

## Harmonic Analysis and Its Applications

In these lectures, we concentrate on the motivations, development and applications of the Calderon-Zygmund operator theory.

### Lecture 1. The differential operators with constant coefficients and the first generation of Calderon-Zygmund operators

Consider the following differential operator with constant coefficients:

$$Lu(x) = \sum_{\alpha} a_{\alpha} \frac{\partial^{\alpha} u}{\partial x^{\alpha}}. \quad (1.1)$$

By taking the Fourier transform,

$$(\widehat{Lu})(\xi) = \sum_{\alpha} a_{\alpha} (-2\pi i \xi)^{\alpha} \hat{u}(\xi). \quad (1.2)$$

This suggests one to consider the following more general Fourier multiplier:

**Definition 1.3:** An operator  $T$  is said to be the Fourier multiplier if

$$(\widehat{Tf})(\xi) = m(\xi) \hat{f}(\xi). \quad (1.4)$$

(1.2) shows that any classical differential operator is a Fourier multiplier.

**Example 1:** Suppose  $f \in L^2(\mathbb{R})$  and  $F$  is an analytic extension of  $f$  on  $\mathbb{R}_+^2$  given by

$$\begin{aligned} F(x + iy) &= -\frac{i}{\pi} \int \frac{1}{x + iy - t} f(t) dt = -\frac{i}{\pi} \int \frac{(x - t) - iy}{(x - t)^2 + y^2} f(t) dt \\ &= \frac{1}{\pi} \int \frac{y}{(x - t)^2 + y^2} f(t) dt - \frac{i}{\pi} \int \frac{(x - t)}{(x - t)^2 + y^2} f(t) dt. \end{aligned} \quad (1.5)$$

Letting  $y \rightarrow 0$ , then  $\frac{1}{\pi} \int \frac{y}{(x-t)^2+y^2} f(t) dt \rightarrow f(x)$  for a. e.  $x$ , and, in general, the second term above has no limit. However, one can show p.v  $\int \frac{1}{x-t} f(t) dt$  exists for a. e.  $x$ . Thus

$$\lim_{y \rightarrow 0} F(x + iy) = f(x) + iH(f)(x) \quad (1.6)$$

where  $H$  is called the Hilbert transform defined by

$$H(f)(x) = -\frac{1}{\pi} \int \frac{f(t)}{x - t} dt. \quad (1.7)$$

**Example 2:** Consider the Laplacian

$$\Delta u = \sum_{j=1}^n \frac{\partial^2 u}{\partial x_j^2}. \quad (1.8)$$

By taking the Fourier transform,

$$(-\widehat{\Delta u})(\xi) = \sum_{j=1}^n (4\pi|\xi_j|)^2 \hat{u}(\xi) = 4\pi|\xi|^2 \hat{u}(\xi).$$

Define the Riesz transforms  $R_j, 1 \leq j \leq n$ , by

$$(\widehat{R_j f})(\xi) = \frac{\xi_j}{|\xi|} \hat{f}(\xi). \quad (1.9)$$

Then

$$\left(\widehat{\frac{\partial^2 u}{\partial x_i \partial x_j}}\right)(\xi) = -4\pi\xi_i \xi_j \hat{u}(\xi) = (R_i \widehat{R_j \Delta u})(\xi). \quad (1.10)$$

Thus,

$$\frac{\partial^2 u}{\partial x_i \partial x_j} = R_i R_j \Delta u. \quad (1.11)$$

Since, it is easy to see that  $H$  and  $R_j, 1 \leq j \leq n$ , are bounded on  $L^2$ , so we obtain

$$\| \lim_{y \rightarrow 0} F(x + iy) \|_2 \leq C \|f\|_2 \quad (1.12)$$

and

$$\left\| \frac{\partial^2 u}{\partial x_i \partial x_j} \right\|_2 = \|R_i R_j \Delta u\|_2 \leq C \| \Delta u \|_2. \quad (1.13)$$

The Hilbert transform, by taking the Fourier transform, can be written as

$$\widehat{H(f)}(\xi) = -i \operatorname{sign}(\xi) \hat{f}(\xi) \quad (1.14)$$

and the Riesz transform, by taking the inverse Fourier transform, can be written as

$$R_j(f)(x) = c_n p.v. \int \frac{y_j}{|y|^{1+n}} f(x-y) dy, 1 \leq j \leq n. \quad (1.15)$$

In 1952, Calderon and Zygmund introduced the following first generation of singular integral operators:

**Definition 1.16:**

$$T(f)(x) = p.v. \int \frac{\Omega(y)}{|y|^n} f(x-y) dy,$$

where  $\Omega$  satisfies the following conditions:

$$\Omega(\lambda y) = \Omega(y) \quad (1.17)$$

for all  $\lambda > 0$ ;

$$\Omega \in C^1(S^{n-1}); \quad (1.18)$$

$$\int \Omega(y) d\sigma(y) = 0. \quad (1.19)$$

First we point out that the first generation of Calderon-Zygmund operators are well defined on  $\mathcal{S}(R^n)$ , the Schwartz test function space. To see this, one can define

$$\begin{aligned} p.v. \int \frac{\Omega(y)}{|y|^n} f(y) dy &= \lim_{\epsilon \rightarrow 0} \int_{|y| > \epsilon} \frac{\Omega(y)}{|y|^n} f(y) dy = \\ &= \int_{|y| < 1} \frac{\Omega(y)}{|y|^n} [f(y) - f(0)] dy + \int_{|y| \geq 1} \frac{\Omega(y)}{|y|^n} f(y) dy \end{aligned} \quad (1.20)$$

for all  $f \in \mathcal{S}$  since  $\Omega$  has zero average. It is then easy to see that both integrals above converge.

**Remark 1.21:** A necessary condition for  $p.v. \int \frac{\Omega(y)}{|y|^n} f(x-y) dy$  exists is that  $\Omega$  has zero average on  $S^{n-1}$ . In fact, let  $f \in \mathcal{S}$  be such that  $f(x) = 1$  for  $|x| \leq 2$ . Then for  $|x| < 1$ ,

$$T(f)(x) = \lim_{\epsilon \rightarrow 0} \int_{\epsilon < |y| < 1} \frac{\Omega(y)}{|y|^n} dy + \int_{|y| \geq 1} \frac{\Omega(y)}{|y|^n} f(x-y) dy. \quad (1.21)$$

The second integral is convergent but the first equals  $\lim_{\epsilon \rightarrow 0} \int_{S^{n-1}} \Omega(y) d\sigma(y) \log(\frac{1}{\epsilon})$ . Thus, if this limit is finite, then the integral of  $\Omega$  on  $S^{n-1}$  is zero.

**Theorem 1.22(Calderon and Zygmund):** If  $T$  is an operator of the first generation of Calderon-Zygmund operator, then  $T$  is bounded on  $L^p$ ,  $1 < p < \infty$ . Moreover, there is a constant  $C$  such that

$$\|T(f)\|_p \leq C \|f\|_p. \quad (1.23)$$

The method of the proof of theorem 1.22 is called the real variable method of Calderon and Zygmund. This method includes the following steps.

**Step 1:**  $T$  is bounded on  $L^2$ . To do this, since  $T$  is a convolution operator, by the Plancherel theorem, it suffices to show the Fourier transform of  $K$ , the kernel of  $T$ , is bounded. In fact, we will show that for  $\xi \in S^{n-1}$ ,

$$m(\xi) = \hat{K}(\xi) = \int_{S^{n-1}} \Omega(t) \left[ \log\left(\frac{1}{|t \cdot \xi|}\right) - i \frac{\pi}{2} \text{sign}(t \cdot \xi) \right] d\sigma(t). \quad (1.24)$$

Indeed, since  $K$  is homogeneous of degree  $-n$ , so  $m(\xi)$  is homogeneous of degree 0. Therefore, we may assume that  $\xi \in S^{n-1}$ . Since  $\Omega$  has zero average,

$$\begin{aligned}
m(\xi) &= \lim_{\epsilon \rightarrow 0} \int_{\epsilon < |y| < \frac{1}{\epsilon}} \frac{\Omega(y)}{|y|^n} e^{-2\pi i y \cdot \xi} dy \\
&= \lim_{\epsilon \rightarrow 0} \int_{S^{n-1}} \Omega(y) \left[ \int_{\epsilon}^1 (e^{-2\pi i t y \cdot \xi} - 1) \frac{dt}{t} + \int_1^{\frac{1}{\epsilon}} e^{-2\pi i t y \cdot \xi} \frac{dt}{t} \right] d\sigma(y) \\
&= \lim_{\epsilon \rightarrow 0} \int_{S^{n-1}} \Omega(y) \left[ \int_{\epsilon}^1 (\cos(2\pi t y \cdot \xi) - 1) \frac{dt}{t} + \int_1^{\frac{1}{\epsilon}} \cos(2\pi t y \cdot \xi) \frac{dt}{t} \right] d\sigma(y) \\
&\quad - i \lim_{\epsilon \rightarrow 0} \int_{S^{n-1}} \Omega(y) \int_{\epsilon}^{\frac{1}{\epsilon}} \sin(2\pi t y \cdot \xi) \frac{dt}{t} d\sigma(y)
\end{aligned}$$

Making the change of variables  $s = 2\pi t y \cdot \xi$  and assume  $y \cdot \xi \neq 0$ , the second term above will be

$$\begin{aligned}
\lim_{\epsilon \rightarrow 0} \int_{S^{n-1}} \Omega(y) \int_{\epsilon}^{\frac{1}{\epsilon}} \sin(2\pi t y \cdot \xi) \frac{dt}{t} d\sigma(y) &= \int_{S^{n-1}} \Omega(y) \int_0^{\infty} \text{sign}(y \cdot \xi) \frac{\sin(s)}{s} ds d\sigma(y) \\
&= \int_{S^{n-1}} \Omega(t) \frac{\pi}{2} \text{sign}(t \cdot \xi) d\sigma(t).
\end{aligned}$$

The first, after the change of variables, will be

$$\begin{aligned}
\lim_{\epsilon \rightarrow 0} \int_{S^{n-1}} \Omega(y) \left[ \int_{\epsilon}^1 (\cos(2\pi y \cdot \xi) - 1) \frac{dt}{t} + \int_1^{\frac{1}{\epsilon}} \cos(2\pi y \cdot \xi) \frac{dt}{t} \right] d\sigma(y) &= \\
= \int_{S^{n-1}} \Omega(y) \left[ \int_{2\pi|y \cdot \xi| \epsilon}^1 (\cos(s) - 1) \frac{ds}{s} + \int_1^{2\pi|y \cdot \xi| \frac{1}{\epsilon}} \cos(s) \frac{ds}{s} - \int_1^{2\pi|y \cdot \xi|} \frac{ds}{s} \right] d\sigma(y) &= \\
= \int_{S^{n-1}} \Omega(t) \left[ \log\left(\frac{1}{|t \cdot \xi|}\right) \right] d\sigma(t) &
\end{aligned}$$

since  $\Omega$  has zero average. Now applying the Plancherel yields

$$\|T(f)\|_2 = \|(\widehat{T(f)})\|_2 = \|\widehat{Kf}\|_2 \leq C\|\widehat{f}\|_2 = C\|f\|_2.$$

**Step 2:** We show  $T$  is of weak type  $(1, 1)$ : There is a constant  $C$  such that

$$|\{x \in R^n : |T(f)(x) > \lambda\}| \leq C \frac{\|f\|_1}{\lambda} \quad (1.25)$$

for any  $\lambda > 0$  and  $f \in L^1 \cap L^2$ . To do this, we need the following Calderon-Zygmund decomposition.

**Calderon-Zygmund decomposition 1.26:** Given  $f \in L^1$  and non-negative, and given a positive  $\lambda$ , there exists a sequence  $\{Q_j\}$  of disjoint cubes such that

$$f(x) \leq \lambda \quad (1.27)$$

for  $x \notin \cup Q_j$ ;

$$|\cup Q_j| \leq \frac{1}{\lambda} \|f\|_1; \quad (1.28)$$

$$\lambda < \frac{1}{|Q_j|} \int_{Q_j} f(x) dx \leq 2^n \lambda. \quad (1.29)$$

The proof of this decomposition is to use the so-called stopping time argument. First, choose a large cube  $Q$  so that  $\frac{1}{|Q|} \int_Q f(x) dx \leq \frac{1}{|Q|} \int f(x) dx \leq \lambda$ . Then divide  $Q$  to  $2^n$  equal subcubes  $Q'$ . Now we use the stopping time argument as follows: if  $\frac{1}{|Q'|} \int_{Q'} f(x) dx > \lambda$ , we

keep this subcube  $Q'$ . Otherwise, divide this subcube as above and keep this procedure. Now we get a sequence  $\{Q_j\}$ . If  $x \notin \cup Q_j$ , this means that there is a sequence  $\{Q_n\}$  with  $|Q_n| \rightarrow 0$  as  $n \rightarrow \infty$ , so that  $x \in Q_n$  for all  $n$  and  $\frac{1}{|Q_n|} \int_{Q_n} f(x) dx \leq \lambda$  which shows

(1.27). To see (1.29), notice that if  $\lambda < \frac{1}{|Q_j|} \int_{Q_j} f(x) dx$  then  $\lambda \geq \frac{1}{|2^n Q_j|} \int_{2^n Q_j} f(x) dx$  which yields (1.29). Finally, from  $\lambda < \frac{1}{|Q_j|} \int_{Q_j} f(x) dx$  we obtain  $|Q_j| < \frac{1}{\lambda} \int_{Q_j} f(x) dx$ . Summing up

shows

$$|\cup Q_j| \leq \frac{1}{\lambda} \sum_j \int_{Q_j} f(x) dx \leq \frac{1}{\lambda} \|f\|_1.$$

Now we apply the Calderon-Zygmund decomposition to show  $T$  is of the weak type  $(1, 1)$ . Define  $g(x) = f(x)$  for  $x \notin \cup Q_j$  and  $g(x) = \frac{1}{|Q_j|} \int_{Q_j} f(x) dx$  for  $x \in Q_j$ , and  $b(x) =$

$f(x) - g(x)$ . Since  $T$  is bounded on  $L^2$ , so

$$\begin{aligned} |\{x \in R^n : T(g)(x) > \frac{\lambda}{2}\}| &\leq C \frac{\|T(g)\|_2^2}{\lambda^2} \leq C \frac{\|g\|_2^2}{\lambda^2} \leq C \frac{1}{\lambda^2} \left[ \int_{(\cup Q_j)^c} |f(x)|^2 dx \right. \\ &+ \left. \sum_j \int_{Q_j} \frac{1}{|Q_j|} \int_{Q_j} f(y) dy dx \right] \leq C \frac{1}{\lambda} \left[ \int_{(\cup Q_j)^c} |f(x)| dx + 2^n |\cup Q_j| \right] \leq C \frac{1}{\lambda} \|f\|_1. \end{aligned}$$

To estimate  $T(b)(x)$ , it is easy to see

$$|\{x \in R^n : |T(b)(x)| > \frac{\lambda}{2}\}| \leq |\cup 2Q_j| + |\{x \in (\cup 2Q_j)^c : |T(b)(x)| > \frac{\lambda}{2}\}|.$$

By (1.28), it suffices to show  $|\{x \in (\cup 2Q_j)^c : |T(b)(x)| > \frac{\lambda}{2}\}| \leq C \frac{1}{\lambda} \|f\|_1$ . Rewrite  $b(x) = \sum_j b_j(x)$ , where  $b_j(x) = [f(x) - \frac{1}{|Q_j|} \int_{Q_j} f(x) dx] \chi_{Q_j}(x)$  and the sum converges in  $L^2$ , as well as pointwise. If  $x \in (\cup 2Q_j)^c$ , then, by the fact that  $b_j$  has zero average,

$$T(b_j)(x) = \int K(x-y) b_j(y) dy = \int [K(x-y) - K(x-x_{Q_j})] b_j(y) dy$$

where  $x_{Q_j}$  is the center of  $Q_j$ . Thus, by the fact that  $x \in (\cup 2Q_j)^c$  and the smoothness of  $K$ ,

$$|T(b_j)(x)| \leq C \int_{Q_j} \frac{|Q_j|}{|x-x_{Q_j}|^{n+1}} |b_j(y)| dy$$

We now estimate the integral

$$\int_{(\cup 2Q_j)^c} |T(b)(x)| dx \leq C \sum_j \int_{(\cup 2Q_j)^c} \int_{Q_j} \frac{|Q_j|}{|x-x_{Q_j}|^{n+1}} |b_j(y)| dy dx.$$

We apply Fubini's theorem to change the order of integration in each of the double integrals and it gives  $C \sum_j \|b_j\|_1$ . This, in turn, can be estimated by using (1.28) and (1.29). So we get

$$\int_{(\cup 2Q_j)^c} |T(b)(x)| dx \leq C \|f\|_1$$

which shows  $|\{x \in (\cup 2Q_j)^c : |T(b)(x)| > \frac{\lambda}{2}\}| \leq C \frac{\|f\|_1}{\lambda}$ . This fact together with (1.28) shows

$$|\{x \in R^n : |T(b)(x)| > \frac{\lambda}{2}\}| \leq C \frac{\|f\|_1}{\lambda}.$$

Noting  $|\{x \in R^n : |T(f)(x)| > \lambda\}| \leq |\{x \in R^n : |T(g)(x)| > \frac{\lambda}{2}\}| + |\{x \in R^n : |T(b)(x)| > \frac{\lambda}{2}\}|$  yields the proof of (1.25) and the proof of step 2 is complete.

