# Topics in Harmonic Analysis
# Lecture 6: Pseudodifferential calculus and almost orthogonality

Po-Lam Yung

The Chinese University of Hong Kong

# Introduction

- While multiplier operators are very useful in studying constant coefficient partial differential equations, one often encounters variable coefficient partial differential equations.
- Thus we consider a variable coefficient generalization of multiplier operators, namely pseudodifferential operators.
- We study compositions and mapping properties of pseudodifferential operators.
- These in turn allow one to construct paramatrices to variable coefficient elliptic PDEs.
- We close this lecture with a beautiful almost orthogonality principle, due to Cotlar and Stein, which will play a crucial role in the proof of the $T(1)$ theorem in the next lecture.

# Outline

- Symbols of pseudodifferential operators
- Kernel representations
- Mapping properties on $L^2$
- Compound symbols
- Closure under adjoints and compositions
- Parametrix construction
- Cotlar-Stein lemma (almost orthogonality)

# Symbols of pseudodifferential operators

- Given a smooth function $a(x, \xi)$ on $\mathbb{R}^n \times \mathbb{R}^n$ (which we think of as the cotangent bundle of $\mathbb{R}^n$), a pseudodifferential operator with symbol $a$ is by definition

$$T_a f(x) = \int_{\mathbb{R}^n} a(x, \xi) \widehat{f}(\xi) e^{2\pi i x \cdot \xi} d\xi.$$

- We will consider only symbols $a$ that satisfy the following differential inequalities:

$$|\partial_x^\alpha \partial_\xi^\beta a(x, \xi)| \lesssim (1 + |\xi|)^{m - |\beta|\gamma + |\alpha|\delta}$$

  for all multiindices $\alpha$ and $\beta$, where $m \in \mathbb{R}$ and $\gamma, \delta \in [0, 1]$ are three fixed parameters.
- Following Hörmander, a symbol $a$ is said to be of class $S^m_{\gamma,\delta}$, if the above differential inequalities are satisfied for all $\alpha$ and $\beta$.
- Usually we consider only the case $\gamma = 1$, $\delta = 0$, in which case we write $S^m$ in place of $S^m_{1,0}$.
- $m$ is called the order of the symbol (or the order of the associated operator).

- Example: If $p(\xi)$ is a polynomial of degree $m$, and

$$a(x,\xi) = p(2\pi i\xi),$$

then $a \in S^m$, and

$$T_a f(x) = p(\partial_x) f(x)$$

is a constant coefficient differential operator of order $m$.

- More generally, if $m \in \mathbb{N}$ and

$$a(x,\xi) = \sum_{|\beta| \leq m} A_\beta(x)(2\pi i\xi)^\beta,$$

where the $A_\beta$'s are all $C_c^\infty$ on $\mathbb{R}^n$, then we have $a \in S^m$ with

$$T_a f(x) = \sum_{|\beta| \leq m} A_\beta(x)\partial_x^\beta f(x)$$

is a variable coefficient partial differential operator of order $m$.

- It is easy to see that if $a \in S^m_{\gamma,\delta}$ for some $m \in \mathbb{R}$, $\gamma, \delta \in [0,1]$, then $T_a$ is a linear map from $\mathcal{S}(R^n)$ into itself, and the map

$$T_a \colon \mathcal{S}(R^n) \to \mathcal{S}(\mathbb{R}^n)$$

is continuous.
- One typical use of pseudodifferential operators is to construct paramatrices (i.e. approximate solutions) to partial differential equations.
- For those we usually need pseudodifferential operators of non-positive orders, which are typically integral operators.
- As before, let $(\mathbb{R}^n \times \mathbb{R}^n)^*$ be $\mathbb{R}^n \times \mathbb{R}^n$ with the diagonal $\{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n \colon x = y\}$ removed.

