\mathcal{C}}$$

for all $1 \le p < \infty$. Here F^* is the non-tangential maximal function

$$F^*(x) = \sup_{y>0} \sup_{z\in B(x,y)} |F(z,y)|.$$

- Apply the second lemma with $d\mu$ being the Carleson measure from the first lemma, $F(x, 2^{-j}) = S_j f(x)$ with $f \in \mathcal{S}(\mathbb{R}^n)$, and $p = 2$, we see that

$$\int_{\mathbb{R}^n} \sum_{j \in \mathbb{Z}} |\Delta_{j+3} a(x) \cdot S_j f(x)|^2 dx \lesssim \|Mf\|_{L^2}^2 \|a\|_{BMO}^2$$

  where M is the Hardy-Littlewood maximal function. Thus by the L^2 boundedness of M, our earlier claim follows.
- It remains to prove the lemmas.
- We will give the proof of the lemmas beginning next slide; let us just pause to mention that there is a converse to the Carleson embedding lemma characterizing BMO. For details, see Stein's *Harmonic Analysis*, Chapter IV, Section 4.3.

- We begin with the proof of the second lemma.
- Let $d\mu$ be a Carleson measure on $\mathbb{R}^{n+1}_+$, and $F(x,y)$ be measurable on $\mathbb{R}^{n+1}_+$. We want to prove

$$\int_{\mathbb{R}^{n+1}_+} |F(x,y)|^p d\mu \leq 3^n \|F^*\|^p_{L^p(\mathbb{R}^n)} \|d\mu\|_{\mathcal{C}}$$

for all $1 \leq p < \infty$.
- Since $(F^*)^p = (|F|^p)^*$, we may assume that $p = 1$.
- By monotone convergence, we may also assume that F is supported on $B(0,R) \times (0,R)$ for some $R > 0$.

- For $t > 0$, let $O_t = \{(x,y) \in \mathbb{R}^{n+1}_+ \colon |F(x,y)| > t\}$. It suffices to show that

$$d\mu(O_t) \leq 3^n \|d\mu\|_{\mathcal{C}} \left|\{x \in \mathbb{R}^n \colon F^*(x) > t\}\right|.$$

(because one gets the desired inequality by integrating in t).
- To do so, let $\mathcal{B}_t$ be a collection of balls in $\mathbb{R}^n$, defined by

$$\mathcal{B}_t = \{B(x,y) \colon (x,y) \in O_t\}.$$

- We claim that there exists a countable collection of pairwise disjoint balls $\{B_i\}_{B_i \in \mathcal{B}_t}$, such that if $B_i = B(x_i, y_i)$, then

$$\bigcup_i B(x_i, 3y_i) \times (0, 3y_i) \quad \text{covers } O_t.$$

- Indeed, the supremum of the radii of all balls in $\mathcal{B}_t$ is finite, since F is supported on $B(0, R) \times (0, R)$. Choose B_1 to be a ball in $\mathcal{B}_t$ so that the radius of B_1 is at least half of the supremum of the radii of all balls in $\mathcal{B}_t$. Remove all balls in $\mathcal{B}_t$ that intersects B_1, and select B_2 in the remaining collection so that the radius of B_2 is at least half of the supremum of the radii of all balls that remained. Repeat, stopping only if there are no balls left.
- If the process never stops, then the supremum of the radii of all remaining balls after the j-th step tends to zero as $j \to \infty$, since F is supported on $B(0, R) \times (0, R)$ for some $R > 0$.
- If $(x, y) \in O_t$, then $B(x, y) \in \mathcal{B}_t$, so $B(x, y)$ intersects one of the chosen B_i's, with $y \leq 2y_i$ where y_i is the radius of B_i. So if $B_i = B(x_i, y_i)$, then $x \in B(x_i, 3y_i)$ and $y \in (0, 3y_i)$. This proves the claim.

- From the claim, we have

$$\begin{aligned} d\mu(O_t) &\leq \sum_i d\mu(B(x_i, 3y_i) \times (0, 3y_i)) \\ &\leq \|d\mu\|_{\mathcal{C}} \sum_i |B(x_i, 3y_i)| \\ &\leq 3^n \|d\mu\|_{\mathcal{C}} \left| \bigcup_i B(x_i, y_i) \right| \end{aligned}$$

  The latter is

$$\leq 3^n \|d\mu\|_{\mathcal{C}} |\{x \in \mathbb{R}^n \colon F^*(x) > t\}|$$

  as promised, since from $(x_i, y_i) \in O_t$, we get $|F(x_i, y_i)| > t$, so for any $x \in \bigcup_i B(x_i, y_i)$, we have $F^*(x) > t$.
- This finishes the proof of the second lemma.