**Step 3:** By Marcinkiewicz's interpolation theorem,  $T$  is bounded on  $L^p$  for  $1 < p \leq 2$ , and

$$\|T(f)\|_p \leq C \|f\|_p.$$

**Step 4:** We observe that the adjoint  $T^*$  of  $T$  satisfies the same conditions as  $T$ , so  $T^*$  is bounded on  $L^p$ ,  $1 < p \leq 2$ . The duality argument shows  $T$  is bounded on  $L^p$ ,  $2 \leq p < \infty$ .

**Example:** Let  $f(x, y)$  be the density of a mass distribution in the plane. Then its Newtonian potential in the half-space  $R_+^3$  is

$$g(x, y, z) = \int_{R^2} \frac{f(u, v)}{[(x-u)^2 + (y-v)^2 + z^2]^{\frac{1}{2}}} dudv.$$

Formally, we have

$$\lim_{z \rightarrow 0} \frac{\partial g}{\partial x}(x, y, z) = - \int_{R^2} \frac{f(u, v)(x-u)}{[(x-u)^2 + (y-v)^2]^{\frac{3}{2}}} dudv.$$

However, this integral does not converge in general. But it exists as a principle value if  $f$  is smooth. This is an operator of the first generation of Calderon-Zygmund operators by considering  $\Omega(x, y) = -\frac{x}{x^2+y^2}$ .

We can consider the first generation of Calderon-Zygmund operators as the Fourier multipliers.

**Theorem 1.30:** If  $m \in C^\infty(R^n \setminus \{0\})$  is a homogeneous function of degree 0, and  $T_m$  is the Fourier multiplier defined by  $(T_m f) = m\hat{f}$ , then there exist  $a$ , a complex number and  $\Omega \in C^\infty(S^{n-1})$  with zero average such that for any  $f \in \mathcal{S}$ ,

$$T_m f = af + p.v. \int \frac{\Omega(y)}{|y|^n} f(x-y) dy. \quad (1.31)$$

Since any homogeneous function of degree 0 is the sum of a constant and a homogeneous function of degree 0 with zero average on  $S^{n-1}$ , theorem 1.30 is a consequence of the following lemma.

**Lemma 1.32:** Let  $m \in C^\infty(R^n \setminus \{0\})$  be a homogeneous function of degree 0, and  $T_m$  is the Fourier multiplier defined by  $(T_m f) = m\hat{f}$ , then there exist  $a$ , a complex number and  $\Omega \in C^\infty(S^{n-1})$  with zero average such that  $\hat{m}(y) = p.v. \frac{\Omega(y)}{|y|^n}$ .

**Proof:** Since  $m$  is a tempered distribution,  $\hat{m}$  exists. Thus,

$$\left(\frac{\partial^n m}{\partial x_i^n}\right)(\xi) = C\xi_i^n \hat{m}(\xi),$$

where  $C$  is a constant. The function  $\frac{\partial^n m}{\partial x_i^n}$  is homogeneous of degree -n, in  $C^\infty(R^n \setminus \{0\})$  and has zero average on  $S^{n-1}$ . Moreover,

$$\frac{\partial^n m}{\partial x_i^n} = p.v. \frac{\partial^n m}{\partial x_i^n} + \sum_{|\alpha| \leq k} C_\alpha D^\alpha \delta, \quad (1.33)$$

where  $\delta$  is the Dirac measure at the origin, since the difference between  $\frac{\partial^n m}{\partial x_i^n}$  and p.v.  $\frac{\partial^n m}{\partial x_i^n}$  is a distribution supported at the origin. By taking the Fourier transform on the

both sizes of (1.33) and note that the left-hand side and the first term on the right-hand side are homogeneous distributions of degree 0, so the polynomial, the Fourier transform of the second term on the right-hand side, is a constant. Thus the right-hand side of (1.33) is a homogeneous function of degree 0 which is in  $C^\infty(R^n \setminus \{0\})$ . Since this valid for  $1 \leq i \leq n$ ,  $\hat{m}$  coincides on  $R^n \setminus \{0\}$  with a homogeneous function of degree -n. We denote its restriction to  $S^{n-1}$  by  $\Omega$ . To see that  $\Omega$  has zero average, fix a radial function  $\phi \in \mathcal{S}$ , which is supported on  $1 \leq |x| \leq 2$  and positive for  $1 < |x| < 2$ . Then

$$\hat{m}(\phi) = \int \frac{\Omega(x)}{|x|^n} \phi(x) dx = c \int \Omega(x) d\sigma(x), \quad (1.34)$$

where  $c > 0$ . On the other hand, since  $\hat{\phi}$  is radial and  $m$  is homogeneous,

$$\hat{m}(\phi) = m(\hat{\phi}) = c' \int m(x) d\sigma(x) = 0$$

which together with (1.34) shows  $\int \Omega(x) d\sigma(x) = 0$ . Finally, to see that  $\hat{m}$  is identical to p.v.  $\frac{\Omega(x)}{|x|^n}$ , consider their difference  $\hat{m} - \text{p.v.} \frac{\Omega(x)}{|x|^n}$ , which is supported at the origin. By taking the Fourier transform of this difference, we get a polynomial which must be a constant because both  $m$  and  $(\text{p.v.} \frac{\Omega(x)}{|x|^n})$  are bounded. Furthermore, this constant must be zero since both  $m$  and  $\Omega$  have zero average on  $S^{n-1}$ .

**Theorem 1.35:** The set  $\mathcal{A}$  of operators defined by theorem 1.30 is a commutative algebra. An element of  $\mathcal{A}$  is invertible if and only if  $m$  is never zero on  $S^{n-1}$ .

**Proof:** If  $T_{m_1}$  and  $T_{m_2}$  are in  $\mathcal{A}$ , then  $T_{m_1} T_{m_2} = T_{m_1 m_2} \cdot T_m$  is invertible if and only if  $T_{\frac{1}{m}} \in \mathcal{A}$ , and hence,  $1/m \in C^\infty(R^n \setminus \{0\})$  which shows  $m(\xi) \neq 0$  for any  $\xi \in S^{n-1}$ . We now return to study  $Lu(x) = \sum_{|\alpha|=m} a_\alpha \frac{\partial^\alpha u}{\partial x^\alpha}$ . Define

$$(\widehat{\Lambda f})(\xi) = 2\pi |\xi| \hat{f}(\xi). \quad (1.36)$$

Then

$$Lu = T \Lambda^m u, \quad (1.37)$$

where the operator  $T$  is defined by

$$(\widehat{Tu})(\xi) = i^m \frac{P(\xi)}{|\xi|^m} \hat{u}(\xi) \quad (1.38)$$

where  $P(\xi) = \sum_{|\alpha|=m} a_\alpha (2\pi i \xi)^\alpha$ . The multiplier  $\frac{P(\xi)}{|\xi|^m}$  is homogeneous of degree zero and it is a  $C^\infty$  function on  $S^{n-1}$ , and hence,  $T$  is an operator in  $\mathcal{A}$ . To solve

$$Lu(x) = \sum_{|\alpha|=m} a_\alpha \frac{\partial^\alpha u}{\partial x^\alpha} = f,$$

it suffices to solve  $(-\Delta)^{\frac{m}{2}} u = T^{-1} f$  because  $\Lambda = (-\Delta)^{\frac{1}{2}}$ .

## Lecture 2. Differential operators with variable coefficients, the second generation of Calderon-Zygmund operators, and pseudo-differential operators

We are interested in studying the following differential operator with variable coefficients:

$$Lu(x) = \sum_{|\alpha|=m} a_\alpha(x) \frac{\partial^\alpha u}{\partial x^\alpha} \quad (2.1)$$

where  $a_\alpha(x) \in C^\infty(\mathbb{R}^n)$ .

Formally, by taking the Fourier transform and then the inverse Fourier transform, we get

$$\begin{aligned} Lu(x) &= \sum_{|\alpha|=m} a_\alpha(x) \int (-2\pi i \xi)^\alpha \hat{u}(\xi) e^{2\pi i \xi \cdot x} d\xi \\ &= \int \left[ \sum_{|\alpha|=m} a_\alpha(x) (-2\pi i \xi)^\alpha \right] \hat{u}(\xi) e^{2\pi i \xi \cdot x} d\xi. \end{aligned} \quad (2.2)$$

(2.2) can be written by more general form:

$$\sigma(x, D)f(x) = \int \sigma(x, \xi) \hat{f}(\xi) e^{2\pi i \xi \cdot x} d\xi. \quad (2.3)$$

Calderon and Zygmund wanted to keep what they did for the first generation of Calderon-Zygmund operators and rewrite (2.3) by

$$Tf(x) = \int L(x, x-y) f(y) dy \quad (2.4)$$

where

$$p.v. \int L(x, y) e^{-2\pi i \xi \cdot y} dy = \sigma(x, \xi). \quad (2.5)$$

The relationship (2.5), discovered by Calderon and Zygmund in the 1950s, opened the way to all later developments in which the pseudo-differential operators were defined using algebras of symbols, without reference to any kernels. After the golden age just described, the two points of view diverged: Kohn and Nirenberg, for their part, and Hormander, for his, systematically favored the definition of pseudo-differential operators by symbols. Research on kernels remained very active in the school of Calderon and Zygmund and led to what we will introduce the third generation of Calderon-Zygmund operators in next lecture.

To see (2.4) and (2.5), formally, by taking the inverse Fourier transform, we obtain

$$\begin{aligned} Tf(x) &= \int \sigma(x, \xi) \hat{f}(\xi) e^{2\pi i \xi \cdot x} d\xi = \int \int \sigma(x, \xi) e^{2\pi i \xi \cdot x} f(y) e^{-2\pi i y \cdot \xi} dy d\xi \\ &= \int \left[ \int \sigma(x, \xi) e^{2\pi i \xi \cdot (x-y)} d\xi \right] f(y) dy = \int L(x, x-y) f(y) dy. \end{aligned}$$

Using the operator  $\Lambda$  again, Calderon and Zygmund rewrote (2.2) as

$$\begin{aligned} Lu(x) &= \int \left( \sum_{|\alpha|=m} a_\alpha(x) (-2\pi i \xi)^\alpha \frac{1}{|\xi|^m} \right) |\xi|^m \hat{u}(\xi) e^{2\pi i \xi \cdot x} d\xi \\ &= T(\Lambda^m u) \end{aligned} \quad (2.6)$$

where  $T$  is defined by

$$Tf(x) = \int \sigma(x, \xi) \hat{f}(\xi) e^{2\pi i \xi \cdot x} d\xi \quad (2.7)$$

with  $\sigma(x, \xi) = \frac{\sum_{|\alpha|=m} a_\alpha(x) (-2\pi i \xi)^\alpha}{|\xi|^m}$ .

Thus,  $\sigma$  is homogeneous of degree 0 in the variable  $\xi$ . By (2.5),

$$Tf(x) = \int K(x, x-y) f(y) dy \quad (2.8)$$

where  $K(x, y) = \int \sigma(x, \xi) e^{2\pi i y \cdot \xi} d\xi$ .

For fixed  $x$ ,  $K(x, \cdot)$  is the inverse Fourier transform of  $\sigma(x, \cdot)$ . By theorem 1.30, for each fixed  $x$  there is a constant  $a(x)$  and a function  $\Omega(x, \cdot) \in C^\infty(S^{n-1})$  with zero average on  $S^{n-1}$  such that

$$K(x, y) = a(x) \delta(y) + p.v. \frac{\Omega(x, y)}{|y|^n}.$$

This was the motivation for Calderon and Zygmund to introduce the second generation of Calderon-Zygmund operators.

**Theorem 2.9:** Suppose that  $\Omega(x, y)$  is a function satisfying the following conditions:

$$\Omega(x, y) = \Omega(x, \lambda y) \quad (2.10)$$

for all  $x \in R^n$  and all  $\lambda > 0$ ;

$$\Omega(x, y) \in C^\infty(R^n \times S^{n-1}); \quad (2.11)$$

$$\int_{S^{n-1}} \Omega(x, y) d\sigma(y) = 0 \quad (2.12)$$

for all  $x \in R^n$ .

The second generation of Calderon-Zygmund operators is the set of  $T$  defined by

$$Tf(x) = p.v. \int \frac{\Omega(x, y)}{|y|^n} f(x-y) dy. \quad (2.13)$$

Then the operator  $T$  given by (2.13) is bounded on  $L^p$ ,  $1 < p < \infty$ .

To show this theorem, by the Calderon-Zygmund real variable method, it suffices to prove the  $L^2$  boundedness of  $T$ . To see this, we follow Calderon and Zygmund, by

performing a spherical harmonic expansion of  $\sigma(x, \xi)$  on the  $S^{n-1}$  for fixed  $x \in R^n$ , where  $\sigma(x, \xi)$  is the Fourier transform of p.v.  $\frac{\Omega(x, y)}{|y|^n}$  with the variable  $y$  whenever  $x$  is fixed. The proof of lemma 1.32 shows that  $\sigma$  satisfies (2.10), (2.11) and (2.13) with  $\Omega$  replaced by  $\sigma$ . By the regularity with respect to  $\xi$ , this gives a norm-convergent sequence

$$\sigma(x, \xi) = \sum_0^{\infty} m_k(x) h_k(\xi)$$

with  $\sum_0^{\infty} \|m_k\|_{\infty} \|h_k\|_{\infty} < \infty$ .

