LECTURE 12: L^2-THEORY OF SEMICLASSICAL PsDOs: COMPACTNESS

1. Compactness of semiclassical PsDOs

¶ Compactness operators.

As we have explained at the very beginning of this course, in the quantum part of the theory, the quantum energies should be discrete. For this purpose, we need to find a criteria for a semiclassical pseudodifferential operator $\widehat{a}^W$ to have discrete spectrum. One very useful way to prove discreteness of spectrum that we learned in functional analysis is through compact operators. Recall:

$$\begin{aligned} &\text{A bounded linear operator } A \in \mathcal{L}(H_1, H_2) \text{ is called a } \textit{compact operator} \\ \Longleftrightarrow\ & A \text{ maps bounded sets in } H_1 \text{ into pre-compact subsets in } H_2 \\ \Longleftrightarrow\ & \text{For every bounded sequence } \{x_k\} \text{ in } H_1, \text{ the image } \{Ax_k\} \text{ contains a} \\ & \text{convergent subsequence.} \end{aligned}$$

For example, if $A \in \mathcal{L}(H_1, H_2)$ has a finite dimensional range, then A is compact. Such operators are called *finite rank operators.* Compact operators are very useful in spectral theory because

- For any compact operator A, the spectrum $\sigma(A)$ consists of eigenvalues with finite multiplicities, and the only possible accumulation point is 0.
- For a self-adjoint operator A to have discrete spectrum, it is enough to prove the compactness of the resolvent $(A - i)^{-1}$.

We list several important facts for compact operators:

- The set of compact operators, $\mathcal{K}(H_1, H_2)$, is a closed vector subspace of $\mathcal{L}(H_1, H_2)$ (with respect to the operator norm topology).
- Any compact operator can be approximated by a sequence of finite rank operators.
- (Schauder) An operator $A \in \mathcal{L}(H_1, H_2)$ is compact if and only if A^* is compact.

¶ Compactness for Schwartz symbols.

We would like to show that if the symbol a is a Schwartz function, then $\widehat{a}^W$ is a compact operator on $L^2(\mathbb{R}^n)$. During the proof we will need to control the L^∞-norm of various functions of the form $x^\alpha \partial_x^\beta \widehat{a}^W u(x)$. So we first prove

Lemma 1.1. *If $a \in \mathscr{S}(\mathbb{R}^{2n})$, then for any multi-indices α and β, there exists constant $C_{\alpha,\beta,n}$ so that for any $u \in L^2(\mathbb{R}^n)$,*

$$\sup_{x\in\mathbb{R}^n} |x^\alpha \partial_x^\beta\ \widehat{a}^W u| \le C_{\alpha,\beta,n}\|u\|_{L^2}. \tag{1}$$

Proof. The Schwartz kernel of $\widehat{a}^W$ is

$$K(x,y) = \frac{1}{(2\pi\hbar)^n}\int_{\mathbb{R}^n} e^{\frac{i}{\hbar}(x-y)\cdot\xi} a(\frac{x+y}{2},\xi)d\xi.$$

Since $a \in \mathscr{S}$, we have $K \in \mathscr{S}$. So for any multi-index α and β,

$$\sup_{x\in\mathbb{R}^n}|x^\alpha\partial_x^\beta\ \widehat{a}^W u| \le \sup_{(x,y)\in\mathbb{R}^{2n}} |x^\alpha \partial_x^\beta \langle y\rangle^n K(x,y)| \int_{\mathbb{R}^n}\langle y\rangle^{-n}|u(y)|dy \le C_{\alpha,\beta,n}\|u\|_{L^2}.$$

□

Now we prove the compactness of $\widehat{a}^W$ for a Schwartz symbol a:

Theorem 1.2. *Suppose $a = a(x,\xi) \in \mathscr{S}(\mathbb{R}^{2n})$. Then*

$$\widehat{a}^W : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$$

is a compact operator.