- We now prove the Carleson embedding lemma.
- Let $a \in BMO(\mathbb{R}^n)$, and

$$d\mu = \sum_{j \in \mathbb{Z}} \delta_{2^{-j}}(y) |\Delta_{j+3} a(x)|^2 dx.$$

- We need to show that $d\mu$ is a Carleson measure on $\mathbb{R}^{n+1}_+$, with $\|d\mu\|_{\mathcal{C}} \lesssim \|a\|_{BMO}^2$. This will follow if we show that

$$\int_{B(x_0,r)} \sum_{2^{-j} \leq r} |\Delta_j a(x)|^2 dx \lesssim r^n \|a\|_{BMO}^2$$

  for any $x_0 \in \mathbb{R}^n$ and any $r > 0$.
- By translation and dilation invariance, we may assume

$$x_0 = 0, \quad r = 1 \quad \text{and} \quad \|a\|_{BMO} = 1.$$

  Write $B = B(x_0, r) = B(0,1)$, and $B^* = B(0,2)$.
- Since $\Delta_j c = 0$ for any constant c, by further subtracting a constant, we may assume $⨍_{B^*} a = 0$.

- Our goal now is to prove that

$$\int_B \sum_{j\geq 0} |\Delta_j a(x)|^2 dx \lesssim 1$$

under the above assumptions.

- Let $a = a_1 + a_2$, where $a_1 = a\chi_{B^*}$ and $a_2 = a\chi_{(B^*)^c}$.
- Since $⨍_{B^*} a = 0$, by John-Nirenberg inequality, we have

$$\|a_1\|_{L^2}^2 = \int_{B^*} |a(y)|^2 dy \lesssim |B^*| \|a\|_{BMO}^2 \lesssim 1.$$

Hence by Plancherel,

$$\int_B \sum_{j\geq 0} |\Delta_j a_1(x)|^2 dx \lesssim \int_{\mathbb{R}^n} \sum_{j\in\mathbb{Z}} |\Delta_j a_1(x)|^2 dx \lesssim \|a_1\|_{L^2}^2 \lesssim 1.$$

- We claim that

$$\int_B \sum_{j\geq 0} |\Delta_j a_2(x)|^2 dx \lesssim 1.$$

Indeed it suffices to show that

$$\sum_{j\geq 0} |\Delta_j a_2(x)|^2 \lesssim 1 \quad \text{for all } x \in B.$$

- But for each $j \geq 0$ and each $x \in B$, we have

$$\Delta_j a_2(x) = \int_{y\in(B^*)^c} a(y)\Phi(2^j(x-y))2^{jn}dy.$$

Since $|\Phi(2^j(x-y))| \lesssim (2^j|y|)^{-(n+1)}$, we have

$$|\Delta_j a_2(x)| \lesssim 2^{-j} \int_{|y|\geq 2} |a(y)| \frac{dy}{|y|^{n+1}}$$

$$|\Delta_j a_2(x)| \lesssim 2^{-j} \int_{|y|\geq 2} |a(y)| \frac{dy}{|y|^{n+1}}$$

- By decomposing $\{|y| \geq 2\}$ into dyadic annuli and noting that $|a_{B(0,2^k)} - a_{B(0,2^{k+1})}| \lesssim 1$ uniformly in k, we see that

$$\int_{|y|\geq 2} |a(y)| \frac{dy}{|y|^{n+1}} \lesssim \|a\|_{BMO} \lesssim 1.$$

(c.f. Question 8(d) of Homework 5.)

- Altogether, this shows

$$\sum_{j\geq 0} |\Delta_j a_2(x)|^2 \lesssim \sum_{j\geq 0} 2^{-2j} \lesssim 1 \quad \text{for all } x \in B,$$

as desired.

- This completes the proof of the Carleson embedding lemma.