We extend  $h_k(\xi)$  as a homogeneous function of degree 0, which is the symbol of an operator we denote by  $H_k$ , and we write  $M_k$  for the operator of pointwise multiplication by  $m_k$ . This yields

$$T = \sum_0^{\infty} M_k H_k$$

and the series of operators is convergent in the  $L^2$  - norm. It is easy to see that the second generation of Calderon-Zygmund operators includes the first generation of Calderon-Zygmund operators.

Now we consider formally a pseudo-differential operator defined by

$$T_{\sigma} f(x) = \int \sigma(x, \xi) \hat{f}(\xi) e^{2\pi i x \cdot \xi} d\xi.$$

If  $\sigma$  is independent of  $\xi$ ,  $\sigma(x, \xi) = a(x)$ , then  $T$  is a multiplication operator:  $Tf(x) = a(x)f(x)$ . When  $\sigma$  is independent of  $x$ ,  $\sigma(x, \xi) = m(\xi)$ , then  $T$  is a Fourier multiplier operator:  $(Tf)(\xi) = m(\xi)\hat{f}(\xi)$  which shows that pseudo-differential operators are generalizations of the Fourier multiplier operators. We shall consider the standard symbol class, denoted by  $S^m$ , which is most common and useful of the general symbol classes.

**Definition 2.16:** A function  $\sigma$  belongs to  $S^m$  (and is said to be of order  $m$ ) if  $\sigma(x, \xi)$  is a  $C^{\infty}$  function of  $(x, \xi) \in R^n \times R^n$  and satisfies the differential inequalities

$$|\partial_x^{\beta} \partial_{\xi}^{\alpha} \sigma(x, \xi)| \leq C_{\alpha, \beta} (1 + |\xi|)^{m - |\alpha|}, \quad (2.17)$$

for all multi-indices  $\alpha$  and  $\beta$ .

It is easy to see that if  $\sigma$  is a polynomial in  $\xi$  and independent of  $x$ , then (2.17) is satisfied. Roughly speaking, the conditions (2.17) mean that the behavior of  $\sigma(x, \xi)$  looks like a polynomial of order  $m$ .

Given a symbol in  $S^m$ , the operator  $T_{\sigma}$  will initially be defined on the Schwartz class of testing functions  $\mathcal{S}$ . In fact, the integral (2.15) converges absolutely and is infinitely differentiable. An integration by parts argument shows that  $T_{\sigma}(f)$  is a rapidly decreasing function. Indeed, note that

$$(I - \Delta_{\xi}) e^{2\pi i x \cdot \xi} = (1 + 4\pi^2 |x|^2) e^{2\pi i x \cdot \xi},$$

and define the operator

$$L_\xi = (1 + 4\pi^2|x|^2)^{-1}(I - \Delta_\xi),$$

then  $(L_\xi)^N e^{2\pi i x \cdot \xi} = e^{2\pi i x \cdot \xi}$ . Inserting this in (2.15) and carrying out the repeated integrations by parts gives

$$T_\sigma f(x) = \int (L_\xi)^N [\sigma(x, \xi) \hat{f}(\xi)] e^{2\pi i x \cdot \xi} d\xi$$

which shows  $T_\sigma f(x)$  is rapidly decreasing. Since this argument works for any partial derivative of  $T_\sigma$ , and, hence,  $T_\sigma$  maps  $\mathcal{S}$  to  $\mathcal{S}$ , and this mapping is continuous. It is worth pointing out that if  $\{\sigma_k\}$  is a pointwise convergent sequence of symbols in  $S^m$  that satisfy the conditions (2.17) uniformly in  $k$ , then  $T_{\sigma_k}(f) \rightarrow T_\sigma(f)$  in  $\mathcal{S}$  for  $f \in \mathcal{S}$ . An alternative way of writing  $T_\sigma$  defined in (2.15) is as a repeated integral

$$T_\sigma f(x) = \int \int \sigma(x, \xi) e^{2\pi i(x-y) \cdot \xi} f(y) dy d\xi. \quad (2.18)$$

However, the integral in (2.18) does not necessarily converge absolutely, even when  $f \in \mathcal{S}$ . To with with this integral, fix a function  $\gamma \in C_0^\infty(R^n \times R^n)$  with  $\gamma(0, 0) = 1$ . Set  $\sigma_\epsilon(x, \xi) = \sigma(x, \xi) \gamma(\epsilon x, \epsilon \xi)$ . Notice that if  $\sigma \in S^m$ , then  $\sigma_\epsilon \in S^m$  and they satisfy the condition (2.17) uniformly in  $\epsilon$ , for  $0 < \epsilon, 1$ . As mentioned above,  $T_{\sigma_\epsilon}(f) \rightarrow T_\sigma(f)$  in  $\mathcal{S}$  when  $f \in \mathcal{S}$ , as  $\epsilon \rightarrow 0$ . Moreover, since the operator  $T_\sigma$  defined in (2.18) converges when  $\sigma$  has compact support, we get that

$$T_\sigma f(x) = \lim_{\epsilon \rightarrow 0} \int \int \sigma_\epsilon(x, \xi) e^{2\pi i(x-y) \cdot \xi} f(y) dy d\xi. \quad (2.19)$$

By the duality relation  $\langle T_\sigma f, g \rangle = \langle f, T_\sigma^* g \rangle$  when  $f, g \in \mathcal{S}$ . The same proof shows  $T_\sigma^*$  also maps  $\mathcal{S}$  to  $\mathcal{S}$ .

**Theorem 2.20:** If  $\sigma \in S^0$ , then the operator  $T_\sigma$  initially defined on  $\mathcal{S}$ , extends to a bounded operator on  $L^2$ .

**Proof:** It suffices to show that

$$\|T_\sigma f\|_2 \leq C \|f\|_2, \quad (2.21)$$

whenever  $f \in \mathcal{S}$ , with  $C$  independent of  $f$ .

Indeed, suppose  $f \in L^2$  and let  $\{f_n\} \in \mathcal{S}$  so that  $f_n \rightarrow f$  in  $L^2$ . Then, by (2.21),  $T_\sigma(f_n)$  converges in  $L^2$  norm, and hence,  $T_\sigma(f_n)$  converges to  $T_\sigma(f)$  in the sense of distributions.

We return to the proof of (2.21). First, we assume that  $\sigma(x, \xi)$  has compact support in  $x$ . We then write

$$\sigma(x, \xi) = \int \hat{\sigma}(\mu, \xi) e^{2\pi i \mu \cdot x} d\mu$$

since  $\sigma$  has compact support in  $x$  variable. An integration by parts shows for each multi-index  $\alpha$ ,

$$(2\pi i \mu)^\alpha \hat{\sigma}(\mu, \xi) = \int [\partial_x^\alpha \sigma(x, \xi)] e^{-2\pi i \mu \cdot x} dx$$

and

$$|(2\pi i\mu)^\alpha \widehat{\sigma}(\mu, \xi)| \leq c_\alpha,$$

uniformly in  $\xi$ . As a result, we obtain

$$\sup_\xi |\widehat{\sigma}(\mu, \xi)| \leq C_N (1 + |\mu|)^{-N}$$

for arbitrary  $N \geq 0$ . Now

$$\begin{aligned} T_\sigma f(x) &= \int \sigma(x, \xi) \widehat{f}(\xi) e^{2\pi i x \cdot \xi} d\xi = \int \int \widehat{\sigma}(\mu, \xi) e^{2\pi i \mu \cdot x} \widehat{f}(\xi) e^{2\pi i x \cdot \xi} d\xi d\mu \\ &= \int (T^\mu f)(x) d\mu, \end{aligned}$$

where

$$(T^\mu f)(x) = e^{2\pi i x \cdot \mu} (T_{\widehat{\sigma}(\mu, \cdot)} f)(x).$$

Since for each  $\mu$ ,  $T_{\widehat{\sigma}(\mu, \cdot)}$  is a Fourier multiplier operator on the Fourier side, by Plancherel's theorem we have that

$$\|T_{\widehat{\sigma}(\mu, \cdot)} f\|_2 \leq \sup_\xi |\widehat{\sigma}(\mu, \xi)| \|\widehat{f}\|_2 = \sup_\xi |\widehat{\sigma}(\mu, \xi)| \|f\|_2 \quad (2.22).$$

By (2.22),  $\|T^\mu\| \leq C_N (1 + |\mu|)^{-N}$ , which yields

$$\|T_\sigma\| \leq C_N \int (1 + |\mu|)^{-N} d\mu < \infty$$

if we choose  $N > n$ . Thus (2.21) is proved when  $\sigma$  has compact support in  $x$ .

The proof for general symbols needs to use the singular integral realization of the operator  $T_\sigma$ . That is, we shall write

$$T_\sigma f(x) = \int k(x, z) f(x - z) dz \quad (2.23)$$

where for each  $x$ ,  $k(x, \cdot)$  is the distribution whose Fourier transform is the function  $\sigma(x, \cdot)$ , formally,

$$\sigma(x, \xi) = \int k(x, z) e^{-2\pi i z \cdot \xi} dz.$$

Thus,  $T_\sigma$  can be interpreted as the convolution of the distribution  $k(x, \cdot)$  with the function  $f \in \mathcal{S}$ , evaluated at the point  $x$ . We first have the following estimate on  $k(x, z)$ .

$$|k(x, z)| \leq C_N |z|^{-N}$$

for all  $|z| \geq 1$  and all  $N > 0$ , uniformly in  $x$ .

To see this, note that  $T_\sigma(f)(x)$  equals  $[k(x, \cdot) * f](x)$ , where  $k(x, \cdot)$  is the distribution whose Fourier transform is the function  $\sigma(x, \cdot)$  as we have written in (2.23). Next,

$(-2\pi iz)^\alpha k(x, \cdot)$ , with the distribution  $k(x, \cdot)$  thought of as acting on functions of  $z$ , equals the inverse Fourier transform of  $\partial_\xi^\alpha \sigma(x, \xi)$ ; by (2.17),  $\partial_\xi^\alpha \sigma(x, \xi)$  is integrable in  $\xi$  whenever  $|\alpha| \geq n + 1$ . This shows that  $k(x, \cdot)$  equals a function away from the origin, and that  $|z|^N |k(x, z)| \leq C_N$  for  $N > n$ , and from this, (2.24) follows.

We return to the proof of theorem 2.20, without assuming that  $\sigma(x, \xi)$  has compact support in  $x$ . To begin with, we will show that, for each  $x_0 \in R^n$ ,

$$\int_{|x-x_0| \leq 1} |T_\sigma f(x)|^2 dx \leq C_N \int_{R^n} \frac{|f(x)|^2}{(1+|x-x_0|)^N} dx \quad (2.25)$$

for all  $N \geq 0$ .

We prove (2.25) first when  $x_0 = 0$ . To do this, we split  $f$  by  $f = f_1 + f_2$ , with  $f_1$  supported in  $B(0, 3)$ ,  $f_2$  supported outside  $B(0, 2)$ ,  $f_1$  and  $f_2$  smooth, and with  $|f_1|, |f_2| \leq |f|$ . We fix  $\eta \in C_0^\infty$  so that  $\eta = 1$  in  $B(0, 1)$ . Then  $\eta T_\sigma(f_1) = T_{\eta\sigma}(f_1)$ , and the symbol  $\eta(x)\sigma(x, \xi)$  has compact support in  $x$ , so the previous result applies. Hence

$$\int_{B(0,1)} |T_\sigma f_1|^2 \leq \int_{R^n} |T_{\eta\sigma} f_1|^2 \leq C \int_{R^n} |f_1|^2 \leq C \int_{R^n} |f|^2 \quad (2.26).$$

If  $x \in B(0, 1)$ , since  $f_2$  is supported away from  $B(0, 2)$ , the representation (2.23) holds:

$$T_\sigma f_2(x) = \int_{(B(0,2))^c} k(x, x-z) f_2(z) dz. \quad (2.27)$$

Since  $|x-z| \geq 1$  when  $x \in B(0, 1)$  and  $z \notin B(0, 2)$ , using the estimate (2.24) yields

$$|T_\sigma f_2(x)| \leq C_N \int_{(B(0,2))^c} |z|^{-N} |f(z)| dz.$$

By Schwartz inequality, we obtain, for  $N > n$ ,

$$\int_{B(0,1)} |T_\sigma f_2(x)|^2 dx \leq C_N \int (1+|x|)^{-N} |f(x)|^2 dx. \quad (2.28)$$

Now  $T_\sigma(f) = T_\sigma(f_1) + T_\sigma(f_2)$ , and combining (2.26) and (2.28) shows (2.25) when  $x_0 = 0$ . The passage to (2.25) for general  $x_0$  can be achieved by noting that, while an individual pseudo-differential operator is not (in general) translation-invariant, the class  $S^m$ , in fact, is. To see this, let  $\tau_h, h \in R^n$ , denote the unitary translation operator given by  $(\tau_h f)(x) = f(x-h)$ . Then,  $\tau_h T_\sigma \tau_{-h} = T_{\sigma_h}$ , where  $\sigma_h(x, \xi) = \sigma(x-h, \xi)$ . Note that the symbol  $\sigma_h$  satisfies the same estimates that  $\sigma$  does, uniformly in  $h$ . Hence, (2.25) holds for  $x_0 = 0$ , with  $\sigma$  replaced by  $\sigma_h$ , with a bound independent of  $h$ . If we set  $h = x_0$ , we see that (2.25) is established. Finally, it is only a matter of integrating (2.25) with respect to  $x_0$ , choosing  $N > n$ , and interchanging order of integration. Theorem is proved.

By combining the  $L^2$  result with the Calderon-Zygmund real variable method, we can also prove the  $L^p$  boundedness of these operators. However, we need to realize pseudo-differential operators in the class  $S^0$  as singular integrals. To be precise, we shall prove the following result.

**Proposition 2.29:** Suppose  $\sigma \in S^m$ . Then the kernel  $k(x, z)$  is in  $C^\infty(R^n \times (R^n \setminus \{0\}))$ , and satisfies

$$|\partial_x^\beta \partial_z^\alpha k(x, z)| \leq C_{\alpha, \beta, N} |z|^{-n-m-|\alpha|-N}, z \neq 0, \quad (2.30)$$

for all multi-indices  $\alpha$  and  $\beta$ , and all  $N \geq 0$  so that  $n + m + |\alpha| + N > 0$ .

**Proof:** The proof uses the so-called dyadic decomposition. We begin by fixing  $\eta \in C_0^\infty$  with the properties that  $\eta(\xi) = 1$  for  $|\xi| \leq 1$ , and  $\eta(\xi) = 0$  for  $|\xi| \geq 2$ . We also define another function  $\delta$ , by  $\delta(\xi) = \eta(\xi) - \eta(2\xi)$ . Then we have the following two partitions of unity of the  $\xi$  - space :

$$1 = \eta(\xi) + \sum_{j=1}^{\infty} \delta(2^{-j}\xi), \quad (2.31)$$

for all  $\xi$ , and

$$1 = \sum_{j=-\infty}^{\infty} \delta(2^{-j}\xi), \quad (2.32)$$

for all  $\xi \neq 0$ .