Proof. Let $\mathcal{B} \subset L^2(\mathbb{R}^n)$ be a bounded set. We need to find a sequence $\{\varphi_k\} \subset \mathcal{B}$ so that $\widehat{a}^W\varphi_k$ converges in L^2. We fix $N > \frac{n}{2}$ and define

$$\psi_k(x) = \langle x\rangle^N \widehat{a}^W\varphi_k(x).$$

Since

$$\|\widehat{a}^W\varphi_k - \widehat{a}^W\varphi_l\|_{L^2} \le \|\langle x\rangle^{-N}\|_{L_2}\cdot\|\psi_k-\psi_l\|_{L^\infty},$$

it is enough to find φ_k so that the sequence ψ_k converges in $L^\infty(\mathbb{R}^n)$.

To find such a sequence, we use the standard diagonalization trick. We denote $\mathbb{Q}^n = \{x_1, x_2, \cdots\}$. We first observe that, according to (1), the family

$$\{\psi(x) := \langle x\rangle^N\widehat{a}^W\varphi(x) \mid \varphi\in\mathcal{B}\}$$

is uniformly bounded. So we can choose a sequence $\{\varphi_k^1\}\subset\mathcal{B}$ so that $\langle x\rangle^N\widehat{a}^W\varphi_k^1$ converges at x_1, and in general we can inductively choose a subsequence $\{\varphi_k^m\}\subset\{\varphi_k^{m-1}\}$ so that $\langle x\rangle^N\widehat{a}^W\varphi_k^m$ converges at the point x_k. If we set $\varphi_k := \varphi_k^k$, then the sequence ψ_k converges at all $x_l\in\mathbb{Q}^n$.

Finally we prove the sequence ψ_k is a Cauchy sequence in $L^\infty(\mathbb{R}^n)$, and thus converges in $L^\infty(\mathbb{R}^n)$. By (1), there exists a constant M so that for all k,

$$|\nabla\psi_k(x)|\le M/3, \quad \langle x\rangle|\psi_k(x)|\le M/2.$$

Fix $\varepsilon>0$ and choose R large enough so that $M/R<\varepsilon$. Choose points $\{y_1,\cdots,y_P\}\subset\mathbb{Q}^n$ such that

$$B(0,R)\subset\cup_{p=1}^P B(y_p,\frac{\varepsilon}{M}).$$

Since each sequence $\{\psi_k(y_p)\}$ is a Cauchy sequence, there exists K so that

$$|\psi_k(y_p) - \psi_l(y_p)| < \frac{\varepsilon}{3}$$

for all $k, l > K$ and all $1 \le p \le P$. Moreover, according to the choice of R,

$$\sup_{|x|\ge R} |\psi_k(x) - \psi_l(x)| \le \frac{2}{R} \sup_{|x|\ge R} \langle x\rangle |\psi_l(x)| < \varepsilon.$$

To estimate $\sup_{|x|<R} |\psi_k(x) - \psi_l(x)|$, for any $x \in B(0,R)$ we choose y_p so that $|x - y_p| < \varepsilon/M$. It follows

$$\begin{aligned}|\psi_k(x) - \psi_l(x)| &\le |\psi_k(x) - \psi_k(y_p)| + |\psi_k(y_p) - \psi_l(y_p)| + |\psi_l(y_p) - \psi_l(x)| \\ &\le \frac{\varepsilon}{3} + (|\nabla\psi_k| + |\nabla\psi_l|)|x - y_p| < \varepsilon.\end{aligned}$$

Thus for any $k, l > K$, we have

$$\|\psi_k - \psi_l\|_{L^\infty} \le \max(\sup_{|x|<R} |\psi_k(x) - \psi_l(x)|, \sup_{|x|\ge R} |\psi_k(x) - \psi_l(x)|) < \varepsilon.$$

This completes the prove. $\square$

¶ Compactness for more general symbols.