# Kernel estimates

## Theorem

*Let $-n < m \leq 0$, and $a \in S^m$. Then there exists a function $K_0 \in C^\infty((\mathbb{R}^n \times \mathbb{R}^n)^*)$ such that*

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

*for all $f \in C_c^\infty(\mathbb{R}^n)$ and all $x$ not in the support of $f$. Furthermore,*

$$|\partial_{x,y}^\lambda K_0(x, y)| \lesssim \frac{1}{|x - y|^{n+m+|\lambda|}}$$

*for all multiindices $\lambda$ and all $x \neq y$.*

- Indeed pick a smooth function $\eta$ with compact support on $\mathbb{R}^n$ with $\eta(0) = 1$. For $x \neq y$ and $\varepsilon > 0$, let

$$K_\varepsilon(x, y) := \int_{\mathbb{R}^n} a(x, \xi) \eta(\varepsilon \xi) e^{2\pi i (x-y) \cdot \xi} d\xi.$$

- We claim that $K_\varepsilon(x,y) \in C^\infty((\mathbb{R}^n \times \mathbb{R}^n)^*)$ for all $\varepsilon > 0$, with

$$|\partial_{x,y}^\lambda K_\varepsilon(x,y)| \lesssim \frac{1}{|x-y|^{n+m+|\lambda|}}$$

  for all multiindices $\lambda$ and all $x \neq y$, where the constants are uniform in $\varepsilon > 0$.
- One sees this by splitting the integral depending on whether $|\xi| \leq |x-y|^{-1}$ or not; when $|\xi| > |x-y|^{-1}$, we integrate by parts using

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

  sufficiently many times to gain enough decay in $|\xi|$. This shows

$$|K_\varepsilon(x,y)| \lesssim \frac{1}{|x-y|^{n+m}}$$

  when $x \neq y$, and similarly one can estimate $\partial_{x,y}^\lambda K_\varepsilon(x,y)$.

- Furthermore, by a similar argument, $K_\varepsilon(x,y)$ converges locally uniformly on $(\mathbb{R}^n \times \mathbb{R}^n)^*$ as $\varepsilon \to 0^+$, and so do $\partial^\lambda_{x,y} K_\varepsilon(x,y)$ for all multiindices $\lambda$.
- For $x \neq y$, let
$$K_0(x,y) := \lim_{\varepsilon \to 0^+} K_\varepsilon(x,y) \in C^\infty((\mathbb{R}^n \times \mathbb{R}^n)^*).$$
- Now note that if $f \in \mathcal{S}(\mathbb{R}^n)$ and $x \in \mathbb{R}^n$, then
$$\begin{aligned} T_a f(x) &= \lim_{\varepsilon \to 0^+} \int_{\mathbb{R}^n} a(x,\xi)\eta(\varepsilon\xi)\widehat{f}(\xi) e^{2\pi i x\cdot\xi} d\xi \\ &= \lim_{\varepsilon \to 0^+} \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} a(x,\xi)\eta(\varepsilon\xi) f(y) e^{2\pi i (x-y)\cdot\xi} dy d\xi \\ &= \lim_{\varepsilon \to 0^+} \int_{\mathbb{R}^n} f(y) K_\varepsilon(x,y) dy. \end{aligned}$$
- If in addition $f \in C^\infty_c(\mathbb{R}^n)$, and $x$ is not in the support of $f$, then the last line is equal to
$$T_a f(x) = \int_{\mathbb{R}^n} f(y) K_0(x,y) dy$$
by the dominated convergence theorem.

- This establishes the desired kernel representation formula for $T_a f(x)$.
- The estimates for $\partial^\lambda_{x,y} K_0(x,y)$ on $(\mathbb{R}^n \times \mathbb{R}^n)^*$ follow from the corresponding uniform estimates for $\partial^\lambda_{x,y} K_\varepsilon(x,y)$.
- We remark that if $|x-y| \gtrsim 1$, the above proof also shows that

$$|\partial^\lambda_{x,y} K_0(x,y)| \lesssim |x-y|^{-N}$$

  for any multiindices $\lambda$ and any $N \in \mathbb{N}$ (i.e. we get rapid decay as $|x-y| \to +\infty$).
- This is closely tied to the *pseudolocality* of psuedodifferential operators: indeed a linear operator $T\colon \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$ is local, if the support of $Tf$ is contained in the support of $f$ for every $f \in \mathcal{S}(\mathbb{R}^n)$. This is the case if the Schwartz kernel of $T$ is supported on the diagonal $\{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n \colon x = y\}$.
- While $K_0(x,y)$ is not supported on the diagonal $\{(x,y) \in \mathbb{R}^n \times \mathbb{R}^n \colon x = y\}$, the above decay of $K_0(x,y)$ away from the diagonal is a close substitute for it.