It is worth to pointing out that for each  $\xi$  there are at most two nonzero terms in the sums (2.31) and (2.32).

Let  $\phi$  be the inverse Fourier transform of  $\eta$ , i.e.,  $\widehat{\phi}(\xi) = \eta(\xi)$ . Then  $\phi \in \mathcal{S}$  and  $\int \phi dx = 1$ . Define  $\psi$  by  $\widehat{\psi}(\xi) = \delta(\xi)$ . Then  $\int \psi dx = 0$ . Writing  $\phi_t(x) = t^{-n} \phi(\frac{x}{t})$  and  $\psi_j = \psi_{2^{-j}}$ , we then have  $\psi_j = \phi_j - \phi_{j-1}$ , while  $\widehat{(\phi_j)}(\xi) = \eta(2^{-j}\xi)$  and  $\widehat{(\psi_j)}(\xi) = \delta(2^{-j}\xi)$ . We now define the operator  $S_j$  by  $S_j(f) = f * \phi_j$  and  $\Delta_j(f) = S_j(f) - S_{j-1}(f) = f * \psi_j$ . In parallel with (2.31) and (2.32), we have the operator identities

$$I = S_0 + \sum_{j=1}^{\infty} \Delta_j \quad (2.33)$$

and

$$I = \sum_{j=-\infty}^{\infty} \Delta_j. \quad (2.34)$$

Note that if  $f$  is a tempered distribution,  $S_j(f)$  is well defined, and

$$S_0(f) + \sum_{j=1}^N \Delta_j(f) = S_N(f) \rightarrow f,$$

as  $N \rightarrow \infty$ , in the sense of distribution. However, it is not true that  $S_M(f) \rightarrow 0$  as  $M \rightarrow -\infty$ , for arbitrary  $f$ . It fails when  $f = 1$ . Thus (2.34) holds only under some restriction on  $f$ .

We now return to the operator  $T_\sigma$ . Using (2.33), we write

$$T_\sigma = TS_0 + \sum_{j=1}^{\infty} T\Delta_j = \sum_{j=0}^{\infty} T_{\sigma_j},$$

where  $\sigma_0(x, \xi) = \sigma(x, \xi)\eta(\xi)$ , and  $\sigma_j(x, \xi) = \sigma(x, \xi)\delta(2^{-j}\xi)$  for  $j \geq 1$ .

Each of the pseudo-differential operators  $T_{\sigma_j}$  will be written in its singular integral form by

$$T_{\sigma_j}f(x) = \int k_j(x, z)f(x-z)dz.$$

Since  $\sigma_j$  have compact  $\xi$ -support and are smooth, the kernel  $k_j$  will also be smooth, and the integrals above will converge for all  $x$ . The kernels  $k_j$  are given by

$$k_j(x, z) = \int \sigma_j(x, \xi)e^{2\pi i\xi \cdot z}d\xi.$$

We claim the following estimate: If  $\sigma \in S^m$ , then

$$|\partial_x^\beta \partial_z^\alpha k_j(x, z)| \leq C_{\alpha, \beta, N} |z|^{-N} 2^{j[n+m-N+|\alpha|]}, \quad (2.35)$$

for all multi-indices  $\alpha, \beta$ , and  $N \geq 0$ , where  $C_{\alpha, \beta, N}$  is independent of  $j \geq 0$ . In fact, observe that

$$(-2\pi iz)^\gamma \partial_x^\beta \partial_z^\alpha k_j(x, z) \leq \int \partial_\xi^\gamma [(2\pi i\xi)^\alpha \partial_x^\beta \sigma_j(x, \xi)] e^{2\pi i\xi \cdot z} d\xi.$$

Noting that the integrand is supported in  $2^{j-1} \leq |\xi| \leq 2^{j+1}$  and estimates on  $\sigma(x, \xi)$  and  $\delta(2^{-j}\xi)$ , we obtain that when  $|\gamma| = M$ ,

$$|z^\gamma \partial_x^\beta \partial_z^\alpha k_j(x, z)| \leq C_{\alpha, \beta, N} 2^{j[n+m-M+|\alpha|]}.$$

Taking the supremum over all  $\gamma$  with  $|\gamma| = M$ , gives (2.35). Since  $k(x, z) = \sum_{j=0}^{\infty} k_j(x, z)$ , it

suffices to show  $\sum_{j=0}^{\infty} |\partial_x^\beta \partial_z^\alpha k_j(x, z)|$  satisfies the estimate given by the right side of (2.29).

Consider the case when  $0 < |z| \leq 1$ , first. We break the above sum into two parts: the first where  $2^j \leq |z|$ , the second where  $2^j > |z|$ . For the first sum we use the estimate (2.35) with  $M = 0$ , which is then majorized by a multiple of  $\sum_{2^j \leq |z|} 2^{j(n+m+|\alpha|)}$ . This in

turn is  $O(|z|^{-n-m-|\alpha|})$  when  $n+m+|\alpha| > 0$ , or  $O(\log(|z|^{-1})+1)$  when  $n+m+|\alpha| \leq 0$ . In either case we get the estimate  $O(|z|^{-n-m-|\alpha|-M})$ , under restriction that  $|z| \leq 1, M \geq 0$  and  $n+m+|\alpha|+M \geq 0$ . Next, for the second sum, we choose  $M > n+m+|\alpha|$ . By (2.35), we get the estimate  $O(|z|^{-M}) \sum_{2^j > |z|^{-1}} 2^{j[n+m+|\alpha|-M]} = O(|z|^{-n-m-|\alpha|})$ . The last term is

$O(|z|^{-n-m-|\alpha|-N})$  if  $N \geq 0$ , since  $|z| \leq 1$ .

Finally, when  $|z| \geq 1$ , and if  $M > n + m + |\alpha| + N$ , then (2.35) shows that the sum is majorized by  $O(|z|^{-M})$ , which is  $O(|z|^{-n-m-|\alpha|-N})$  for every  $N$ , since  $|z| \geq 1$ . The proof of the proposition is therefore concluded.

Now the  $L^2$  boundeness and the estimates of kernel of  $T_\sigma$  allow us to apply the Calderon-Zygmund real variable method, and hence, the  $L^p, 1 < p < \infty$ , boundedness follows.

As mentioned before, after Calderon and Zygmund introduced the second generation of Calderon-Zygmund operators, they seek to compose the operators in order to obtain an algebra with a precise symbolic calculus. To get this, it remains to be seen whether the operators in question can be defined by symbols satisfying simple conditions of regularity and rate of growth at infinity. There are two such algebras  $\mathcal{A}_\infty$  and  $\mathcal{B}_\infty$ . The algebra  $\mathcal{A}_\infty$  is the set of all operators  $T$ , where the symbols of  $T$  and its adjoint  $T^*$  satisfy the following illicit estimates:

$$|\partial_x^\beta \partial_\xi^\alpha \sigma(x, \xi)| \leq C_{\alpha, \beta} |\xi|^{-(|\alpha| - |\beta|)}. \quad (2.36)$$

The algebra  $\mathcal{B}_\infty$  is the set of all operators  $T$ , where the symbols of  $T$  and its adjoint  $T^*$  satisfy the following illicit estimates:

$$|\partial_x^\beta \partial_\xi^\alpha \sigma(x, \xi)| \leq C_{\alpha, \beta} (1 + |\xi|)^{-(|\beta| - |\alpha|)}. \quad (2.37)$$

Several years later (1965), Calderon took another look at the problem of the symbolic calculus when he sought conditions of minimal regularity with respect to  $x$ . In applications to partial differential equations, the regularity with respect to  $x$  is given by that of the coefficients  $a_\alpha(x)$  of the differential operators  $\sum a_\alpha(x) \partial^\alpha$ . This problem reduces to the study of the commutators  $[A, T]$ , where  $A$  is the operator of pointwise multiplication by a function  $a(x)$ , and where  $T = (\frac{\partial}{\partial x_1})T_1 + \dots + (\frac{\partial}{\partial x_n})T_n$ , the  $T_j, 1 \leq j \leq n$ , being first generation Calderon-Zygmund operators.

In dimension 1,  $T$  is replaced by  $DH$ , where  $D = -i \frac{d}{dx}$  and  $H$  is the Hilbert transform. In 1965, Calderon showed that the commutator  $[A, DH]$  was bounded on  $L^2(R)$  if and only if the function  $a(x)$  was Lipschitz, that is, there is a constant  $C$  such that  $|a(x) - a(y)| \leq C|x - y|$ , for all  $x, y \in R$ . It is easy to see that this condition is necessary. The reverse implication is deep, and Calderon's proof relies on the characterization, established by Calderon for this purpose, of complex Hardy space  $H^1$  by the integrability of Lusin's area function. This remarkable result lead to the third generation Calderon-Zygmund operators which do not belong to the pseudo-differential operators. A further operator belonging to Calderon's program is related the classical method of using a double layer potential to solve the Dirichlet and the Neumann problem in a Lipschitz domain. The operator involved, similar to the commutator  $[A, DH]$ , is given in local coordinates by

$$Tf(x) = \frac{1}{\omega_n} p.v. \int K(x, y) f(y) dy$$

where

$$K(x, y) = \frac{a(x) - a(y) - (x, y) \cdot \nabla a(y)}{(|x - y|^2 + (a(x) - a(y))^2)^{(n+1)/2}},$$

and  $f \in L^2(\mathbb{R}^n)$ . The Lipschitz domain is defined (locally) by  $t > a(x)$ , where  $x \in \mathbb{R}^n, t \in \mathbb{R}$ , and the function  $a(x)$  is Lipschitz.

If  $n = 1$ , this kernel is precisely the real part of the kernel of the Cauchy integral on a Lipschitz curve. All these new operators are non-convolution operators. The method to obtain the  $L^2$  boundedness for first and second generation Calderon-Zygmund operators does not work anymore. This leads to third generation Calderon-Zygmund operators.

### Lecture 3. Littlewood-Paley Theory and Function Spaces

There is a number of ways to set up the Littlewood-Paley theory on  $R^n$ . One standard way is as follows. Let  $\phi(\xi)$  be a real radial bump function supported on  $\{\xi \in R^n : |\xi| \leq 2\}$  which equals 1 on  $\{\xi \in R^n : |\xi| \leq 1\}$ . Let  $\psi$  be the function:  $\psi(\xi) = \phi(\xi) - \phi(2\xi)$ . Thus  $\psi$  is a bump function supported on  $\{\xi \in R^n : \frac{1}{2} \leq |\xi| \leq 2\}$ . By construction we have

$$\sum_k \psi(\xi/2^k) = 1$$

for all  $\xi \neq 0$ . Thus we can partition unity into the function  $\psi(\xi/2^k)$  for integers  $k$ , each of which is supported on an annulus of the form  $|\xi| \sim 2^k$ . We now define the Littlewood-Paley projection operators  $Q_k$  and  $P_k$  by

$$(\widehat{Q_k f})(\xi) = \psi(\xi/2^k) \hat{f}(\xi) \tag{3.1}$$

$$(\widehat{P_k f})(\xi) = \phi(\xi/2^k) \hat{f}(\xi). \tag{3.2}$$

Informally,  $Q_k$  is a frequency projection to time annulus  $\{|\xi| \sim 2^k\}$ , while  $P_k$  is a frequency projection to the ball  $\{|\xi| \leq 2^k\}$ . Observe that  $Q_k = P_k - P_{k-1}$ . Also, if  $f \in L^2$ , then  $P_k(f) \rightarrow 0$  in  $L^2$  as  $k \rightarrow -\infty$ , and  $P_k(f) \rightarrow f$  in  $L^2$  as  $k \rightarrow \infty$ , which follows from the Plancherel theorem. By telescoping the series, we thus have the Littlewood-Paley decomposition

$$f = \sum_k Q_k(f) \tag{3.3}$$

for all  $f \in L^2$ , where the series converges in the  $L^2$  norm.

We are now interested in how the  $L^p$  behavior of the Littlewood-Paley pieces  $Q_k f$  relate to the behavior of  $f$ . First, we write

$$P_k f(x) = f * (2^{nk} \phi(2^k \cdot)) = \int f(x + 2^{-k} y) \phi(y) dy. \tag{3.4}$$

Note that  $\phi$  is a Schwartz function with the total mass  $\int \phi(y) dy = \phi(0) = 1$ . Thus the function  $P_k f$  is an average of  $f$  localized to physical scales  $\leq 2^{-k}$ . In particular, we expect  $P_k f$  to be essentially constant at scales  $< 2^{-k}$ . What does a function  $Q_k f$  look like? Since  $Q_k f = P_{k+2} Q_k f$ , we see from (3.4) that

$$Q_k f(x) = \int Q_k f(x + 2^{-2-k} y) \phi(y) dy. \tag{3.5}$$

Thus,  $Q_k f$  is essentially constant at physical scales  $< 2^{-k}$ . On the other hand, we have  $P_{k-2} Q_k f = 0$ , so from (3.3),

$$\int Q_k f(x + 2^{2-k} y) \phi(y) dy = 0$$

for all  $x \in R^n$ .

This roughly assert that  $Q_k f$  has mean zero at scales  $\leq 2^{-k}$ .

From (3.3) and the Minkowski's inequality, for  $1 \leq p \leq \infty$ , we have

$$\|P_k f\|_p \leq \int \|f(x + 2^{-k}y)\|_p |\phi(y)| dy \leq C \|f\|_p. \quad (3.6)$$

Thus,  $P_k f$  does not get any bigger than  $f$  itself as measured in  $L^p$ , or in any translation invariant Banach space. Similarly,  $\|Q_k f\|_p \leq C \|f\|_p$ . On the other hand, we have  $f = \sum_k Q_k f$ , so by the triangle inequality, we obtain the cheap Littlewood-Paley inequality

$$\sup_k \|Q_k f\|_p \leq C \|f\|_p \leq C \sum_k \|Q_k f\|_p. \quad (3.7)$$

As the name suggests, the cheap Littlewood-Paley inequality is not the sharpest statement one can make connecting the  $L^p$  norms of  $Q_k f$  with those of  $f$ . By using the Fourier transform, when  $p = 2$ , we get

$$\|f\|_2 \sim \left( \sum_k \|Q_k f\|_2^2 \right)^{\frac{1}{2}}. \quad (3.8)$$

In fact, to see this, square both sides and take Plancherel to obtain

$$\|\hat{f}\|_2^2 \sim \sum_k \|[\psi(\cdot/2^k)\hat{f}(\cdot)]\|_2^2.$$

Observe that for each  $\xi \neq 0$ , there are only two values of  $\psi(\xi/2^k)$  which do not vanish, and these two add up to 1.

We can rewrite (3.8) as

$$\|f\|_2 \sim \left\| \left( \sum_k |Q_k f|^2 \right)^{\frac{1}{2}} \right\|_2. \quad (3.9)$$

The quantity  $\left( \sum_k |Q_k f|^2 \right)^{\frac{1}{2}}$  is known as the Littlewood-Paley square function.

Now define  $Sf$  for the vector-valued function by  $Sf(x) = \{Q_k f\}_k$ , and  $|Sf| = \left( \sum_k |Q_k f|^2 \right)^{\frac{1}{2}}$  is the  $\ell^2$  norm of  $Sf$ .