Now we want to study the compactness of $\widehat{a}^W$ for more general symbols $a \in S(m)$, where m is some order function. The idea is similar to the proof of Calderon-Vaillancourt's theorem, namely we first use the periodic partition of unity

$$1 = \sum_{\alpha\in\mathbb{Z}^{2n}} \chi_\alpha$$

that we constructed last time to decompose the symbol a into countably many almost disjoint compact supported symbols which are uniformly controllable:

$$a(x,\xi) = \sum_{\alpha\in\mathbb{Z}^{2n}} a_\alpha(x,\xi),$$

where as last time, $a_\alpha := \chi_\alpha a$. Since each a_α is compactly supported, the operator $\widehat{a}_\alpha^W$ is compact. As a result, any finite sum $\sum_{|\alpha|\le M} \widehat{a_\alpha}^W$ is compact. Since the set of compact operators is closed in the set of bounded linear operators, to prove that $\widehat{a}^W$ is compact, it is enough to prove the convergence of $\sum_\alpha \widehat{a_\alpha}^W$. So, again we need the Cotlar-Stein lemma to add these compact operators:

Lemma 1.3 (Cotlar-Stein Lemma)**.** *Let H_1, H_2 be Hilbert spaces. For $j \in \mathbb{N}$ let $A_j \in \mathcal{L}(H_1, H_2)$ be bounded linear operators satisfying*

$$\sup_j \sum_{k=1}^\infty \|A_j^* A_k\|^{1/2} < C \qquad \text{and} \qquad \sup_j \sum_{k=1}^\infty \|A_j A_k^*\|^{1/2} \le C.$$

The the series $\sum_{j=1}^\infty A_j$ converges in the strong operator topology to $A \in \mathcal{L}(H_1, H_2)$ with $\|A\| \le C$.

However, this time the convergence in strong operator norm is not enough for us, since we need $\sum_\alpha \widehat{a_\alpha}^W$ to converge in operator norm topology. Fortunately, the Cotlar-Stein Lemma can also be applied to prove convergence in operator norm topology:

$$\begin{aligned} &\text{To prove } \sum_j A_j \text{ converges to } A \text{ in the operator norm topology} \\ \Longleftrightarrow\ &\text{To prove } \sum_{j\le M} A_j - A = \sum_{j>M} A_j \text{ converges to } B_M \text{ with } \|B_M\| \le \varepsilon \\ \Longleftrightarrow\ &\text{Apply Cotlar-Stein to } \sum_{j>M} A_j \text{ for } M\to\infty \text{ and try to get bounds} \\ &\text{that tends to zero.} \end{aligned}$$

So we still need the "almost orthogonality" of $\widehat{a_\alpha}^W$ to control on the "interactions" $\widehat{b_{\alpha\beta}}^W = (\widehat{a_\alpha}^W)^* \circ \widehat{a_\beta}^W$. But this time, we need a better control. Fortunately, a better estimate exists, as long as we use a nicer symbol class $S(m)$ instead of $S(1)$:

Again suppose $\hbar = 1$.

Lemma 1.4 ("Almost orthogonality" for $S(m)$)**.** *Suppose $a \in S(m)$. Then for each N and each multi-index γ, there is a constant*

$$C = C(\gamma, N, n) \sum_{|\alpha|\le 2N+4n+1+|\gamma|} \sup_{\mathbb{R}^n} |\partial^\alpha a|,$$

such that for all $z = (x,\xi) \in \mathbb{R}^{2n}$,

$$|\partial^\gamma b_{\alpha\beta}(z)| \le C m(\alpha) m(\beta) \langle \alpha-\beta\rangle^{-N} \langle z - \frac{\alpha+\beta}{2}\rangle^{-N}. \tag{2}$$

As a consequence, we get from the explicit formula of the $\mathcal{L}(L^2)$-bound of $\widehat{a}^W$ with Schwartz symbols (c.f. Proposition 1.4 in Lecture 11) that

$$\|\widehat{b_{\alpha\beta}}^W\|_{\mathcal{L}(L^2(\mathbb{R}^n))} \le C m(\alpha) m(\beta) \langle \alpha-\beta\rangle^{-N}. \tag{3}$$