The proof of $T(1)$ theorem

- Let's recap what we have so far.
- For $a \in BMO(\mathbb{R}^n)$, $f \in \mathcal{S}(\mathbb{R}^n)$, we defined the paraproduct
$$L_a(f) = \sum_{j \in \mathbb{Z}} \Delta_{j+3} a \cdot S_j f,$$
and the above showed that L_a extends as a bounded linear operator on L^2.
- Earlier we also saw that the integral kernel of L_a satisfies appropriate differential inequalities, so that L_a is a standard Calderón-Zygmund operator on $\mathbb{R}^n$.
- We also have $L_a(1) = a$, $L_a^*(1) = 0$.
- This shows that in proving the $T(1)$ theorem, we may assume in addition that $T(1) = T^*(1) = 0$.
- To do so we decompose T into almost orthogonal pieces, and use Cotlar-Stein; indeed when $T(1) = T^*(1) = 0$, T will be shown to be similar enough to a translation-invariant singular integral, so that almost orthogonality works.

- Let's modify our earlier definitions of S_j and Δ_j as follows.
- Let $\Psi \in C_c^\infty(\mathbb{R}^n)$ be supported on $B(0,1)$, with $\Psi \equiv 1$ on $B(0,1/2)$, and $\int_{\mathbb{R}^n} \Psi(x)dx = 1$. Let

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

  so that $\int_{\mathbb{R}^n} \Phi(x)dx = 0$.
- For $j \in \mathbb{Z}$, write $\Psi_j(x) = 2^{jn}\Psi(2^j x)$ and $\Phi_j(x) = 2^{jn}\Phi(2^j x)$. For $f \in \mathcal{S}(\mathbb{R}^n)$, also let

$$S_j f := f * \Psi_j, \quad \Delta_j f := f * \Phi_j$$

  so that $\Delta_j = S_j - S_{j-1}$.
- While $S_j f$ and $\Delta_j f$ does not have compact Fourier support, morally speaking they are frequency localized to a ball of radius 2^j and an annulus of size 2^j respectively.

- If $f \in \mathcal{S}(\mathbb{R}^n)$, we claim that

$$Tf = \lim_{J \to +\infty} S_J T S_J f \quad \text{and} \quad \lim_{J \to +\infty} S_{-J} T S_{-J} f = 0,$$

where convergence are in the topology of $\mathcal{S}'(\mathbb{R}^n)$.

- To prove this claim, first note that since $T \colon \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ is continuous and linear, by the Schwartz kernel theorem, there exists a tempered distribution K on $\mathbb{R}^n \times \mathbb{R}^n$ such that

$$\langle Tf, g \rangle = \langle K, f \otimes g \rangle \quad \text{for all } f, g \in \mathcal{S}(\mathbb{R}^n);$$

here the pairing on the left is a pairing of $Tf \in \mathcal{S}'(\mathbb{R}^n)$ with $g \in \mathcal{S}(\mathbb{R}^n)$, while the pairing on the right is between $K \in \mathcal{S}'(\mathbb{R}^n \times \mathbb{R}^n)$ with $f \otimes g(x, y) := f(x)g(y) \in \mathcal{S}(\mathbb{R}^n \times \mathbb{R}^n)$.

- Since $S_J f \otimes S_J g \to f \otimes g$ in the topology of $\mathcal{S}(\mathbb{R}^n \times \mathbb{R}^n)$ as $J \to \infty$, it follows that

$$\langle S_J T S_J f, g \rangle = \langle K, S_J f \otimes S_J g \rangle \to \langle K, f \otimes g \rangle = \langle Tf, g \rangle$$

as $J \to \infty$.

- The second part of the claim amounts to saying that

$$\lim_{J\to+\infty}\langle TS_{-J}f, S_{-J}^{*}g\rangle = 0$$

for all $f, g \in \mathcal{S}(\mathbb{R}^n)$. By our previous argument and the density of $C_c^\infty(\mathbb{R}^n)$ in $\mathcal{S}(\mathbb{R}^n)$, we assume $f, g \in C_c^\infty(\mathbb{R}^n)$.
- In that case, for J large enough, $S_{-J}f$ is a normalized bump function on $B(0, 2^{J+1})$, and so is $S_{-J}^{*}g$.
- Since T is weakly bounded, this shows

$$|\langle TS_{-J}f, S_{-J}^{*}g\rangle| \lesssim 2^{-nJ} \to 0$$

as $J \to +\infty$, as desired.

- So we have proved that for $f \in \mathcal{S}(\mathbb{R}^n)$

$$Tf = \lim_{J \to +\infty} (S_J T S_J f - S_{-J-1} T S_{-J-1} f)$$
$$= \lim_{J \to +\infty} \sum_{|j| \leq J} (\Delta_j T S_j f + S_{j-1} T \Delta_j f)$$

where the convergence is in $\mathcal{S}'(\mathbb{R}^n)$.

