The notes in this section are compiled from:

- Notes from a graduate course I took at Berkeley from Don Sarason in 2006,
- Pedersen's *Analysis Now*, and

2 Hilbert space basics

For this section, H is a Hilbert space. Recall the polarization identity, which holds for any sesquilinear form:

$$\langle \eta, \xi \rangle = \frac{1}{4}\sum_{k=0}^{3} i^k \langle \eta + i^k\xi, \eta + i^k\xi\rangle \qquad \forall\, \eta, \xi \in H. \tag{2.0.1}$$

Exercise 2.0.2. Prove that a positive sesquilinear form is self adjoint.

The adjoint is defined via the Riesz-Representation Theorem, i.e., if $x \in B(H \to K)$, for all $\xi \in K$, $\eta \mapsto \langle x\eta, \xi\rangle_K$ is a bounded linear functional on H, so there is a unique $x^*\xi \in H$ such that

$$\langle x\eta, \xi\rangle_K = \langle \eta, x^*\xi\rangle_H \qquad \forall\, \eta \in H,\ \forall\, \xi \in K.$$

The assignment $\xi \mapsto x^*\xi$ is linear and bounded, so $x^* \in B(H)$.

Exercise 2.0.3. Explain the relationship between $x, x^*, \overline{x}, x^t$ where $\overline{x}\colon \overline{H} \to \overline{K}$ is the conjugate operator given by $\overline{x}(\overline{\eta}) := \overline{x\eta}$, and x^t is the transpose, given by the Banach adjoint $K^* \to H^*$ by $\langle \xi | \mapsto \langle \xi | \circ x$.

2.1 Operators in $B(H)$

We have various types of operators as in the C*-algebra notes. We call $x \in B(H)$:

- self-adjoint if $x = x^*$,
- positive if there is a $y \in B(H)$ such that $x = y^*y$,
- normal if $xx^* = x^*x$,
- a projection if $x = x^* = x^2$,
- an isometry if $x^*x = 1$,
- a unitary if $x^*x = 1 = xx^*$ (equivalently, an invertible isometry),
- a partial isometry if x^*x is a projection.

Here are some elementary properties about $B(H)$:

(B1) $\ker(x^*) = (xH)^\perp$.

Proof. Since $\langle x\eta, \xi\rangle = \langle \eta, x^*\xi\rangle$, we have $\xi \perp xH$ if and only if $x^*\xi \perp H$ if and only if $x^*\xi = 0$. □

(B2) $x = y$ if and only if $\langle x\xi, \xi\rangle = \langle y\xi, \xi\rangle$ for all $\xi \in H$.

Proof. Replacing x with $x - y$, we may assume $y = 0$. The forward direction is trivial. Suppose $\langle x\xi, \xi\rangle = 0$ for all $\xi \in H$. Polarization (2.0.1) applied to the form $\langle x\,\cdot\,, \cdot\rangle$ implies $\langle x\eta, \xi\rangle = 0$ for all $\eta, \xi \in H$. Thus $x\eta \perp H$ for all $\eta \in H$, so $x = 0$. □

(B3) x is normal if and only if $\|x\xi\| = \|x^*\xi\|$ for all $\xi \in H$.

Proof. By (B2), $x^*x = xx^*$ if and only if $\langle x^*x\xi, \xi\rangle = \langle xx^*\xi, \xi\rangle$ for all $\xi \in H$. But this holds if and only if $\|x\xi\|^2 = \|x^*\xi\|^2$ for all $\xi \in H$. □

(B4) $x \in B(H)$ is self-adjoint if and only if $\langle x\xi, \xi\rangle \in \mathbb{R}$ for all $\xi \in H$.

Proof. Homework. □

2.2 Normal operators

We now prove some elementary properties about normal operators. For the following properties, $x \in B(H)$ is normal.

(N1) $x\xi = \lambda\xi$ if and only if $x^*\xi = \overline{\lambda}\xi$.

Proof. Immediate from (B3) applied to $x - \lambda$. □

(N2) $x\eta = \lambda\eta$ and $x\xi = \mu\xi$ with $\lambda \neq \mu$ implies $\eta \perp \xi$.

(N3) Every $\lambda \in \mathrm{sp}(x)$ is an approximate eigenvalue of x, i.e., there is a sequence of unit vectors (ξ_n) such that $(x - \lambda)\xi_n \to 0$.