Let's put the above ideas together. We want

$$\|\sum_{|\alpha|>M} \widehat{a_\alpha}^W\|_{\mathcal{L}(L^2(\mathbb{R}^n))} \to 0 \quad \text{as } M\to\infty. \tag{4}$$

So instead of applying the Cotlar-Stein Lemma to the whole sum $\sum_\alpha \widehat{a}^W$, we will apply it to the sum $\sum_{|\alpha|>M} \widehat{a_\alpha}^W$, which gives us

$$\left\|\sum_{|\alpha|>M} \widehat{a_\alpha}^W\right\|_{\mathcal{L}(L^2)} \le \max\left(\sup_{|\alpha|>M} \sum_{|\beta|>M} \|\widehat{a_\alpha}^W (\widehat{a_\beta}^W)^*\|^{1/2}, \sup_{|\alpha|>M} \sum_{|\beta|>M} \|(\widehat{a_\alpha}^W)^* \widehat{a_\beta}^W\|^{1/2} \right).$$

Since $(\widehat{a_\alpha}^W)^* \circ \widehat{a_\beta}^W = \widehat{b_{\alpha\beta}}^W$, we have

$$\sup_{|\alpha|>M} \sum_{|\beta|>M} \|(\widehat{a_\alpha}^W)^* \circ \widehat{a_\beta}^W\|_{\mathcal{L}(L^2)}^{1/2} \le C \sup_{|\alpha|>M} \sum_{|\beta|>M} \sqrt{m(\alpha)m(\beta)} \langle \alpha-\beta\rangle^{-N/2} \le C \sup_{|\alpha|>M} m(\alpha).$$

Similarly

$$\sup_{|\alpha|>M} \sum_{|\beta|>M} \|\widehat{a_\alpha}^W (\widehat{a_\beta}^W)^*\|^{1/2} \le C \sup_{|\alpha|>M} m(\alpha).$$

So for (4) to hold, it is enough to assume

$$m(z) \to 0 \qquad \text{as } z \to \infty!$$

Of course we assumed $\hbar = 1$ in these arguments. However, it is easy to extend the arguments to general $\hbar$ by using the re-scaling trick as we did last time. We can also use the change of quantization formula to extend the theorem to other semiclassical t-quantizations. Thus we get

Theorem 1.5. *Suppose $m = m(x, \xi)$ is an order function on $\mathbb{R}^{2n}$ such that*

$$\lim_{z\to\infty} m(z) = 0. \tag{5}$$

Then for any $a \in S(m)$, the operator $\widehat{a}^W : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ is a compact operator.

2. Proof of Lemma 1.4.

Note: In this section we take $\hbar = 1$.

Proof. Recall (See Lecture 9, page 5)

$$\begin{aligned}\bar{a}_\alpha \star a_\beta(x,\xi) &= e^{\frac{i\hbar}{2}(D_\xi\cdot D_y - D_x\cdot D_\eta)}[\bar{a}_\alpha(x,\xi)a_\beta(y,\eta)]\Big|_{y=x,\eta=\xi} \\ &= \frac{1}{(\pi\hbar)^{2n}}\int_{\mathbb{R}^{4n}} e^{-\frac{2i}{\hbar}(\tilde{\xi}\cdot\tilde{y}-\tilde{x}\cdot\tilde{\eta})}\bar{a}_\alpha(x+\tilde{x},\xi+\tilde{\xi})a_\beta(x+\tilde{y},\xi+\tilde{\eta})d\tilde{x}d\tilde{\xi}d\tilde{y}d\tilde{\eta}.\end{aligned}$$