# Mapping properties on $L^2$

- We now focus on pseudodifferential operators of order 0.

**Theorem**

*Let $a \in S^0$. Then $T_a$ extends to a bounded operator on $L^2(\mathbb{R}^n)$.*

- In view of the kernel representation theorem above, and the variable coefficient singular integral theorem from Lecture 4, this establishes the following corollary.

**Corollary**

*Let $a \in S^0$. Then $T_a$ extends to a bounded operator on $L^p(\mathbb{R}^n)$ for all $1 < p < \infty$.*

- One direct proof of the theorem proceeds via *pseudolocality*.
- For $j \in \mathbb{Z}^n$, let $B_j$ be the open ball of radius 2 centered at $j$. Then $\{B_j\}_{j \in \mathbb{Z}^n}$ covers $\mathbb{R}^n$.
- Let $1 = \sum_j \phi_j^2$ be a smooth partition of unity subordinate to the above cover, so that $\phi_j \in C_c^\infty(B_j)$ for every $j$.
- Then

$$\begin{aligned} \|T_a f\|_{L^2}^2 &= \sum_j \|\phi_j T_a f\|_{L^2}^2 \\ &= \sum_j \|\phi_j T_a(\chi_{2B_j} f)\|_{L^2}^2 + \sum_j \|\phi_j T_a(\chi_{(2B_j)^c} f)\|_{L^2}^2. \end{aligned}$$

- From our earlier kernel estimates when $|x-y| \gtrsim 1$, we get

$$|\phi_j(x) T_a(\chi_{(2B_j)^c} f)(x)| \lesssim \chi_{B_j}(x) \int_{y \notin 2B_j} |f(y)||x-y|^{-N} dy,$$

so choosing $N > n$ and using Cauchy-Schwarz, we get

$$|\phi_j(x) T_a(\chi_{(2B_j)^c} f)(x)|^2 \lesssim \chi_{B_j}(x) \int_{y \notin 2B_j} |f(y)|^2 |x-y|^{-N} dy.$$

Integrating both sides gives

$$\|\phi_j T_a(\chi_{(2B_j)^c} f)\|_{L^2}^2 \lesssim \int_{|y-j| \gtrsim 1} |f(y)|^2 |j-y|^{-N} dy,$$

so summing over $j$ gives

$$\sum_j \|\phi_j T_a(\chi_{(2B_j)^c} f)\|_{L^2}^2 \lesssim \sum_j \int_{|y-j| \gtrsim 1} |f(y)|^2 |j-y|^{-N} dy \lesssim \|f\|_{L^2}^2.$$

- It remains to show

$$\sum_j \|\phi_j T_a(\chi_{2B_j} f)\|_{L^2}^2 \lesssim \|f\|_{L^2}^2.$$

  This follows if we can show

$$\|\phi_j T_a\|_{L^2 \to L^2} \lesssim 1$$

  since $\sum_j \|\chi_{2B_j} f\|_{L^2}^2 \lesssim \|f\|_{L^2}^2$.
- But $\phi_j T_a$ is a pseudodifferential operator with symbol $\phi_j(x)a(x,\xi)$. The latter is just another symbol in $S^m$, except now it has compact $x$-support inside some ball of radius 2.
- Hence it remains to prove our theorem for $a \in S^0$, under the additional assumption that $a(x,\xi)$ has compact $x$-support inside some unit cube.
- This we obtain by expanding $a(x,\xi)$ as Fourier series in $x$.