- Using $T(1) = 0$ and Cotlar-Stein, we will show that

$$\left| \left\langle \sum_{|j| \leq J} \Delta_j T S_j f, g \right\rangle \right| \lesssim \|f\|_{L^2} \|g\|_{L^2}$$

for any $f, g \in \mathcal{S}(\mathbb{R}^n)$, uniformly in $J \in \mathbb{N}$; similarly, using $T^*(1) = 0$ instead, one has a corresponding estimate for $\left| \left\langle \sum_{|j| \leq J} S_{j-1} T \Delta_j f, g \right\rangle \right|$. This would prove that

$$|\langle Tf, g \rangle| \lesssim \|f\|_{L^2} \|g\|_{L^2}$$

for all $f, g \in \mathcal{S}(\mathbb{R}^n)$, completing the proof of $T(1)$ theorem.

- For $j \in \mathbb{Z}$, let $T_j := \Delta_j T S_j$. Our goal is to show that

$$\left|\left\langle T^{(J)} f, g \right\rangle\right| \lesssim \|f\|_{L^2} \|g\|_{L^2} \quad \text{for any } f, g \in C_c^\infty(\mathbb{R}^n)$$

uniformly in $J \in \mathbb{N}$, where $T^{(J)} := \sum_{|j| \leq J} T_j$.

- First we determine the integral kernel of the operator T_j. Let

$$\Psi_j^y(w) = 2^{jn} \Psi(2^j(w-y)) \quad \text{and} \quad \Phi_j^x(v) = 2^{jn} \Phi(2^j(x-v)).$$

These are normalized bump functions in $B(y, 2^{-j})$ and $B(x, 2^{-j})$. Define, for $x, y \in \mathbb{R}^n$,

$$k_j(x,y) := \langle T(\Psi_j^y), \Phi_j^x \rangle.$$

Then using the continuity of $T \colon \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$, one can check that for $f, g \in C_c^\infty(\mathbb{R}^n)$, we have

$$\langle T_j f, g \rangle = \int_{\mathbb{R}^n \times \mathbb{R}^n} k_j(x,y) f(y) g(x) dx dy.$$

- Let δ be the Hölder exponent in the assumed estimates for the kernel K_0 of T. We claim

$$|k_j(x,y)| \lesssim \frac{2^{jn}}{(1+2^j|x-y|)^{n+\delta}} \tag{1}$$

$$|\partial_x^\alpha \partial_y^\beta k_j(x,y)| \lesssim \frac{2^{j(n+|\alpha|+|\beta|)}}{(1+2^j|x-y|)^{n+\delta}} \tag{2}$$

$$\int_{\mathbb{R}^n} k_j(x,y)dy = 0 \quad \text{for all } x \in \mathbb{R}^n \tag{3}$$

$$\int_{\mathbb{R}^n} k_j(x,y)dx = 0 \quad \text{for all } y \in \mathbb{R}^n \tag{4}$$

- Indeed (1) and (2) will follow from the weak boundedness of T and the kernel estimates for K_0, while (3) will follow from the assumption $T(1) = 0$ and kernel estimates for K_0.
- (4) is similar to (3) except that it is easier; one will not need the assumption on $T(1)$.
- The above claims in turn allow us to invoke a proposition from the end of the last lecture, which shows that Cotlar-Stein applies, and $T^{(J)}$ is uniformly bounded on L^2, as desired.

- It remains to prove the claims. Recall

$$k_j(x,y) := \langle T(\Psi_j^y), \Phi_j^x \rangle.$$

- First, to prove

$$|k_j(x,y)| \lesssim \frac{2^{jn}}{(1+2^j|x-y|)^{n+\delta}},$$

note that this follows from weak boundedness of T if $|x-y| \lesssim 2^{-j}$. Otherwise the supports of Ψ_j^y and Φ_j^x are at a distance $\gtrsim 2^{-j}$. In that case $k_j(x,y)$ can be written as

$$\int_{\mathbb{R}^n \times \mathbb{R}^n} K_0(v,w)\Psi_j^y(w)\Phi_j^x(v)dvdw.$$

Since $\int_{\mathbb{R}^n} \Phi_j^x(w)dw = 0$, this shows

$$k_j(x,y) = \int_{\mathbb{R}^n \times \mathbb{R}^n} (K_0(v,w) - K_0(x,w))\Psi_j^y(w)\Phi_j^x(v)dvdw$$