With the notation $z = (x, \xi)$, $w_1 = (-\tilde{x}, -\tilde{\xi})$, $w_2 = (-\tilde{y}, -\tilde{\eta})$ and $\varphi(w_1, w_2) := \tilde{\xi}\cdot\tilde{y} - \tilde{x}\cdot\tilde{\eta}$ one gets

$$b_{\alpha,\beta}(z) = \bar{a}_\alpha \star a_\beta(z) = \frac{1}{\pi^{2n}}\int_{\mathbb{R}^{4n}} e^{-2i\varphi(w_1,w_2)}\bar{a}_\alpha(z-w_1)a_\beta(z-w_2)dw_1dw_2.$$

Our aim is to find an estimation of $|\partial^\gamma b_{\alpha\beta}|$, which should be "small" for $|\alpha-\beta|$ large. Since each a_α is supported near the lattice point α, it is natural to separate the integral above into two parts, ... So we choose a cut-off function

$$\zeta = \zeta(w_1, w_2) \in C_0^\infty(\mathbb{R}^{4n})$$

such that $0 \le \zeta \le 1$ on $\mathbb{R}^{4n}$, $\zeta = 1$ on $B^{4n}(0,1)$ and $\zeta = 0$ on $\mathbb{R}^{4n} \setminus B^{4n}(0,2)$. By plugging $1 = \zeta + (1-\zeta)$ in the integrand, we can split the above integral into two parts:

$$b_{\alpha\beta}(z) = I_1 + I_2.$$

It remains to estimate

$$|\partial^\gamma I_1| \qquad \text{and} \qquad |\partial^\gamma I_2|.$$

In estimating these terms, we also need to use the following elementary estimation (which, of course, is a mathematical way to say that the partition of unity we chose is "uniformly controllable"): for any multi-index γ, there exists a constant C_γ such that

$$|\partial^\gamma a_\alpha(w)| = |\partial^\gamma[\chi_0(w-\alpha)a(w)]| \le C_\gamma \sup_{\rho\le\gamma} |\partial^\rho \chi_0(w-\alpha)| m(w) \le C_\gamma m(\alpha),$$

where in the last step we used the fact that $\chi_0(w-a) = 0$ for $|w-a| > \sqrt{2n}$, and for $|w-a| < \sqrt{2n}$ we have

$$m(w) \le Cm(\alpha)$$

by the definition of an order function m.

Estimate I_1.

For the first part

$$I_1 = \frac{1}{\pi^{2n}} \int_{\mathbb{R}^{4n}} e^{-2i\varphi(w_1,w_2)} \zeta(w_1,w_2) \bar{a}_\alpha(z-w_1) a_\beta(z-w_2) dw_1 dw_2,$$

we have

$$\begin{aligned} |I_1| &\le C \int_{B^{4n}(0,2)} |\bar{a}_\alpha(z-w_1)||a_\beta(z-w_2)| dw_1 dw_2 \\ &= C \int_{B^{4n}(0,2)} \chi_0(z-w_1-\alpha)\chi_0(z-w_2-\beta)|a(z-w_1)||a(z-w_2)| dw_1 dw_2. \end{aligned}$$

Recall that χ_0 is supported in $B(0,\sqrt{2n})$. So the integrand vanishes unless

$$|z-w_1-\alpha| \le \sqrt{2n} \quad \text{and} \quad |z-w_2-\beta| \le \sqrt{2n}$$

for some $|w| \le 2$. It follows

$$|\alpha-\beta| \le 4 + 2\sqrt{2n} \quad \text{and} \quad \left| z - \frac{\alpha+\beta}{2} \right| \le 2 + \sqrt{2n}.$$

Since $a \in S(m)$ and $|z-w_1-\alpha| \le \sqrt{2n}$, we have

$$|\partial^\alpha a(z-w_1)| \le C_\alpha m(z-w_1) \le Cm(\alpha).$$

So for each N one could find $C = C_{N,n}$ so that

$$|I_1| \le C_{N,n} \langle \alpha-\beta \rangle^{-N} m(\alpha) m(\beta) \langle z - \frac{\alpha+\beta}{2} \rangle^{-N}.$$