- Without loss of generality, suppose $a(x,\xi) \in S^0$ and has compact $x$-support on the unit cube $B$ centered at 0. Then

$$a(x,\xi) = \sum_{\eta \in \mathbb{Z}^n} \widehat{a}(\eta,\xi) e^{2\pi i \eta \cdot x}$$

where $\widehat{a}$ is the Fourier transform of $a$ in the first variable. Thus

$$\begin{aligned} T_a f(x) &= \sum_{\eta \in \mathbb{Z}^n} e^{2\pi i \eta \cdot x} \int_{\mathbb{R}^n} \widehat{a}(\eta,\xi) \widehat{f}(\xi) e^{2\pi i x \cdot \xi} d\xi \\ &= \sum_{\eta \in \mathbb{Z}^n} \frac{e^{2\pi i \eta \cdot x}}{(1 + 4\pi^2 |\eta|^2)^n} \int_{\mathbb{R}^n} \widehat{\Delta_x^n a}(\eta,\xi) \widehat{f}(\xi) e^{2\pi i x \cdot \xi} d\xi. \end{aligned}$$

- But $\widehat{\Delta_x^n a}(\eta,\xi)$ is a bounded multiplier on $L^2$ uniformly in $\eta$.
- Thus triangle inequality gives

$$\|T_a f\|_{L^2} \lesssim \sum_{\eta \in \mathbb{Z}^n} (1 + |\eta|)^{-2n} \|f\|_{L^2} \lesssim \|f\|_{L^2},$$

which finishes the proof of the Theorem.

## Compound symbols

- We will turn soon to the adjoints and compositions of pseudodifferential operators whose symbols are in $S^m$ for some $m \in \mathbb{R}$.
- A convenient tool is the concept of compound symbols.
- Let $m \in \mathbb{R}$, $\gamma \in [0,1]$, $\delta \in [0,1)$. A compound symbol of class $CS^m_{\gamma,\delta}$ is a smooth function $c(x,y,\xi)$ on $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ where

$$|\partial^\alpha_{x,y} \partial^\beta_\xi c(x,y,\xi)| \lesssim (1+|\xi|)^{m-|\beta|\gamma+|\alpha|\delta}$$

  for all multiindices $\alpha$ and $\beta$.
- To every $c \in CS^m_{\gamma,\delta}$, we associate an operator $T_{[c]}$ on $\mathcal{S}(\mathbb{R}^n)$ by

$$T_{[c]} f(x) = \lim_{\varepsilon \to 0^+} T_{[c],\varepsilon} f(x), \quad \text{where}$$

$$T_{[c],\varepsilon} f(x) := \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} c(x,y,\xi) \eta(\varepsilon\xi) f(y) e^{2\pi i (x-y)\cdot\xi} d\xi dy$$

  and $\eta$ is a fixed function in $C^\infty_c(\mathbb{R}^n)$ with $\eta(0)=1$.

- Since $\delta < 1$, one can show that for $f \in \mathcal{S}(\mathbb{R}^n)$, $T_{[c],\varepsilon} f$ defines a Schwartz function on $\mathbb{R}^n$ for every $\varepsilon > 0$, and that $T_{[c],\varepsilon} f$ converges in the topology of $\mathcal{S}(\mathbb{R}^n)$ as $\varepsilon \to 0^+$. Indeed this follows from multiple integrating by parts via

$$e^{2\pi i(x-y)\cdot\xi} = \frac{(I - \Delta_y)e^{2\pi i(x-y)\cdot\xi}}{1 + 4\pi^2|\xi|^2}.$$

- Thus $T_{[c]}$ defines a linear mapping

$$T_{[c]} \colon \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n);$$

  it is easy to check that this map is also continuous.
- We are mainly interested in $CS^m_{\gamma,\delta}$ when $\gamma = 1$ and $\delta = 0$.
- We write $CS^m$ for $CS^m_{1,0}$ if $m \in \mathbb{R}$.
- The main theorem about compound symbols is the following:

### Theorem

*If $c \in CS^m$ for some $m \in \mathbb{R}$, then there exists $a \in S^m$ such that*

$$T_{[c]}f = T_a f$$

*for all $f \in \mathcal{S}(\mathbb{R}^n)$. Also, we have the asymptotic expansion*

$$a(x,\xi) \sim \sum_\gamma \frac{(2\pi i)^{-|\gamma|}}{\gamma!} \partial_y^\gamma \partial_\xi^\gamma \, c(x,y,\xi)|_{y=x}\,,$$

*in the sense that the sum of those terms with $|\gamma| < N$ on the right hand side differ from $a(x,\xi)$ by a symbol in $S^{m-N}$ for all $N \in \mathbb{N}$.*

- A proof is outlined in Homework 6.
- We note that different compound symbols $c$ may give rise to the same symbol $a$ in the above theorem.
- In particular, the map $c \mapsto T_{[c]}$ is not injective.
- We will use this non-injectivity to our advantage in what follows.