**Theorem 3.10:** For  $1 < p < \infty$ , then  $\|Sf\|_p \sim \|f\|_p$  with the implicit constant depending on  $p$ .

The proof of theorem 3.10 follows from the Calderon-Zygmund real variable method. In fact, we have the  $L^2$  result, and it suffices to see that  $S$  is a vector-valued Calderon-Zygmund operator with vector-valued kernel  $K(x) = (2^{nk}\psi(2^k x))_k$ . Since  $\psi$  is a Schwartz function, it is easy to check that  $K(x)$  satisfies the size and smoothness conditions for the

first generation Calderon-Zygmund operators. Now the  $L^2$  result implies the  $L^p, 1 < p < \infty$ , results. By duality we also have  $\|S^*(f_k)\|_p \leq C\|(f_k)\|_p$ . Thus,

$$\left\| \sum_k Q_k f_k \right\|_p \leq C \left\| \left( \sum_k |f_k|^2 \right)^{\frac{1}{2}} \right\|_p.$$

Similarly,

$$\left\| \sum_k \tilde{Q}_k f_k \right\|_p \leq C \left\| \left( \sum_k |f_k|^2 \right)^{\frac{1}{2}} \right\|_p,$$

where  $\tilde{Q}_k = P_{k+2} - P_{k-2}$ .

We apply this with  $f_k = Q_k f$ , since  $\tilde{Q}_k Q_k = Q_k$ , we obtain

$$\|f\|_p \leq \left\| \left( \sum_k |Q_k f|^2 \right)^{\frac{1}{2}} \right\|_p.$$

As an application, we give a proof of the Hormander-Mikhlin multiplier theorem.

**Theorem 3.11:** Let  $m(\xi)$  be a multiplier such that

$$|\partial^\alpha m(\xi)| \leq C|\xi|^{|\alpha|} \quad (3.12)$$

for all  $|\alpha| \geq 0$ , where the constant  $C$  depends on  $\alpha$ . Let  $T_m$  be the Fourier multiplier with symbol  $m : (\widehat{T_m f})(\xi) = m(\xi)\hat{f}(\xi)$ . Then  $T_m$  is bounded on  $L^p, 1 < p < \infty$ .

**Proof:** We have

$$T_m = \sum_{k,k'} Q_k T_m Q_{k'} = \sum_k \tilde{Q}_k \bar{Q}_k$$

where  $\tilde{Q}_k = Q_k T_m$  and  $\bar{Q}_k = \sum_{k-2 < k' < k+2} Q_{k'}$ . From (3.12) we see that  $\tilde{Q}_k$  and  $\bar{Q}_k$  are smooth frequency localization operators to the annuli  $\{|\xi| \sim 2^k\}$ . Then the theorem 3.11 follows from the Littlewood-Paley  $L^p$  inequality by composing  $\tilde{Q}_k$  and  $\bar{Q}_k$ .

A typical consequence of the Hormander-Mikhlin multiplier theorem is the estimate

$$\|\partial_{x_i} \partial_{x_j} f\|_p \leq C \|\Delta f\|_p$$

for all  $1 < p < \infty$ , where  $\Delta = \sum_{i=1}^n \partial_{x_i}^2$  is the Laplacian on  $R^n$ . This follows because

$$(\partial_{x_i} \partial_{x_j} \widehat{f})(\xi) = \frac{\xi_i \xi_j}{|\xi|^2} (\widehat{\Delta f})(\xi), \text{ and the symbol } m(\xi) = \frac{\xi_i \xi_j}{|\xi|^2} \text{ satisfies the condition (3.12).}$$

The Littlewood-Paley  $L^p, 1 < p < \infty$ , inequality suggest that one can consider the similar inequality for  $0 < p \leq 1$ . However, this fails even when  $p = 1$ . Notice that  $\psi$  defined in the Littlewood-Paley  $S$  function, is in  $\mathcal{S}$ , the Schwartz test function space, so for any  $f \in \mathcal{S}'$ , the temperate distribution space,  $Q_k f$  is well defined. This means that for any  $f \in \mathcal{S}'$ ,  $Sf$  is well defined. Now we introduce the hardy space  $H^p$  as follows.

**Definition 3.13:** For  $0 < p < \infty$ ,  $H^p = \{f \in \mathcal{S}' : Sf \in L^p\}$  and if  $f \in H^p$ , the norm of  $f$  is defined by  $\|f\|_{H^p} = \|Sf\|_p$ .

It is easy to see that when  $1 < p < \infty$ ,  $H^p = L^p$ , by the Littlewood-Paley  $L^p$  inequality. For  $0 < p \leq 1$ ,  $H^p$  is a new space which is different from  $L^p$ . The best way to see this is to get the so-called atomic decomposition of  $H^p$ . Here we only consider  $p = 1$ . To do this, let  $\mathcal{S}_\infty = \{f \in \mathcal{S} : \int f(x)x^\alpha dx = 0, \text{ for all } |\alpha| \geq 0\}$ . We then have

**Theorem 3.14:** Suppose  $f \in \mathcal{S}_\infty$  and  $\sum_k |\widehat{\psi}(2^k \xi)|^2 = 1$ , where  $\psi$  is a bump function supported on the annulus  $\{|x| \sim 2^{-k}\}$ . Then

$$f = \sum_k Q_k Q_k f,$$

where the series converges in the topology of  $\mathcal{S}$ .

By a duality argument, for  $f \in \mathcal{S}'$  and  $g \in \mathcal{S}_\infty$  we have

$$\langle \sum_k Q_k Q_k f, g \rangle = \langle f, \sum_k Q_k Q_k g \rangle = \langle f, g \rangle.$$

Hence, for any  $f \in \mathcal{S}'$ ,  $f = \sum_k Q_k Q_k f$  in  $(\mathcal{S}_\infty)' \approx \mathcal{S}'/\mathcal{P}$ ,  $\mathcal{S}'$  modulo polynomials.

**Definition 3.15:** A function  $a(x)$  is said to be an atom if  $a(x)$  satisfies

- (i)  $\text{Supp } a \subseteq Q$ , a cube in  $R^n$ ;
- (ii)  $\|a\|_2 \leq |Q|^{-\frac{1}{2}}$ ;
- (iii)  $\int a(x) dx = 0$ .

**Theorem 3.15:**  $f \in H^1$  if and only if  $f = \sum_k \lambda_k a_k$ , where  $a_k$  are atoms and  $\sum_k |\lambda_k| < \infty$ .

Moreover,

$$\|f\|_{H^1} \approx \inf \left\{ \sum_k |\lambda_k| : \text{for all } f = \sum_k \lambda_k a_k \right\}.$$

As mentioned before,  $H^p$  is well defined for  $0 < p < \infty$ , but not for  $p = \infty$ . Next, we consider the replacement of  $H^p$  when  $p = \infty$ .

**Definition 3.16:** For  $f \in L^1_{loc}(R^n)$  we let  $\|f\|_* = \sup_Q m_Q |f - m_Q f|$ , where  $m_Q f$  is the average of  $f$  over  $Q$ , and we define the space  $BMO(R^n)$  (functions of bounded mean oscillation) to consist of those functions  $f$  such that

$$\|f\|_* < \infty.$$

$(BMO(R^n))$  is a semi-normed vector space, with the seminorm vanishing on the constant functions. If we let  $\mathcal{C}$  denote the vector space of constant functions, then the quotient of  $BMO$  by  $\mathcal{C}$  is a Banach space, which we also denote by  $BMO$ . This space  $BMO$  was originally introduced by John and Nirenberg. They proved the following John-Nirenberg inequality.

**Theorem 3.17:** There exist two positive constants  $\lambda > 0$  and  $C > 0$  such that for any  $f \in BMO$ ,

$$\sup_Q m_Q \left( \exp\left(\frac{\lambda}{\|f\|_*} |f - m_Q f|\right) \right) \leq C. \quad (3.18)$$

**Proof:** We assume that  $f$  is bounded, so that the above supremum makes sense for all  $\lambda$ , and we shall prove the theorem by finding a bound independent of  $\|f\|_\infty$ .

Let  $Q_0$  be a fixed cube and  $Q$  some dyadic cube. Recall that  $\tilde{Q}$  is the unique dyadic cube which contains  $Q$  and lies in the previous generation. It is easy to see that

$$|m_Q f - m_{\tilde{Q}} f| \leq 2^n \|f\|_* . \quad (3.19)$$

Consider now the Calderon-Zygmund decomposition of the function  $(f - m_{Q_0} f)\chi_{Q_0}$  for  $\lambda = 2\|f\|_*$ . This yields a collection of dyadic cubes  $Q_i$ , maximal with respect to inclusion, satisfying

$$m_{Q_i} |(f - m_{Q_0} f)\chi_{Q_0}| > 2\|f\|_* \quad (3.20)$$

and

$$|(f - m_{Q_0} f)\chi_{Q_0}| \leq 2\|f\|_* \quad (3.21)$$

on  $(\cup Q_i)^c$ .

Clearly,  $Q_i \subseteq Q_0$  for each  $i$ , and  $|\cup Q_i| \leq \frac{\|(f - m_{Q_0} f)\chi_{Q_0}\|_1}{2\|f\|_*} \leq \frac{|Q_0|}{2}$ .

Since the  $Q_i$ 's are maximal,  $m_{\tilde{Q}_i} |(f - m_{Q_0} f)| \leq 2\|f\|_*$ , and (3.19) gives

$$|m_{Q_i} f - m_{Q_0} f| \leq (2^n + 2)\|f\|_* .$$

Let  $X(\lambda) = \sup_Q m_Q \exp(\frac{\lambda}{\|f\|_*} |f - m_Q f|)$ , which is finite since we are assuming that  $f$  is bounded. We obtain

$$\begin{aligned} m_{Q_0} (\exp(\frac{\lambda}{\|f\|_*} |f - m_{Q_0} f|)) &\leq \frac{1}{|Q_0|} \int_{Q_0 \setminus \cup Q_i} e^{2\lambda} dx \\ &+ \frac{1}{|Q_0|} \sum_i \frac{|Q_i|}{|Q_i|} \left[ \int_{Q_i} \exp(\frac{\lambda}{\|f\|_*} |f - m_{Q_i} f|) dx e^{(2^n + 2)\lambda} \right] \\ &\leq e^{2\lambda} + \frac{1}{2} [\exp(2^n + 2)] X(\lambda) . \end{aligned}$$

From taking the supremum over all cube  $Q_0$  it follows that  $X(\lambda)[1 - \frac{1}{2}\exp(2^n + 2)] \leq e^{2\lambda}$ , which implies that  $X(\lambda) \leq C$ , if  $\lambda$  is small enough, which proves the theorem.

A consequence of the theorem, which in fact is equivalent to it, is the following: There exist positive constant  $\lambda$  and  $C$  such that for every cube  $Q$  and every  $t > 0$ ,

$$|\{x \in Q : |f(x) - m_Q f| > t\|f\|_*\}| \leq C e^{-\lambda t} |Q|. \quad (3.22)$$

For  $1 \leq p < \infty$ , then  $\|f\|_{p,*} = \sup_Q [m_Q |f - m_Q f|^p]^{\frac{1}{p}}$  are equivalent.

Now we are ready to prove the duality of  $H^1$  and BMO. We shall see that each continuous linear functional on  $H^1$  can be realized as a mapping

$$\ell(g) = \int f(x)g(x)dx, g \in H^1,$$

when suitably defined, where  $f$  is a function in BMO.

For general  $f \in BMO$  and  $g \in H^1$ , the integral in (3.23) does not converge absolutely. For this reason, we take  $g \in H^1$  has finite atomic decomposition. We denote this subspace by  $H_a^1$  which is dense in  $H^1$ .

**Theorem 3.24:** (a) Suppose  $f \in BMO$ . Then the linear functional given by (3.23), initially defined on  $H_a^1$ , has a unique bounded extension to  $H^1$  and satisfies

$$\|\ell\| \leq c\|f\|_*.$$

(b) Conversely, every continuous linear functional on  $H^1$  can be realized as above, with  $f \in BMO$ , and with

$$\|f\|_* \leq c\|\ell\|.$$

**Proof:** To see (a), note that if  $a$  is an atom supported on  $Q$ , then

$$\left| \int f(x)a(x)dx \right| = \left| \int [f(x) - m_Q f]a(x)dx \right| \leq |Q|^{\frac{1}{2}}\|f\|_{2,*}\|a\|_2 \leq c\|f\|_*.$$

Thus, if  $g \in H_a^1$ , then  $g = \sum_k \lambda_k a_k$  with the sum having a finite terms and  $a_k$  are atoms.

So

$$\begin{aligned} \left| \int f(x)g(x)dx \right| &= \left| \sum_k \lambda_k \int [f(x) - m_{Q_k} f]a_k(x)dx \right| \\ &\leq \sum_k |\lambda_k| |Q_k|^{\frac{1}{2}} \|f\|_{2,*} \|a_k\|_2 \leq c\|f\|_* \sum_k |\lambda_k|. \end{aligned}$$

To show (b), fix a cube  $Q$  and let  $L_Q^2$  be the space of all square integrable functions supported on  $Q$ . Let  $L_{Q,0}^2 = \{f \in L_Q^2 : \int f(x)dx = 0\}$ . Note that every  $g \in L_{Q,0}^2$  is a multiple of an atom and  $\|g\|_{H^1} \leq c|Q|^{\frac{1}{2}}\|g\|_2$ . Thus if  $\ell$  is a given linear functional on  $H^1$  with the norm  $\leq 1$ , then  $\ell$  extends to a linear functional on  $L_{Q,0}^2$  with norm at most  $c|Q|^{\frac{1}{2}}$ . By the Riesz representation theorem for the Hilbert space  $L_{Q,0}^2$ , there exists an element  $F^Q \in L_{Q,0}^2$  so that

$$\ell(g) = \int F^Q(x)g(x)dx, \tag{3.25}$$

if  $g \in L_{Q,0}^2$ , with

$$\left( \int |F^Q(x)|dx \right)^{\frac{1}{2}} \leq c|Q|^{\frac{1}{2}}. \tag{3.26}$$

Hence for each  $Q$ , we get such a function  $F^Q$ . We want to have a single function  $f$  so that, on each  $Q$ ,  $f$  differs from  $F^Q$  by a constant. To construct this  $f$ , observe that if  $Q_1 \subseteq Q_2$ , then  $F^{Q_1} - F^{Q_2}$  is constant on  $Q_1$ . Indeed, both  $F^{Q_1}$  and  $F^{Q_2}$  give the same functional on  $L_{Q_1,0}^2$ , so they must differ by a constant on  $Q_1$ . We can modify  $F^Q$ , replacing it with  $f^Q = F^Q + c_Q$ , where  $c_Q$  is a constant chosen so that  $f^Q$  has the average zero over the

unit cube centered at the origin. It follows that  $f^{Q_1} = f^{Q_2}$  on  $Q_1$ , if  $Q_1 \subseteq Q_2$ . Finally, we define  $f$  on  $R^n$  by taking  $f(x) = f^Q(x)$  for  $x \in Q$ . Observe that

$$\frac{1}{|Q|} \int_Q |f(x) - c_Q| dx \leq \left( \frac{1}{|Q|} \int_Q |f(x) - c_Q|^2 dx \right)^{\frac{1}{2}} = \left( \frac{1}{|Q|} \int_Q |F^Q|^2 dx \right)^{\frac{1}{2}} \leq c,$$

which shows  $f \in BMO$  with  $\|f\|_* \leq c$ . Also, by (3.25),  $\ell(g) = \int F^Q(x)g(x)dx$ , if  $g \in L_{Q,0}^2$ , for some  $Q$ , in particular, this representation holds for all  $g \in H_a^1$ . The converse (b) of the theorem is proved.