Similarly for each multi-index γ one could find $C = C_{N,\gamma,n}$ so that

$$|\partial^\gamma I_1| \le C_{N,\gamma,n,a} m(\alpha) m(\beta) \langle \alpha-\beta \rangle^{-N} \langle z - \frac{\alpha+\beta}{2} \rangle^{-N},$$

where the dependence of $C_{N,\gamma,n,a}$ on a is clear: C is bounded by a constant multiple of $\sup_{\rho\le\gamma} |\partial^\rho a|$

Estimate I_2.

To estimate the second part

$$I_2 = \frac{1}{\pi^{2n}} \int_{\mathbb{R}^{4n}} e^{-2i\varphi(w_1,w_2)} (1-\zeta(w_1,w_2)) \bar{a}_\alpha(z-w_1) a_\beta(z-w_2) dw_1 dw_2,$$

we use integration by parts trick. As in Lecture 5 we introduce the operator

$$L = 4 + \langle \nabla\psi, D_w\rangle,$$

where $w = (w_1, w_2)$ and $\psi(w) := -2\varphi(w_1, w_2) = -2(\tilde{\xi}\cdot\tilde{y} - \tilde{x}\cdot\tilde{\eta})$. Since $|\nabla\psi| = 2|\omega|$, we have

$$\frac{1}{4\langle w\rangle^2} L(e^{i\psi}) = e^{i\psi}.$$

Since the integrant of I_2 vanishes for $|w| \le 1$, we can integrate by parts via L many times to get

$$|I_2| \le C_M \int_{|w|\ge 1} \langle\omega\rangle^{-M} \bar{c}_{\alpha,M}(z-\omega_1)c_{\beta,M}(z-\omega_2)d\omega_1 d\omega_2$$

for some smooth $c_{\alpha,M}$ supported in $B(\alpha, \sqrt{2n})$ and $c_{\beta,M}$ supported in $B(\beta, \sqrt{2n})$. So again the integrand vanishes unless

$$|z - w_1 - \alpha| \le \sqrt{2n} \quad \text{and} \quad |z - w_2 - \beta| \le \sqrt{2n}$$

for some $|w| \ge 1$, which implies

$$|\alpha - \beta| \le 2\sqrt{2n} + |w_1| + |w_2| \le (2\sqrt{2n} + 2)\langle w\rangle$$

and

$$\left| z - \frac{\alpha+\beta}{2} \right| \le \sqrt{2n} + \frac{|w_1| + |w_2|}{2} \le (\sqrt{2n} + 1)\langle w\rangle.$$

So if we take $M \ge 2N + 4n + 1$,

$$\begin{aligned} |I_2| &\le C_M \int_{|w|\ge 1} \langle w\rangle^{2N-M} \langle\alpha-\beta\rangle^{-N} \langle z - \frac{\alpha+\beta}{2}\rangle^{-N} dw \\ &\le C_M \langle\alpha-\beta\rangle^{-N} \langle z - \frac{\alpha+\beta}{2}\rangle^{-N}. \end{aligned}$$

Similarly if we estimate the derivative $|\partial^\gamma I_2|$, we will need to take no more than $M + |\gamma|$ derivatives of a and the conclusion follows. □

3. Appendix: Proof of the Cotlar-Stein Lemma

Let H_1, H_2 be Hilbert spaces and $A \in \mathcal{L}(H_1, H_2)$ a bounded linear operator. Recall that the norm of A is by definition

$$\|A\| = \sup_{|x|=1} |Ax|. \tag{6}$$

If we let $A^* \in \mathcal{L}(H_2, H_1)$ be the adjoint of A, then one has (exercise!)