# Closure under adjoints and compositions

- We will prove two theorems using compound symbols.

## Theorem

*Let $m \in \mathbb{R}$. If $a \in S^m$, then there exists a symbol $a^* \in S^m$, such that the formal adjoint of $T_a$ is $T_{a^*}$, in the sense that*

$$\int_{\mathbb{R}^n} T_a f(x) \overline{g(x)} dx = \int_{\mathbb{R}^n} f(x) \overline{T_{a^*} g(x)} dx \quad \text{for all } f, g \in \mathcal{S}(\mathbb{R}^n).$$

*Also, we have the asymptotic expansion*

$$a^*(x,\xi) \sim \sum_{\gamma} \frac{(2\pi i)^{-|\gamma|}}{\gamma!} \partial_y^\gamma \partial_\xi^\gamma \, \overline{a(y,\xi)} \Big|_{y=x}.$$

*In particular,*

$$a^*(x,\xi) = \overline{a(x,\xi)} \quad (\text{mod } S^{m-1}).$$

### Theorem

*Let $m_1, m_2 \in \mathbb{R}$. If $a_1 \in S^{m_1}$ and $a_2 \in S^{m_2}$, then there exists a symbol $a \in S^{m_1+m_2}$, such that*

$$T_{a_1} T_{a_2} f = T_a f \quad \text{for all } f \in \mathcal{S}(\mathbb{R}^n).$$

*Also, we have the asymptotic expansion*

$$a(x,\xi) \sim \sum_{\gamma} \frac{(2\pi i)^{-|\gamma|}}{\gamma!} \partial_\xi^\gamma a_1(x,\xi) \partial_x^\gamma a_2(x,\xi).$$

*In particular,*

$$a(x,\xi) = a_1(x,\xi) a_2(x,\xi) \quad (\text{mod } S^{m_1+m_2-1}).$$

- Indeed, let $m \in \mathbb{R}$, $a \in S^m$ and $f, g \in \mathcal{S}(\mathbb{R}^n)$. Let $\eta \in C_c^\infty$ on $\mathbb{R}^n$ with $\eta(0) = 1$. Then by dominated convergence,

$$\int_{\mathbb{R}^n} T_a f(x)\overline{g(x)}dx = \lim_{\varepsilon \to 0^+} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} a(x,\xi)\eta(\varepsilon\xi)\widehat{f}(\xi)\overline{g(x)}e^{2\pi i x\cdot\xi}d\xi dx.$$

- Writing $c(x,y,\xi) := \overline{a(y,\xi)}$, the above limit is equal to

$$\begin{aligned} &\lim_{\varepsilon \to 0^+} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} a(x,\xi)\eta(\varepsilon\xi)f(y)\overline{g(x)}e^{2\pi i(x-y)\cdot\xi}dyd\xi dx \\ = &\lim_{\varepsilon \to 0^+} \int_{\mathbb{R}^n} f(y)\overline{T_{[c],\varepsilon}g(y)}dy. \end{aligned}$$

- Since $T_{[c],\varepsilon}g$ converges to $T_{[c]}g$ in $\mathcal{S}(\mathbb{R}^n)$ as $\varepsilon \to 0^+$, the above limit is just

$$\int_{\mathbb{R}^n} f(y)\overline{T_{[c]}g(y)}dy.$$

- It remains to write $T_{[c]}$ as $T_{a^*}$ for some $a^* \in S^m$, using our previous theorem about compound symbols.