We now discuss the relationship between BMO and Carleson measures and Littlewood-Paley square functions.

**Definition 3.27:** A Borel measure  $d\mu$  on  $R_+^{n+1}$  is said to be a Carleson measure if

$$\sup_Q \frac{1}{|Q|} \int_{T(Q)} |d\mu| \leq C < \infty. \quad (3.28)$$

If  $d\mu$  is a Carleson measure, we denote  $\|d\mu\|_C$ , the Carleson norm of  $d\mu$ , by the smallest constant  $C$  in (3.28).

**Theorem 3.29:** Suppose  $\phi \in \mathcal{S}$  with  $\int \phi(x)dx = 1$ . Then  $d\mu$  is a Carleson measure if and only if

$$\int_{R_+^{n+1}} |\phi_t * f(x)|^2 d\mu \leq C \|f\|_2^2. \quad (3.30)$$

**Theorem 3.31:** Let  $\psi \in \mathcal{S}$  with  $\int_0^\infty |\widehat{\psi}(t\xi)|^2 \frac{dt}{t} = 1$  for all  $\xi \neq 0$ . Then  $f \in BMO$  if and only if  $|f * \psi_t(x)|^2 \frac{dxdt}{t}$  is a Carleson measure.

Littlewood-Paley theory allows us to consider a large range of classical function spaces within a single framework. The general classes of spaces we will define are the homogeneous Besov spaces  $\dot{B}_p^{\alpha,q}$  and Triebel-Lizorkin spaces  $\dot{F}_p^{\alpha,q}$  as well as their inhomogeneous analogs.

Let us choose  $\phi \in \mathcal{S}$  so that  $\text{Supp } \widehat{\phi} \subseteq \{\xi : \frac{1}{2} \leq |\xi| \leq 2\}$  and  $|\widehat{\phi}(\xi)| \geq c > 0$  if  $\frac{3}{5} \leq |\xi| \leq \frac{5}{3}$ . For  $\alpha \in R, p \neq \infty, 0 < p, q \leq \infty$  and  $f \in \mathcal{S}'$  we define

$$\|f\|_{\dot{F}_p^{\alpha,q}} = \left\| \left\{ \sum_k (2^{k\alpha} |\phi_k * f|)^q \right\}^{\frac{1}{q}} \right\|_p, \quad (3.32)$$

and, for the same indices, and including  $p = \infty$ ,

$$\|f\|_{\dot{B}_p^{\alpha,q}} = \left\{ \sum_k (2^{k\alpha} \|\phi_k * f\|_p)^q \right\}^{\frac{1}{q}}, \quad (3.33)$$

where  $\phi_k(x) = 2^{kn} \phi(2^k x)$ .

Note that  $\phi_k * f$  is a smooth function when  $\phi \in \mathcal{S}$  and  $f \in \mathcal{S}'$ . Also  $\|f\|_{\dot{F}_p^{\alpha,q}} = 0$  or  $\|f\|_{\dot{B}_p^{\alpha,q}} = 0$  if and only if  $\phi_k * f$  is the zero function for all  $k$ . But this is equivalent to having  $\hat{f}(\xi)\hat{\phi}(2^k\xi)$  be zero for all  $k$ . because of the conditions on  $\hat{\phi}$ , this, in turn, is equivalent to  $\text{Supp } \hat{f} = \{0\}$ . Finally, this means that the distribution  $f$  is a polynomial. Thus, we work modulo polynomials when considering (3.32) and (3.33); that is,  $f \in \mathcal{S}'/\mathcal{P}$  in these equalities, where  $\mathcal{P}$  denote the class of polynomials on  $R^n$ . In particular, we define  $\dot{F}_p^{\alpha,q}$  and  $\dot{B}_p^{\alpha,q}$  to be the set of all such  $f$  for which the expression (3.32) and (3.33) is finite. It is not difficult to see that these expressions are norms when  $1 \leq p, q \leq \infty$  and quasinnorms in general. We are not include the case  $p = \infty$  in the definition of the Triebel-Lizorkin spaces. In this case, the  $L^\infty$  norm should be replaced by a Carleson measure condition. The spaces defined by the finiteness of the these norms are called the homogeneous Triebel-Lizorkin and Besov spaces, respectively. The inhomogeneous versions of these spaces are obtained by adding the term  $\|\Phi * f\|_p$  to variants of the above expressions, where  $\Phi \in \mathcal{S}$ ,  $\text{Supp } \Phi \subseteq \{\xi : |\xi| \leq 2\}$  and  $|\hat{\Phi}(\xi)| \geq c > 0$  if  $|\xi| \leq \frac{5}{3}$ . The variants in question are as in (3.32) and (3.33) with  $\sum_k$  replaced by  $\sum_{k \geq 1}$ . These spaces are denoted by  $F_p^{\alpha,q}$  and  $B_p^{\alpha,q}$  and they are spaces of tempered distributions; the necessity of considering such distributions modulo polynomials disappears since  $\hat{\Phi}(0) \neq 0$ .

By the results mentioned above by the Littlewood-Paley theory, we obtain the following identifications:

$$L^p \sim \dot{F}_p^{0,2}$$

when  $1 < p < \infty$ ;

$$H^p \sim \dot{F}_p^{0,2}$$

when  $0 < p \leq 1$ ;

$$BMO \sim \dot{F}_\infty^{0,2}$$

when  $\dot{F}_\infty^{0,2}$  is defined by the Carleson measure. We also can show  $L_\alpha^p \sim F_p^{\alpha,2}$  and  $\dot{L}_\alpha^p \sim \dot{F}_p^{\alpha,2}$  when  $\alpha > 0$  and  $1 < p < \infty$ ;

$$\Lambda_\alpha \sim F_\infty^{\alpha,\infty}$$

and

$$\dot{\Lambda}_\alpha \sim \dot{F}_\infty^{\alpha,\infty}$$

when  $\alpha > 0$ . Suppose  $\tilde{\phi}$  and  $\tilde{\Phi}$  are two other functions satisfying the properties of  $\phi$  and  $\Phi$  announced above. One can show that replacing  $\phi$  and  $\Phi$  with  $\tilde{\phi}$  and  $\tilde{\Phi}$  in the definitions yields the same spaces with equivalent norms.

## Lecture 4. Third generation Calderon-Zygmund operators and the T1 theorem

The pseudo-differential calculus is like that mythological bird. Its first birth was at the end of the 1930s, the founding fathers being Giraud and Marcinkiewicz. The second birth took place at the end of the 1950s, as we discussed in the Lecture 2, and it clearly benefited from the theory of distributions, developed by Schwartz during the 1940s. The third birth is the one to claim in this lecture. In order to deal with linear partial differential equations having coefficients which are only slightly regular and. in order to approach the problem of the regularity of solutions of non-linear partial differential equations, Calderon decided to make the pseudo-differential calculus include the operators  $A$  of pointwise multiplication by functions  $a(x)$  which are only slightly regular with respect to  $x$ . Of course, Calderon wanted to keep what had been gained during the previous decades: the classical pseudo-differential operators. An important step was taken in 1965 when Calderon proved that the commutator  $[A, DH] = ADH - DHA$  between the pointwise multiplication operator  $A$  by the function  $a(x)$  and the operator  $DH$ , where  $D = -i \frac{d}{dx}$  and  $H$  is the Hilber transform, is bounded on  $L^2$  if  $a$  is a Lipschitz function, i.e.,  $|a(x) - a(y)| \leq c|x - y|$  for  $x, y \in R$ . We note that the commutator  $[A, DH](f)(x)$  is given by

$$[A, DH](f)(x) = p.v. \int \frac{A(x) - A(y)}{(x - y)^2} f(y) dy. \quad (4.1)$$

This operator is called Calderon's first commutator. To see why this operator plays an important role in the study of linear partial differential operators with variable coefficients, we follow Calderon's 1978 International Congress lecture. Let  $L$  be an operator defined by

$$Lf(x) = \sum_{j=0}^m a_j(x) \frac{d^j f}{dx^j}. \quad (4.2)$$

As we did in lecture 2, by Fourier transform and Fourier inversion,

$$Lf(x) = \frac{1}{2\pi} \int \sum_{j=0}^m a_j(x) (i\xi)^j \hat{f}(\xi) e^{ix\xi} d\xi. \quad (4.3)$$

The idea behind pseudo-differential operators is to replace the function  $a_j(x)(i\xi)^j$  by more general functions  $\sigma(x, \xi)$  is such a way that the resulting class of operators is closed under composition, adjunction, and other basic operations. If we want this class of operators to be closed under composition, and, in particular, be able to freely compose linear differential operators  $L$ , then it will only contain differential operators with infinitely differentiable coefficients, i.e.,  $a \in C^\infty$ .

There is another algebra of operators, however, that can be used in the study of operators  $L$  as above with nonsmooth coefficients. Let  $\Lambda$  be the operator defined by  $(\widehat{\Lambda f})(\xi) = \psi(\xi) \hat{f}(\xi)$  where  $\psi$  is an infinitely differentiable function with  $\psi(\xi) = |\xi|$  if  $|\xi| \geq 1$ , and let

$$Tf(x) = \int q(x, \xi) \hat{f}(\xi) e^{ix\xi} d\xi + Rf(x), \quad (4.4)$$

where

$$Rf(x) = \int r(x, \xi) \hat{f}(\xi) e^{ix\xi} d\xi. \quad (4.5)$$

Here  $q(x, \xi) = |\xi|^{-m} a_m(x) (i\xi)^m$  and  $r(x, \xi)$  is defined by the relation

$$\frac{1}{2\pi} \sum_{j=0}^m a_j(x) (i\xi)^j = (q(x, \xi) + r(x, \xi)) (\psi(\xi))^m.$$

Now we can write

$$Lf = T\Lambda^m f.$$

It is easy to show that the operator  $R$  and  $R\frac{d}{dx}$  are bounded on  $L^2$ , say, provided the coefficients  $a_j$  are bounded. The function  $q(x, \xi)$  is regular, homogeneous of degree 0 in  $\xi$ , and bounded. The corresponding operator  $T$  can be generalized by allowing  $q(x, \xi)$  to be a general function with these three properties, and allowing  $R$  to be any operator such that  $R$  and  $R\frac{d}{dx}$  are bounded on  $L^2$ . To avoid some pathologies it turns out to be necessary to restrict the class slightly and assume, in addition, that  $q(x, \xi)$  is Lipschitz in  $x$ . The class of operators  $L$ , given by (4.6), with  $T$  in this more general class, at least contains the linear differential operators whose coefficients are bounded, and, for the highest terms, bounded and Lipschitz. Let  $A$  be the operator corresponding to multiplication by the Lipschitz function  $a(x)$ , and let  $H$  be the Hilbert transform. Obviously,  $A$  is one of the operators  $T$ , and if we recall that

$$Hf(x) = c \int \text{sign}(\xi) \hat{f}(\xi) e^{ix\xi} d\xi, \quad (4.7)$$

then it becomes clear that  $H$  is in this class as well. To prove that the class of operators  $T$  as above is closed under composition, it is necessary to show that  $AH$  and  $HA$  are also of the same general type. For  $AH$  this is trivial. For  $HA$ , if we write  $HA = AH + (HA - AH)$ , then it becomes clear that  $HA$  is also of the right type if  $(AH - HA)D$  is bounded on  $L^2$ . Now

$$(AH - HA)D = [A, HD] + HDA - HAD$$

and  $DA - AD$  is just multiplication by  $a'(x)$ , which of course is bounded on  $L^2$  since  $a'$  is a bounded function. Hence,  $HA$  belongs to the class if and only if  $[A, HD]$  is bounded on  $L^2$ , and this is Calderon's result since  $HD = DH$ . Now to show that the composition of two general operators  $T$  in the class is still in the class can be reduced to the special cases we just considered. The fact that the class is closed under composition can be used to prove existence and uniqueness results, a priori estimates, etc. for partial differential equations.

There are also many other operators which are not convolution operators that arise naturally in analysis. Calderon's  $k$ th commutator, for example, is given by

$$C_k f(x) = p.v. \int \left( \frac{a(x) - a(y)}{x - y} \right)^k \frac{f(y)}{x - y} dy, \quad k \geq 1. \quad (4.9)$$

These operators are closely related to the boundary behavior of analytic functions given by Cauchy integrals

$$f(z(x)) = \frac{1}{2\pi i} \int \frac{1}{z(y) - z(x)} f(z(y)) z'(y) dy \quad (4.10)$$

on Lipschitz curve  $z(x) = x + ia(x)$ ,  $a' \in L^\infty$ . Another nonconvolution operator is the double layer potential associated with a domain  $\Omega$ . In local coordinates this operator takes the form

$$Tf(x) = p.v. \frac{1}{\omega_n} \int_{R^n} \frac{a(x) - a(y) - (x - y) \cdot \nabla a(y)}{(|x - y|^2 + (a(x) - a(y))^2)^{\frac{n+1}{2}}} f(y) dy, \quad (4.11)$$

where  $\omega_n$  is the area of the unit sphere in  $R^n$ . To solve the Dirichlet problem in a Lipschitz domain by the method of layer potentials, one needs to prove the boundedness of  $L^2$  of the above operator with a Lipschitz.

We emphasize that while for convolution operators, boundedness on  $L^2$  is a simple application of Plancherel's theorem, the  $L^2$  - boundedness for non-convolution operators like the ones above is highly nontrivial.

In 1978 Coifman and Meyer introduced the third generation Calderon-Zygmund operators. Let  $T$  be a continuous linear operator from the Schwartz class  $\mathcal{S}$  of test functions to its dual  $\mathcal{S}'$ . By the Schwartz kernel theorem there is a distribution  $K$  in  $\mathcal{S}'$  such that

$$(Tf, g) = (K, g \otimes f)$$

for all  $f, g \in \mathcal{S}$  and here  $(, )$  denotes the distribution pairing, linear in each coordinate, rather than the pairing  $\langle, \rangle$ , which is conjugate linear in the second coordinate, and  $g \otimes f(x, y) = g(x)f(y)$ . The distribution  $K$  is called the kernel of  $T$ .

**Definition 4.12:** We say that  $K$  is a Calderon-Zygmund kernel if its restriction to the set  $\Omega = \{(x, y) \in R^n \times R^n : x \neq y\}$  is a continuous function  $K(x, y)$  which satisfies

$$|K(x, y)| \leq C \frac{1}{|x - y|^n}, \quad (4.13)$$

$$|K(x, y) - K(x', y)| \leq C \frac{|x - x'|^\epsilon}{|x - y|^{n+\epsilon}}, \quad (4.14)$$

if  $|x - x'| \leq \frac{1}{2}|x - y|$ ,

$$|K(x, y) - K(x, y')| \leq C \frac{|y - y'|^\epsilon}{|x - y|^{n+\epsilon}}, \quad (4.15)$$

if  $|y - y'| \leq \frac{1}{2}|x - y|$ ,

for some constant  $C$  and some  $\epsilon$  in  $(0, 1]$ . We call  $T$  a Calderon-Zygmund singular integral operator, and write  $T \in CZSIO$ , or  $T \in CZSIO(\epsilon)$ , if the kernel of  $T$  satisfies these conditions. In particular, if  $T \in CZSIO$ , then

$$(Tf, g) = \int \int K(x, y) g(x) f(y) dy dx \quad (4.16)$$

whenever  $f, g \in \mathcal{S}$  and  $\text{supp}f \cap \text{supp}g = \emptyset$ .