$$\|A\| = \|A^*\| \qquad \text{and} \qquad \|A\|^2 = \|A^*A\|. \tag{7}$$

In the case $H_1 = H_2 = H$ so that $A^m \in \mathcal{L}(H, H)$ for all $m \in \mathbb{N}$, one has

$$\|A^m\| \le \|A\|^m.$$

(One could have strictly inequality above, for example, when A is nilpotent.) However, if we assume A is self-adjoint, i.e. $A = A^*$, then for all $m \in \mathbb{N}$,

$$\|A^m\| = \|A\|^m.$$

Proof of Cotlar-Stein Lemma. Let's first assume that $A_j = 0$ for all $j > N$ so that A is obviously a well-defined bounded linear operator. Since A^*A is self-adjoint,

$$\|A\|^{2m} = \|A^*A\|^m = \|(A^*A)^m\| = \left\| \sum_J A_{j_1}^* A_{j_2} \cdots A_{j_{2m-1}}^* A_{j_{2m}} \right\|,$$

where the summation is over $J = (j_1, \cdots, j_{2m}) \in [1, \cdots, N]^{2m}$. According to the Cauchy-Schwarz inequality,

$$\|A_{j_1}^* A_{j_2} \cdots A_{j_{2m-1}}^* A_{j_{2m}}\| \le \|A_{j_1}^* A_{j_2}\| \cdots \|A_{j_{2m-1}}^* A_{j_{2m}}\|$$

and

$$\|A_{j_1}^* A_{j_2} \cdots A_{j_{2m-1}}^* A_{j_{2m}}\| \le \|A_{j_1}^*\| \|A_{j_2} A_{j_3}^*\| \cdots \|A_{j_{2m}}\|.$$

Together with the facts $\|A_1\| \le C$ and $\|A_{j_{2m}}\| \le C$, one gets

$$\|A_{j_1}^* A_{j_2} \cdots A_{j_{2m-1}}^* A_{j_{2m}}\| \le C \|A_{j_1}^* A_{j_2}\|^{1/2} \|A_{j_2} A_{j_3}^*\|^{1/2} \cdots \|A_{j_{2m-1}}^* A_{j_{2m}}\|^{1/2}$$

It follows

$$\begin{aligned} \|A\|^{2m} &\le C \sum_J \|A_{j_1}^* A_{j_2}\|^{1/2} \|A_{j_2} A_{j_3}^*\|^{1/2} \cdots \|A_{j_{2m-1}}^* A_{j_{2m}}\|^{1/2} \\ &= C \sum_{j_1, j_2} \|A_{j_1}^* A_{j_2}\|^{1/2} \sum_{j_3} \|A_{j_2} A_{j_3}^*\|^{1/2} \cdots \sum_{j_{2m}} \|A_{j_{2m-1}}^* A_{j_{2m}}\|^{1/2} \\ &\le N C^{2m} \end{aligned}$$

and thus (since N is fixed and m is arbitrary)

$$\|A\| \le C.$$

To prove the general case, we first take $x \in H_1$ so that $x = A_k^* y$ for some k and some $y \in H_2$. Then

$$\left\| \sum_{j=1}^\infty A_j x \right\| = \left\| \sum_{j=1}^\infty A_j A_k^* y \right\| \le \sum_{j=1}^\infty \|A_j A_k^*\|^{1/2} \, \|A_j A_k^*\|^{1/2} \, \|y\| \le C^2 \|y\|.$$

It follows that the series $\sum_j A_j x$ converges for all

$$x \in U := \operatorname{span}\{A_k^* y \mid y \in H_2, k \ge 1\}.$$

Since we have already proved $\| \sum_{j=1}^N A_j \| \le C$ uniformly for all N, one could commute the summation with limit and conclude that the series $\sum_j A_j x$ converges for all $x \in \overline{U}$, and $\| \sum_j A_j x \| \le C \|x\|$.

Finally if $x \in \overline{U}^\perp$, then $x \in \ker(A_k)$ for all k, so we trivially have $\sum A_j x = 0$. $\square$