- Next, let $m_1, m_2 \in \mathbb{R}$, $a_1 \in S^{m_1}$, $a_2 \in S^{m_2}$.
- By the previous theorem, there exists $a_2^* \in S^{m_2}$, such that the formal adjoint of $T_{a_2^*}$ is $T_{a_2}$, which in view of the computation on the previous page implies

$$T_{a_2} f(x) = \lim_{\varepsilon \to 0^+} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \overline{a_2^*(y,\xi)} \eta(\varepsilon\xi) f(y) e^{2\pi i (x-y)\cdot\xi} dy d\xi.$$

  for all $f \in \mathcal{S}(\mathbb{R}^n)$ and all $x \in \mathbb{R}^n$ (limit taken in $\mathcal{S}(\mathbb{R}^n)$).
- Hence $T_{a_1} T_{a_2} f(x)$ is given by

$$\lim_{\varepsilon \to 0^+} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} a_1(x,\xi) \overline{a_2^*(y,\xi)} \eta(\varepsilon\xi) f(y) e^{2\pi i (x-y)\cdot\xi} dy d\xi.$$

- The latter is $\lim_{\varepsilon \to 0^+} T_{[c],\varepsilon} f(x)$ if

$$c(x,y,\xi) := a_1(x,\xi) \overline{a_2(y,\xi)}.$$

  Since such $c \in CS^{m_1+m_2}$, by the previous theorem about compound symbols, there exists $a \in S^{m_1+m_2}$, such that the above limit is equal to $T_a f(x)$.

# Parametrix construction

- Let $m \in \mathbb{N}$. Let

$$P(D) = \sum_{|\alpha| \leq m} p_\alpha(x) \partial_x^\alpha$$

  be a differential operator of order $m$ with $C^\infty$ coefficients.
- $P(D)$ is a pseudodifferential operator with symbol

$$p(x, \xi) := \sum_{|\alpha| \leq m} p_\alpha(x)(2\pi i \xi)^\alpha.$$

- It is said to be elliptic, if there exists a constant $C > 0$, such that

$$|p(x, \xi)| \geq C|\xi|^m$$

  for all $x \in \mathbb{R}^n$ and all $\xi$ with $|\xi| \geq 1$.
- We use the following theorem to construct parametrices of such elliptic partial differential operators.

**Theorem**

*Let $m \in \mathbb{R}$. Given a sequence of symbols $a_0, a_1, \ldots,$ with*

$$a_k \in S^{m-k} \quad \text{for every } k \geq 0,$$

*then there exists a symbol $a \in S^m$, such that for every $N \in \mathbb{N}$, there exists $e_N \in S^{m-N}$ with*

$$a(x,\xi) = \sum_{k=0}^{N-1} a_k(x,\xi) + e_N(x,\xi).$$

- See Homework 6 for its proof.

- Let $P(D)$ be an elliptic partial differential operator of order $m$ with $C^\infty$ coefficients. Let $p(x,\xi)$ be its symbol.
- Let $\varphi(\xi)$ be a smooth function that is identically 0 on $B(0,1)$, and identically 1 outside $B(0,2)$.
- Let $T_{a_0}$ be the pseudodifferential operator of order $-m$ with symbol

$$a_0(x,\xi) := \frac{\varphi(\xi)}{p(x,\xi)}.$$

- Then the composition theorem shows that

$$P(D)T_{a_0} = I - E_{-1}$$

for some pseudodifferential operator of order $-1$.
- We compose both sides on the right by $E_{-1}^k$ (where $k \in \mathbb{N}$), and get

$$P(D)T_{a_0}E_{-1}^k = E_{-1}^k - E_{-1}^{k+1}.$$

- Summing over $k$ and telescoping, we get

$$P(D)[T_{a_0} + T_{a_0}E_{-1} + \cdots + T_{a_0}E_{-1}^N] = I - E_{-1}^{N+1}$$

for any $N \in \mathbb{N}$.

- Using the composition theorem again, for any $k \in \mathbb{N}$, there exists a symbol $a_k \in S^{-m-k}$, such that

$$T_{a_0} E_{-1}^k = T_{a_k}.$$

- From the previous theorem, there exists a symbol $a \in S^{-m}$, such that for every $N \in \mathbb{N}$, there exists $e_N \in S^{-m-N}$ such that

$$T_a = \sum_{k=0}^{N-1} T_{a_k} + T_{e_N}.$$

- Let $T_e$ be the pseudodifferential operator defined by

$$P(D) T_a = I + T_e.$$

- The calculation on the previous slide shows that $e \in S^{-m-N}$ for any $N \in \mathbb{N}$.
- In this sense $T_a$ is an approximate solution to $P(D)$, aka a parametrix for $P(D)$.