$T$  is said to be a third generation Calderon-Zygmund operator if  $T \in CZSIO$  and it can be extended to a bounded operator on  $L^2$ . We write  $T \in CZO$ .

By the Calderon-Zygmund real variable method, we can easily get the following theorem.

**Theorem 4.17:** If  $T \in CZO$ , then  $T$  is bounded on  $L^p$  for  $1 < p < \infty$ , from  $H^1$  to  $L^1$ , and from  $L^\infty$  to BMO.

A basic, important problem is how to understand when a Calderon-Zygmund singular integral operator is bounded on  $L^2$ . As Meyer wrote " To go beyond the context of convolution operators, it becomes indispensable to have a criterion for  $L^2$  continuity, without which the theory collapses like a house of cards. This problem was solved by the celebrated  $T1$  theorem of David-Journe in 1983. This theorem represents the culmination of the theory started by Calderon and Zygmund some thirty years earlier. Let's go back to the Calderon commutator. In 1965, Calderon proved the  $L^2$  boundedness of Calderon commutator. After 9 years, in 1974, Coifman and Meyer proved the same result for the second order Calderon commutator. in 1977, Calderon proved the  $L^2$  boundedness for the Cauchy integral on a Lipschitz curve with the Lip norm  $< \epsilon$ , where  $\epsilon$  is a small number. Coifman, McIntosh and Meyer, in 1981, proved the  $L^2$  boundedness for the Cauchy integral on all Lipschitz curves.

Before we can state the David-Journe  $T1$  theorem, we require some preliminary definition and concepts. We begin by assigning a meaning to  $T1$  and  $T^*1$  for  $T \in CZSIO$ . Even in the classical case of  $T \in CASIO$ , this requires care. Denote  $\mathcal{D}$  the set of all  $f \in \mathcal{S}$  with the compact support and  $\mathcal{D}_0 = \{f \in \mathcal{D} : \int f(x)dx = 0\}$ . We now can define  $T1$  as a linear functional on  $\mathcal{D}_0$  as follows. For any given  $f \in \mathcal{D}_0$ , choose  $\chi_1 \in \mathcal{D}$  such that  $\chi_1(x) = 1$  for  $x \in 2Q$ , where  $Q$  is a cube containing support of  $f$ , and  $1 = \chi_1 + \chi_2$ . Since  $\chi_1 \in \mathcal{D}$ , so  $\langle T\chi_1, f \rangle$  is well defined. Note that, formally,  $\langle T\chi_2, f \rangle = \langle \chi_2, T^*f \rangle$  and this makes sense if  $\langle \chi_2, T^*f \rangle$  is well defined. Indeed, by the fact that  $f \in \mathcal{D}_0$ , we have

$$\langle \chi_2, T^*f \rangle = \int \int \chi_2(x)K(x, y)f(y)dydx = \int \chi_2(x) \int [K(x, y) - K(x, y_Q)]f(y)dydx.$$

By the conditions on  $K$ , we obtain the absolute value of the last  $\|f\|_1$ . This means  $\langle \chi_2, T^*f \rangle$  is well defined. Now we define  $\langle T1, f \rangle = \langle T\chi_1, f \rangle + \langle \chi_2, T^*f \rangle$ . It is easy to see that the equality does not depend the choice of  $\chi$ . We also define  $T1 = 0$  by the fact that  $\int T^*f(x)dx = 0$  for all  $f \in \mathcal{D}_0$ , and  $T^*1 = 0$  is defined similarly. Note that  $\mathcal{D}_0$  is a subset, so  $T1$  is a BMO function means that  $T1$ , as a linear functional on  $\mathcal{D}_0$ , can be extended to be a linear functional on  $H^1$ , and it is similar for  $T^*1$  being a BMO function. Next we consider a condition known as the weak boundedness property.

**Definition 4.18:** We say that a linear continuous operator  $T : \mathcal{S} \rightarrow \mathcal{S}'$  satisfies the weak boundedness property, and we write  $T \in WBP$ , if for each bounded subset  $\mathcal{B}$  of  $\mathcal{S}$  there exists a constant  $C = C(\mathcal{B})$  such that for all  $f, g \in \mathcal{B}$ ,

$$|\langle T(f_t^z), g_t^z \rangle| \leq Ct^{-n}, z \in R^n, t > 0, \quad (4.19)$$

where  $f_t^z(x) = t^{-n}f(\frac{x-z}{t})$ .

Note that if  $T$  is bounded on  $L^p$ , for some fixed  $1 < p < \infty$ , then  $T \in WBP$ . As an example, a rather large class of operators which satisfy the weak boundedness property is the class of CZSIO's with antisymmetric kernel  $K$ , i.e., such that  $K(x, y) = -K(y, x)$ . In this case,

$$\langle Tf, g \rangle = \lim_{\epsilon \rightarrow 0} \frac{1}{2} \int \int_{|x-y|>\epsilon} K(x, y)[f(y)g(x) - f(x)g(y)] dy dx,$$

and, using the definition of a Calderon-Zygmund singular integral kernel, it is easy to see that  $T \in WBP$ . This class of operators includes the all Calderon commutators.

We are now ready to state the  $T1$  theorem of David-Journe.

**Theorem 4.20:** Suppose  $T \in CZSIO$ . Then  $T \in CZO$  if and only if  $T1 \in BMO, T^*1 \in BMO$ , and  $T \in WBP$ .

**Proof: Step 1.** Consider first the case when  $T1 = T^*1 = 0$ . BY the Littlewood-paley theory and the Calderon reproducing formula

$$\|Tf\|_2 \leq C \|S(Tf)\|_2 = C \left\| \left\{ \int_0^\infty |\psi_t * (T \int_0^\infty \psi_s * \psi_s * f \frac{ds}{s})| \frac{dt}{t} \right\}^{\frac{1}{2}} \right\|_2.$$

A basic estimate under the condition of  $T$  is the following inequality:

$$|\psi_t * (T\psi_s)(x, y)| \leq C \left(\frac{s}{t} \wedge \frac{t}{s}\right)^\epsilon \frac{(t \vee s)^\epsilon}{(t \vee s + |x - y|)^{n+\epsilon}}, \quad (4.21)$$

where  $\psi_t * (T\psi_s)(x, y)$  is the kernel of  $\psi_t * (T\psi_s)$ ,  $a \wedge b = \min(a, b)$ , and  $a \vee b = \max(a, b)$ .

Using the estimate in (4.21) yields

$$\int_0^\infty |\psi_t * (T \int_0^\infty \psi_s * \psi_s * f \frac{ds}{s})|^2 \frac{dt}{t} \leq C \int_0^\infty M^2(\psi_s * f) \frac{ds}{s}$$

and then applying the Fefferman-Stein vector-valued maximal function inequality and, again, the Littlewood-paley theory, we obtain

$$\|Tf\|_2 \leq C \left\| \left\{ \int_0^\infty |\psi_s * f|^2 \frac{ds}{s} \right\}^{\frac{1}{2}} \right\|_2 \leq C \|f\|_2.$$

**Step 2.** To reduce the general case to the case in step 1, we need the so-called para-product operator.

**Definition 4.22:** Suppose  $b \in BMO$  and  $\psi$  and  $\phi$  are functions as defined in the Littlewood-Paley theory. The para-product operator  $\Pi_b$ , is defined by

$$\Pi_b f(x) = \int_0^\infty \psi_t * [\psi_t * b(\cdot) \phi_t * f(\cdot)] \frac{dt}{t}. \quad (4.23)$$

**Theorem 4.24:** If  $b \in BMO$ , then  $\Pi_b \in CZO$ . Moreover,

$$\|\Pi_b\| \leq C\|b\|_*.$$

Now define  $\tilde{T} = T - \Pi_{T_1} - (\Pi_{T^*_1})^*$ . Note that  $\Pi_b(1) = b$  and  $(\Pi_b)^*(1) = 0$ . It follows that  $\tilde{T} \in CZSIO$  and  $\tilde{T}(1) = (\tilde{T})^*(1) = 0$ . By step 1,  $\tilde{T}$  is bounded on  $L^2$ , and, hence,  $T$  is bounded on  $L^2$ .

As a beautiful application of the  $T1$  theorem, we prove all Calderon's commutators are bounded on  $L^2$ . Recall that the  $k$ th order Calderon commutator is defined by

$$C_k f(x) = p.v. \int \left( \frac{a(x) - a(y)}{x - y} \right)^k \frac{f(y)}{x - y} dy. \quad (4.24)$$

particularly,  $C_0 = H$ , the Hilbert transform. So  $C_0$  is a  $CZO$ . To see  $C_1$  is bounded on  $L^2$ , by the  $T1$  theorem it suffices to show  $C_1(1)$  is a  $BMO$  function because all kernels of  $C_k$  for  $k \geq 0$ , are antisymmetric, and, as mentioned above,  $C_k \in WBP$  for all  $k \geq 0$ . However, by the integration by parts,  $C_1(1)(x) = H(a')(x)$  which shows  $C_1(1) \in BMO$ . repeating this procedure, we prove  $C_k \in CZO$ . Moreover,

$$\|C_k\| \leq (C\|a'\|_\infty)^k$$

for some constant  $C$ . Finally, the Cauchy integral on a Lipschitz curve can be decomposed by

$$\int \frac{1}{x - y + i(a(x) - a(y))} f(y) dy = \sum_{k=0}^{\infty} (-i)^k C_k f(x)$$

which shows that if  $\|a'\|_\infty < \epsilon$ , then the Cauchy integral is bounded on  $L^2$ .

Calderon's research program was motivated by the study of elliptic partial differential equations in domains with irregular boundary. Calderon's method consists of replacing the partial differential equation on the interior by a pseudo-differential equation on the boundary. But if the boundary is only Lipschitz, the nature of this equation changes: the operators which appear are no longer pseudo-differential but are of the third generation of Calderon-Zygmund operators. When Calderon inaugurated his program, there were two difficulties. The very existence of the operators, needed for the method, was problematic. And, supposing that, in the fullness of time, such operators could be constructed, it would be necessary to solve the equations on the boundary that this process led to. The regrettable absence of such a symbolic calculus was signaled above, and is the second problem of Calderon's program. We now illustrate these remarks by examining a classical problem, which goes back to Poincare, Neumann, and Hilbert. This is the solution, by the double-layer potential method, of the Dirichlet and Neumann problem for a domain  $\Omega$  in  $R^{n+1}$ .

When  $\Omega$  is a bounded, regular, open set, the operator of Calderon's method are classical pseudo-differential operators. They are also singular integral operators whose kernels are given by a double-layer potential. After the reduction given by Calderon's method, resolving a Dirichlet or Neumann problem just amounts to inverting an operator

of the form  $\frac{1}{2} + K$  acting on the boundary. The ambient Banach space will be  $L^2(\partial\Omega, d\sigma)$ , where  $d\sigma$  is the surface measure on the boundary  $\partial\Omega$  of  $\Omega$ .

When  $\Omega$  is a bounded, regular, open set (of class  $C^{1+\alpha}$ , for some  $\alpha > 0$ ), the operator  $K : L^2(\partial\Omega, d\sigma) \rightarrow L^2(\partial\Omega, d\sigma)$  is compact, so Fredholm theory allows one to invert  $\frac{1}{2} + K$ .

When  $\Omega$  is just of class  $C^1$ , the principal difficulty is to prove that  $K$  is continuous on the space  $L^2(\partial\Omega, d\sigma)$ . Once this continuity had been established (Calderon, 1977), the operator  $K$  was still compact, so that the rest could be done just as in the regular case. However,  $K$  no longer was a classical pseudo-differential operator. Finally, for  $\Omega$  a bounded, Lipschitz, open set, the continuity of  $K$  was established in 1981. But the operator  $K$  is no longer compact. This essential difficulty was overcome by showing that  $\frac{1}{2} + K$  was invertible, using the Jerison and Kenig energy inequality.

## Lecture 5. Spaces of Homogeneous Type and nonhomogeneous space

The Calderon-Zygmund operator theory may be substantially widened by replacing  $R^n$  with an abstract space having a few simple properties. This is space of homogeneous type introduced by Coifman and Weiss in 1970.

A quasimetric  $\rho$  on a set  $X$  is a function  $\rho : X \times X \rightarrow [0, \infty]$  satisfying

$$\rho(x, y) = 0 \tag{5.1}$$

if and only if  $x = y$ ,

$$\rho(x, y) = \rho(y, x) \tag{5.2}$$

for all  $x, y$ ,

$$\rho(x, y) \leq A[\rho(x, z) + \rho(z, y)] \tag{5.3}$$

for all  $x, y, z$ , where  $A < \infty$  is a constant.

Any quasimetric defines a topology, for which the balls  $B(x, r) = \{y \in X : \rho(x, y) < r\}$  form a base. However, the balls themselves need not to be open when  $A > 1$ . A general setting for the Calderon-Zygmund operator theory is as follows.

**Definition 5.4:** A space of homogeneous type  $(X, \rho, d\mu)$  is a set  $X$  together with a quasimetric  $\rho$  and a nonnegative measure  $d\mu$  on  $X$  such that  $\mu(B(x, r)) < \infty$  for all  $x \in X$  and  $r > 0$ , and  $d\mu$  satisfies the doubling condition, that is, there exists a constant  $C < \infty$  such that for all  $x \in X$  and  $r > 0$ ,

$$\mu(B(x, 2r)) \leq C\mu(B(x, r)). \tag{5.5}$$

It is unsettling that the balls associated with a quasimetric need not to be open, in general. Fortunately, a result of Macias and Segovia asserts that the given quasimetric may always be replaced by an equivalent one with better properties.

**Theorem 5.6:** Let  $\rho$  be a quasimetric on a set  $X$ . Then there exist  $C < \infty, \epsilon > 0$  and a quasimetric  $\rho'$  on  $X$ , equivalent to  $\rho$  in the sense that  $C^{-1}\rho'(x, y) \leq \rho(x, y) \leq C\rho'(x, y)$  for all  $x, y \in X$ , such that for all  $x, y, z \in X$ ,

$$|\rho'(x, z) - \rho'(y, z)| \leq C\rho'(x, y)^\epsilon [\rho'(x, z) + \rho'(y, z)]^{1-\epsilon}. \tag{5.7}$$

Then the balls associated with  $\rho'$  also form a base for the original topology, but are easily seen to be open as a consequence of (5.7). It is also easy to verify, using (5.7), the existence of  $\delta_0 > 0$  such that  $(\rho')^{\delta_0}$  is a metric, not merely a quasi-metric.

Before we are going to set up the Calderon-Zygmund operator theory on space of homogeneous type, let us examine some examples.