- Since on the supports of the integral, we have $|v - x| \lesssim 2^{-j}$ and $|v - w| \simeq |x - w| \simeq |x - y|$, we have, by our assumption on K_0, that

$$|k_j(x,y)| \lesssim 2^{-j\delta}|x-y|^{-n-\delta} = \frac{2^{jn}}{(2^j|x-y|)^{n+\delta}},$$

  as desired. This proves (1).
- (2) follows by observing that

$$|\partial_x^\alpha \partial_y^\beta k_j(x,y)| = \left|\langle T(\partial^\beta \Psi_j^y), \partial^\alpha \Phi_j^x \rangle\right|$$

  and modifying the argument that proved (1); the key is that $2^{-j|\beta|}\partial^\beta \Psi_j^y(w)$ and $2^{-j|\alpha|}\partial^\alpha \Phi_j^x(v)$ are also normalized bump functions on $B(y, 2^{-j})$ and $B(x, 2^{-j})$ respectively.

- To prove (3), it suffices to prove that

$$\lim_{R\to+\infty}\int_{|y|\leq R} k_j(x,y)dy = 0.$$

But since $k_j(x,y) = \langle T(\Psi_j^y), \Phi_j^x\rangle$, by continuity of $T\colon \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$, we have

$$\int_{|y|\leq R} k_j(x,y)dy = \langle T(h_{j,R}), \Phi_j^x\rangle$$

where

$$h_{j,R}(w) := \int_{|y|\leq R} \Psi_j^y(w)dy = 2^{jn}\int_{|y|\leq R} \Psi(2^j(w-y))dy.$$

- Clearly

$$h_{j,R}(w) = 0 \quad \text{whenever } |w| \leq R + 2^{-j}.$$

We also have

$$h_{j,R}(w) = 1 \quad \text{whenever } |w| \leq R - 2^{-j}.$$

- This suggests one to use the condition $T(1) = 0$, and write

$$\int_{|y|\leq R} k_j(x,y)dy = \langle T(h_{j,R}-1), \Phi_j^x \rangle.$$

- If $R - 2^{-j} \geq 2(|x| + 2^{-j})$, then $h_{j,R} - 1$ and Φ_j^x have disjoint supports, so

$$\begin{aligned}
&|\langle T(h_{j,R}-1), \Phi_j^x \rangle| \\
=&\left| \int_{\mathbb{R}^n \times \mathbb{R}^n} [K_0(v,w) - K_0(x,w)](h_{j,R}-1)(w)\Phi_j^x(v)dvdw \right| \\
\leq& \int_{|w|\geq R-2^{-j}} \int_{|v-x|\leq 2^{-j}} |K_0(v,w) - K_0(x,w)||\Phi_j^x(v)|dvdw \\
\lesssim& 2^{-j\delta}(R-2^{-j})^{-\delta} \to 0
\end{aligned}$$

as $R \to +\infty$. This completes the proof of (3).

- Similarly,
$$\int_{|x|\le R} k_j(x,y)dx = \langle T(\Psi_j^y), H_{j,R}\rangle$$
where
$$H_{j,R}(v) := \int_{|x|\le R} \Phi_j^x(w)dx = 2^{jn}\int_{|x|\le R} \Phi(2^j(x-v)).$$
- $H_{j,R}(v)$ is supported on the annulus $R-2^{-j}\le |v|\le R+2^{-j}$.
- If $R-2^{-j}\ge 2(|y|+2^{-j})$, then $H_{j,R}$ and Ψ_j^y have disjoint supports, so
$$\begin{aligned}
&|\langle T(\Psi_j^y), H_{j,R}\rangle| \\
=&\left|\int_{\mathbb{R}^n\times\mathbb{R}^n} K_0(v,w)\Psi_j^y(w)H_{j,R}(v)dvdw\right| \\
\le&\int_{R-2^{-j}\le|v|\le R+2^{-j}}\int_{|w-y|\le 2^{-j}} |K_0(v,w)||\Psi_j^y(w)|dvdw \\
\lesssim& 2^{-j}R^{n-1}(R-2^{-j})^{-n}\to 0
\end{aligned}$$
as $R\to+\infty$. This completes the proof of (4).