# Cotlar-Stein lemma (almost orthogonality)

- Earlier we proved $L^2$ boundedness of psuedodifferential operators with symbols in $S^0$ by using the Fourier transform.
- We now describe another important tool about establishing $L^2$ boundedness of linear operators, namely Cotlar-Stein lemma.
- This can be used to prove a theorem of Calderón and Vaillaincourt, namely $T_a$ is bounded on $L^2$ whenever $a \in S^0_{\gamma,\gamma}$ for all $\gamma \in [0, 1)$ (see Homework 6); in particular, this recovers the $L^2$ boundedenss of psuedodifferential operators with symbols in $S^0$.
- The Cotlar-Stein lemma also plays a key role in the proof of many celebrated theorems.
- We will prove a proposition this time, which will play a crucial role in the proof of the $T(1)$ theorem in the next lecture.

## Theorem (Cotlar-Stein)

*Suppose $\{T_j\}$ is a sequence of bounded linear operators between two Hilbert spaces. If there exist constants $A$ and $B$ such that*

$$\sup_j \sum_i \|T_j T_i^*\|^{1/2} \leq A \quad \textit{and} \quad \sup_i \sum_j \|T_i^* T_j\|^{1/2} \leq B$$

*for all $i, j$, then $\sum_j T_j$ converges strongly to a bounded linear operator $T$ between the two Hilbert spaces, with $\|T\| \leq \sqrt{AB}$.*

- This is called an almost orthogonality lemma, because one situation where the hypothesis are fulfilled are when all $\|T_j\| \leq B$, the images of the different $T_j$'s are orthogonal, and the images of the different $T_i^*$'s are orthogonal. (Indeed then $T_i^* T_j = 0$ and $T_j T_i^* = 0$ whenever $j \neq i$.)
- The proof of Cotlar-Stein involves a classic application of the *tensor power trick*.

- To prove Cotlar-Stein, suppose first $\{T_j\}$ is a finite sequence (say $J$ terms). Then for any positive integer $N$,

$$\|T\|^{2N} = \left\|(TT^*)^N\right\| = \sum_{j_1,\dots,j_{2N}} \|T_{j_1} T_{j_2}^* T_{j_3} T_{j_4}^* \dots T_{j_{2N-1}} T_{j_{2N}}^*\|$$

- But the summand is bounded by both

$$\|T_{j_1} T_{j_2}^*\| \, \|T_{j_3} T_{j_4}^*\| \dots \|T_{j_{2N-1}} T_{j_{2N}}^*\|$$

and

$$\|T_{j_1}\| \, \|T_{j_2}^* T_{j_3}\| \dots \left\|T_{j_{2N-2}}^* T_{j_{2N-1}}\right\| \, \|T_{j_{2N}}^*\| \, .$$

- So taking geometric average, the sum is bounded by

$$\sum_{j_1,\dots,j_{2N}} \|T_{j_1}\|^{1/2} \, \|T_{j_1} T_{j_2}^*\|^{1/2} \, \|T_{j_2}^* T_{j_3}\|^{1/2} \dots \|T_{j_{2N-1}} T_{j_{2N}}^*\|^{1/2} \, \|T_{j_{2N}}^*\|^{1/2}$$

$$\leq J \max_{j_1} \|T_{j_1}\|^{1/2} A^N B^{N-1} \max_{j_{2N}} \|T_{j_{2N}}\|^{1/2}$$

- Taking $2N$-th root and letting $N \to +\infty$ yields $\|T\| \leq \sqrt{AB}$.