**Example 1.**  $R^n$ , with the Euclidean metric and Lebesgue measure.

**Example 2.** Any  $C^\infty$  compact Riemannian manifold, with the Riemannian metric and volume.

**Example 3.** The graph of a Lipschitz function  $F: R^n \rightarrow R$ , with the induced Euclidean metric and with  $\mu(F(E)) = |E|$ , the Lebesgue measure of  $E \subseteq R^n$ . Similarly, the boundary of any bounded Lipschitz domain in  $R^n$ , that is, any bounded open set whose boundary may be represented locally as the graph of a Lipschitz function after a suitable rotation of coordinates. This includes all polygons in  $R^2$ .

**Example 4.** Let  $n \geq 1$  and  $a_1, \dots, a_n > 0$ . Set  $\rho(x, y) = \sum_{j=1}^n |x_j - y_j|^{a_j}$  for  $x, y \in R^n$ . Then  $\rho$  is a quasimetric, though not in general a metric, and  $R^n$  becomes a space of homogeneous type when equipped with  $\rho$  and Lebesgue measure.

**Example 5.** Let  $n \geq 1$  be an integer and  $d \in (0, n]$  a real number. Let  $E \subseteq R^n$  be a closed subset whose  $d$ -dimensional Hausdorff measure  $\Lambda_d(E)$  is finite and positive. Suppose it happens that there exists  $C < \infty$  so that for every  $x \in E$  and  $r > 0$ ,

$$C^{-1}r^d \leq \Lambda_d(E \cap B(x, r)) \leq Cr^d.$$

Then with the Euclidean metric and  $\mu = \Lambda_d$ ,  $E$  becomes a space of homogeneous type.

**Example 6.** Let  $G$  be a nilpotent group with a left-invariant Riemannian metric  $\rho$ . Let  $\mu$  be the induced measure. Then  $(G, \rho, \mu)$  is a space of homogeneous type.

**Example 7.** Let  $\Omega$  be an open set in  $R^n$  and let  $X_1, \dots, X_k$  be  $C^\infty$  vector fields in  $\Omega$ . Suppose that  $\{X_j\}$  satisfy the condition of Hormander, which is that they, together with all their commutators of all orders, span the tangent space to  $R^n$  at each  $x \in \Omega$ . Say that a Lipschitz curve  $\gamma: [0, 1] \rightarrow \Omega$  is admissible if for almost every  $t$ ,

$$\frac{d\gamma}{dt} = \sum_{j=1}^k c_j(t) X_j(\gamma(t))$$

where  $\sum |c_j(t)|^2 \leq 1$ . Define  $\rho(x, y)$  to be the infimum of the set of all  $r$  for which there exists an admissible curve with  $\gamma(0) = x$  and  $\gamma(r) = y$ . It is a theorem that such an admissible curve exists for any  $x, y \in \Omega$ , provided  $\Omega$  is connected. Then  $\rho$  is clearly a metric. If one stays away from the boundary of  $\Omega$ , then with Lebesgue measure  $d$ ,  $(\Omega, \rho, d)$  becomes a space of homogeneous type.

**Example 8.** This is a special case of Example 7. Let  $U$  be a bounded open subset of  $C^2$  with  $C^\infty$  boundary. At any  $z \in \partial U$  the vector space of linear combinations

$$a_1 \frac{\partial}{\partial \bar{z}_1} + a_2 \frac{\partial}{\partial \bar{z}_2}$$

belonging to the complexified tangent space to  $\partial U$  has dimension 1 over  $C$ .

Fix a nonvanishing  $C^\infty$  complex vector field

$$\bar{L}(z) = a_1(z) \frac{\partial}{\partial \bar{z}_1} + a_2(z) \frac{\partial}{\partial \bar{z}_2}$$

which is tangent to  $\partial U$  at every point  $z$  in some open set  $V \subseteq \partial U$ . Write  $\bar{L} = X + iY$  where  $X, Y$  are real vector fields. If  $U$  is strictly pseudoconvex, then the pair  $X, Y$  satisfies the Hormander condition and  $V$  becomes a space of homogeneous type.

We set down the basic definitions and results concerning the Calderon-Zygmund operator theory on spaces of homogeneous type. The first ingredient of the general theory is of course the analogue of the maximal function of Hardy and Littlewood.

**Definition 5.8:** For  $f \in L^1_{loc}(X)$ , the maximal function of  $f$  is defined by

$$Mf(x) = \sup_{r>0} \frac{1}{\mu(B(x, r))} \int_{B(x, r)} |f(y)| d\mu(y). \quad (5.9)$$

**Theorem 5.10:**  $M$  is bounded on  $L^p(X)$  for all  $1 < p < \infty$ , and is of weak type  $(1, 1)$ .

We also have the Calderon-Zygmund decomposition.

**Theorem 5.11:** Let  $f \in L^1(X)$  and  $\lambda > 0$ , and assume that  $\mu(X) > \lambda^{-1} \|f\|_1$ . Then  $f$  may be decomposed as  $f = g + b$  where

$$\|g\|_2^2 \leq C\lambda \|f\|_1, \quad (5.12)$$

$$b = \sum_j b_j \quad (5.13)$$

where each  $b_j$  is supported on some ball  $B(x_j, r_j)$ ,

$$\int b_j(x) d\mu(x) = 0, \quad (5.14)$$

$$\|b_j\|_1 \leq C\lambda \mu(B(x_j, r_j)), \quad (5.15)$$

$$\sum_j \mu(B(x_j, r_j)) \leq C\lambda^{-1} \|f\|_1. \quad (5.16)$$

**Definition 5.17:** A kernel  $K : X \times X \setminus \{x = y\} \rightarrow C$  is said to be a singular integral kernel if there exist  $\epsilon > 0$  and  $C < \infty$  such that for all  $x \neq y \in X$  and  $z$  with  $\rho(x, z) \leq \frac{1}{2A} \rho(x, y)$ ,

$$|K(x, y)| \leq C\mu(B(x, r))^{-1} \quad (5.18)$$

where  $r = \rho(x, y)$ ,

$$|K(x, y) - K(z, y)| + |K(y, x) - K(y, z)| \leq C \left( \frac{\rho(x, z)}{\rho(x, y)} \right)^\epsilon \mu(B(x, r))^{-1} \quad (5.19)$$

where  $r = \rho(x, y)$ .

**Definition 5.20:** A continuous linear operator  $T : \Lambda_\delta \rightarrow (\Lambda_\delta)'$ , where  $\Lambda_\delta$  is the space of all bounded function on  $X$ , Holder continuous of order  $\delta > 0$  with respect to  $\rho$  and having

bounded support, and  $(\Lambda_\delta)'$  is its dual space, is said to be a singular integral operator if it is associated with a singular integral kernel  $K$  such that

$$\langle Tf, g \rangle = \int \int K(x, y) f(y) g(x) d\mu(y) d\mu(x) \quad (5.21)$$

for all  $f, g \in \Lambda_\delta$  whose supports are separated by a positive distance.

**Theorem 5.22:** Any singular integral operator which is bounded on  $L^2$  is also bounded on  $L^p$ ,  $1 < p < \infty$ , and is of weak type  $(1, 1)$ .

**Definition 5.23:** If  $f \in L^1_{loc}(X)$ , then

$$\|f\|_{BMO} = \sup_{x,r} \inf_c \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f(y) - c| d\mu(y)$$

and BMO is the set all all equivalence classes modulo constants of locally integrable functions with finite BMO norm.

**Theorem 5.24:** A singular integral operator which is bounded on  $L^2$  maps  $L^\infty$  boundedly to BMO.

To formulate the  $T1$  theorem, we need the notion of weak boundedness. For  $\delta \in (0, 1]$ ,  $x \in X$  and  $r > 0$ , define  $A(\delta, x, r)$  to be a set of all  $\phi \in \Lambda_\delta$  supported in  $B(x, r)$  satisfying  $\|\phi\|_\infty \leq 1$  and  $|\phi(y) - \phi(z)| \leq r^{-\delta} \rho(y, z)^\delta$  for all  $y, z \in X$ .

**Definition 5.25:** A singular integral operator  $T$  is weak bounded if there exist  $\delta \in (0, \delta_0]$  and  $C < \infty$  such that for all  $x \in X, r > 0$ , and  $\phi, \psi \in A(\delta, x, r)$

$$|\langle T\phi, \psi \rangle| \leq C\mu(B(x, r)). \quad (5.26)$$

Of course, the inequality would be an immediate consequence of  $L^2$  boundedness. It hold automatically for singular integral operators with antisymmetric kernels.

**Theorem 5.27:** A singular integral operator  $T$  on a space of homogeneous type is bounded on  $L^2$  if and only if it is weak bounded and  $T1, T^*1 \in BMO$ .

We also have analogue of dyadic cubes for a space of homogeneous type.

**Theorem 5.28:** let  $(X, \rho, d\mu)$  be a space of homogeneous type. Then there exists a family of subsets  $Q_\alpha^k \subseteq X$ , defined for all integers  $k$ , and constants  $\delta, \epsilon > 0, C < \infty$  such that

$$\mu(X \setminus \cup_\alpha Q_\alpha^k) = 0 \quad (5.29)$$

for all  $k$ ,

$$\text{for any } \alpha, \beta, k, \text{ with } \ell \geq k, \text{ either } Q_\beta^\ell \subseteq Q_\alpha^k \text{ or } Q_\beta^\ell \cap Q_\alpha^k = \emptyset, \quad (5.30)$$

$$\text{each } Q_\beta^k \text{ has exactly one parent for all } k \geq 1, \quad (5.31)$$

each  $Q_\beta^k$  has at least one child, (5.32)

$$Q_\alpha^{k+1} \subseteq Q_\beta^k \text{ then } \mu(Q_\alpha^{k+1}) \geq \epsilon \mu(Q_\beta^k), \quad (5.33)$$

for each  $(\alpha, k)$  there exists  $x_{\alpha,k} \in X$  such that  $B(x_{\alpha,k}, \delta^k) \subseteq Q_\alpha^k \subseteq B(x_{\alpha,k}, C\delta^k)$ , (5.34)

$$\mu\{y \in Q_\alpha^k : \rho(y, X \setminus Q_\alpha^k) \leq t\delta^k\} \leq Ct^\epsilon \mu(Q_\alpha^k) \text{ for } 0 < t \leq 1 \text{ and all } \alpha, k. \quad (5.35)$$

For most of the examples of spaces of homogeneous type we indicate natural examples of associated singular integral operators.

**Example 1.** Compact manifolds. Any pseudo differential operators with symbol in the class  $S_{1,0}^0$ , on a compact  $C^\infty$  manifold without boundary.  $L^2$  boundedness always holds.

**Example 2.** Parabolic dilations in  $R^n$ . Let  $K$  be a tempered distribution on  $R^n$  with is  $C^\infty$  away from the origin, and satisfies  $K(\delta_t x) = t^{-d}K(x)$  for all  $x \in R^n \setminus \{0\}$  and  $t > 0$ , where  $d = \sum_j a_j$  is called the homogeneous dimension. Then convolution with  $K$  defines a

singular integral operator  $T$ . By translation-invariance,  $T$  maps constant to constant, so that  $T(1) = 0$  in BMO. The same is true for the transpose. It is also easy to verify the weak boundedness. For a special example from differential equations, take  $a_j = 1$  for all  $j \leq n - 1$ , and  $a_n = 2$ . Take  $K$  to be the tempered distribution whose Fourier transform is the function  $\frac{\xi_i \xi_j}{(|\xi'|^2 + i\tau)}$  for some  $i, j \leq n - 1$ , where the Fourier transform variable is written  $\xi = (\xi', \tau) \in R^{n-1} \times R$ . This multiplier arises from regularity estimates for the heat equation in  $n-1$  space variables.

**Example 3.**  $d$ -dimensional subset of Euclidean space. The Cauchy integral on an Ahlfors-regular one-dimensional subset of the complex plane.

**Example 4.** Nilpotent Lie groups. Let  $X_1, \dots, X_k$  be a basis for the linear space of left-invariant vector fields on a simply connected nilpotent Lie group. Let  $L = -\sum_j X_j^2$ , a

left-invariant second-order differential operator. Then  $L^{-1}$  and  $L^{-\frac{1}{2}}$  may be defined by means of spectral theory, and for any  $i$  and  $j$ ,  $X_j X_i L^{-1}$  and  $X_j L^{-\frac{1}{2}}$  are singular integral operators.

In recent years it has been ascertained that central results of classical Calderon-Zygmund theory hold true in very general situations in which the standard doubling condition on the underlying measure is not satisfied. This has come as a great surprise to that homogeneous spaces were not only a convenient setting for developing Calderon-Zygmund theory, but that they were essentially the right context. The motivation of non-homogeneous space come from the so-called problem of analytic capacity. A compact subset  $E$  of the plane is said to be removable for bounded analytic function if given any open set  $\Omega$  and any bounded analytic function  $f$  on  $\Omega \setminus E$ , then  $f$  extends to an analytic function on  $\Omega$ . This can be easily shown to be equivalent to any bounded analytic function on the complement of  $E$  being constant. For example, a set reduced to a point is removable and a closed disc is not. Painleve proved more than one hundred years ago that if  $E$  is a compact set of zero length, then  $E$  is removable. He also asked for a metric, or, even better,

geometric characterization of removable sets. This was called the Painleve problem. If one restricts the attention to subsets of the real line, then the converse of Painleve's result is true, so that the removable subset of the real line are precisely those of zero length.

Denjoy believe that the same would happen if the real axis were replaced by a general rectifiable curve, but the argument he found had a gap. A complete proof of the " Denjoy conjecture did not arrive until Calderon obtained in 1977 his famous result on the  $L^2$  boundedness of the Cauchy integral on Lipschitz graphs with small Lipschitz constant. Notice that the solution of the Denjoy conjecture provides an answer to Painleve's problem for a restricted class of compact sets. In 1967 Vitushkin conjectured that among sets of finite length removability is characterized by pure unrectifiability. That removability implies pure unrectifiability is again a consequence of Calderon's theorem, by essentially the same argument that gives the Denjoy conjecture. The other implication is the difficult one. Roughly speaking, if one wants to proof that a pure unrectifiable set  $E$  is removable, by using at some point Calderon-Zygmund theory, you are in trouble because your set is irregular, that is, homogeneity is missing in this case. To be precise, let  $\mu$  be a positive Borel measure in the plane. Given  $\epsilon > 0$  and a compact supported function in  $L^1(\mu)$ , set

$$C_\epsilon(f\mu)(z) = \int_{|\xi-z|>\epsilon} \frac{f(\xi)}{\xi-z} d\mu(\xi), z \in C.$$

One says that the above Cauchy integral maps boundedly  $L^p(\mu)$  into  $L^p(\mu)$  whenever

$$\|C_\epsilon(f\mu)\|_p \leq C(p)\|f\mu\|_p,$$

where  $C(p)$  is a positive constant independent of  $\epsilon$  and of the compactly supported function  $f \in L^p(\mu)$ .

The key facts that  $\mu$  satisfies only the growth condition

$$\mu(D) \leq C \text{ radius } (D),$$

for each disc  $D$ .

In recent years, the list of results that one can prove without resorting to the doubling condition is amazing: the Calderon-Zygmund decomposition and the derivation of weak type  $(1, 1)$ , and  $L^p$  bounds,  $1 < p < \infty$ , from  $L^2$  bounds, Cotlar's inequality for maximal singular integral, the  $T1$  and  $Tb$  theorems, and many others.