- The general case when we have infinitely many operators $T_j$ follows once we have the following lemma:

**Lemma**

*Let $\{f_j\}$ be a sequence in a Hilbert space, and $A \in \mathbb{R}$. Suppose for any sequence $\{\varepsilon_j\}$ that has only finitely many non-zero terms, and that satisfies $|\varepsilon_j| \leq 1$ for all $j$, we have*

$$\left\| \sum_j \varepsilon_j f_j \right\| \leq A.$$

*Then $\sum_j f_j$ converges in the Hilbert space.*

- See Homework 6 for the proof of this lemma.
- We close this lecture by the following application of Cotlar-Stein.

## Proposition

*Suppose $\delta > 0$. Let $\{k_j(x,y)\}_{j\in\mathbb{Z}}$ be a sequence of $C^\infty$ kernels on $\mathbb{R}^n \times \mathbb{R}^n$ that satisfies*

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

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

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

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

*For each $j \in \mathbb{Z}$, let $T_j f(x) = \int_{\mathbb{R}^n} f(y)k_j(x,y)dy$. Then*

$$\|T_j T_i^*\|_{L^2\to L^2} + \|T_i^* T_j\|_{L^2\to L^2} \lesssim 2^{-\delta|i-j|},$$

*so $\sum_j T_j$ converges strongly to a bounded linear operator $T$ on $L^2$.*

- One can use this to prove, for instance, that the Hilbert transform is bounded on $L^2$; see Homework 6.
- Indeed, morally speaking, (1) to (4) says that $\sum_j T_j$ is almost like a translation-invariant singular integral we studied in Lecture 4.
- Next time we use this proposition to prove the $T(1)$ theorem.
- We prove this proposition as follows.
- We will only show that

$$\|T_j T_i^*\|_{L^2 \to L^2} \lesssim 2^{-\delta|i-j|}$$

since the bound for $\|T_i^* T_j\|_{L^2 \to L^2}$ is similar.
- The kernel for $T_j T_i^*$ is

$$K_{j,i}(x,y) := \int_{\mathbb{R}^n} k_j(x,z)\overline{k_i}(y,z)dz,$$

and if $i \geq j$, then we rewrite this using (3) as

$$K_{j,i}(x,y) = \int_{\mathbb{R}^n} (k_j(x,z) - k_j(x,y))\overline{k_i}(y,z)dz.$$

- From (1) and (2), we get

$$
\begin{aligned}
&|k_j(x,z) - k_j(x,y)| \\
\lesssim &(2^j|z-y|)^{\frac{\delta}{2}} \left( \frac{2^{jn}}{(1+2^j|x-z|)^{n+\delta}} + \frac{2^{jn}}{(1+2^j|x-y|)^{n+\delta}} \right);
\end{aligned}
$$

indeed if $2^j|y-z| \geq 1$, this follows by estimating

$$
|k_j(x,z) - k_j(x,y)| \leq |k_j(x,z)| + |k_j(x,y)|
$$

and using (1), whereas if $2^j|y-z| \leq 1$, then we use (2) and the mean-value theorem, noting that if $u$ is on the straight line segment connecting $y$ and $z$, then

$$
\min\{|x-y|, |x-z|\} \lesssim 2^{-j} + |x-u|,
$$

which implies

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

- Now we invoke an elementary inequality, which says

$$\int_{\mathbb{R}^n} \frac{2^{jn}}{(1+2^j|x-z|)^{n+\frac{\delta}{2}}} \frac{2^{in}}{(1+2^i|z|)^{n+\frac{\delta}{2}}} dz \lesssim \frac{2^{jn}}{(1+2^j|x|)^{n+\frac{\delta}{2}}}$$

  if $i \geq j$ (this can be proved by rescaling to $j = 0$, and then dividing the domain of integration into 2 parts, depending on whether $|z-x| \geq \frac{|x|}{2}$ or $|z-x| \leq \frac{|x|}{2}$).
- Hence if $i \geq j$, then

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

  similarly if $j \geq i$, then

$$|K_{j,i}(x,y)| \lesssim 2^{-(j-i)\frac{\delta}{2}} \frac{2^{in}}{(1+2^i|x-y|)^{n+\frac{\delta}{2}}}.$$

- From Young's inequality, it follows that

$$\|T_j T_i^*\|_{L^2 \to L^2} \lesssim 2^{-|i-j|\frac{\delta}{2}}.$$