Proof. Suppose λ is not an approximate eigenvalue of x. Then there is a $\varepsilon > 0$ such that $\|(x - \lambda)\xi\| \geq \varepsilon\|\xi\|$ for all $\xi \in H$. Then $x - \lambda$ is injective with closed range, and by (B3), so is $x^* - \overline{\lambda}$. But $0 = \ker(x^* - \overline{\lambda}) = ((x - \lambda)H)^\perp$ by (B1). Thus $x - \lambda$ is surjective, and thus $x - \lambda$ is bijective and bounded, hence invertible. Thus $\lambda \notin \mathrm{sp}(x)$. □

(N4) $\|x\| = \sup\{|\langle x\xi, \xi\rangle| \,|\, \|\xi\| = 1\}$

Proof. Since $r(x) = \|x\|$, there is a $\lambda \in \operatorname{sp}(x)$ such that $|\lambda| = \|x\|$. Then since λ is an approximate eigenvalue by (N3), there is a sequence (ξ_n) of unit vectors such that $(x-\lambda)\xi_n \to 0$. Thus

$$\begin{aligned} |\langle x\xi_n, \xi_n\rangle - \lambda| &= |\langle x\xi_n, \xi_n\rangle - \lambda\langle \xi_n, \xi_n\rangle| \\ &= |\langle (x-\lambda)\xi_n, \xi_n\rangle| \\ &\underset{(\text{CS})}{\leq} \|x\xi_n - \lambda\xi_n\| \cdot \underbrace{\|\xi_n\|}_{=1} \xrightarrow{n\to\infty} 0. \end{aligned}$$

□

(N5) If $x = x^*$,

$$\begin{aligned} \sup\{\langle x\xi, \xi\rangle \,|\, \|\xi\| = 1\} &= \max\{\lambda \,|\, \lambda \in \operatorname{sp}(x)\} \qquad \text{and} \\ \inf\{\langle x\xi, \xi\rangle \,|\, \|\xi\| = 1\} &= \min\{\lambda \,|\, \lambda \in \operatorname{sp}(x)\} \end{aligned}$$

Proof. Set $M := \max\{\lambda \,|\, \lambda \in \operatorname{sp}(x)\}$. By the Spectral Mapping Theorem, $\operatorname{sp}(x + \|x\|) = \operatorname{sp}(x) + \|x\| \subset [0, \infty)$, and thus $x + \|x\|$ is (spectrally) positive. Then

$$\begin{aligned} M + \|x\| &\underset{(\text{SMT})}{=} \max\{\lambda \,|\, \lambda \in \operatorname{sp}(x + \|x\|)\} \underset{(\text{N4})}{=} \sup\{\langle (x + \|x\|)\xi, \xi\rangle \,|\, \|\xi\| = 1\} \\ &= \sup\{\langle x\xi, \xi\rangle \,|\, \|\xi\| = 1\} + \|x\|. \end{aligned}$$

The proof for the second is similar swapping min and inf for max and sup, and subtracting $\|x\|$. □

Remark 2.2.1. The set

$$R(x) := \{\langle x\xi, \xi\rangle \,|\, \|\xi\| = 1\}$$

is called the *numerical range* of $x \in B(H)$. It is always convex subset of $\mathbb{C}$; this is easy to show when x is self-adjoint. Indeed, since $\xi \mapsto \langle x\xi, \xi\rangle$ is continuous and the unit sphere is connected, $R(T)$ is then a connected subset of $\mathbb{R}$, i.e., an interval.

Proposition 2.2.2. *The following are equivalent for $x \in B(H)$.*

1. $\langle x\xi, \xi\rangle \geq 0$ *for all* $\xi \in H$.

2. x *is normal and* $\operatorname{sp}(x) \subset [0, \infty)$.

3. x *is positive.*

Proof.

(1) ⇒ (2): Assuming (1), we have

$$\langle x\xi,\xi\rangle = \overline{\langle x\xi,\xi\rangle} = \langle \xi, x\xi\rangle = \langle x^*\xi,\xi\rangle \qquad \forall \xi\in H,$$

so $x = x^*$ by (B2). By (N4),

$$\operatorname{sp}(x) \subset \overline{R(x)} \subset [0,\infty).$$

(2) ⇒ (3): Since x is normal and $\operatorname{sp}(x)\subset[0,\infty)$, we can use the continuous functional calculus to get a self-adjoint operator $\sqrt{x}\in B(H)$ such that $\sqrt{x}^2 = x$.

(3) ⇒ (1): If $x = y^*y$ for some $y\in B(H)$, then

$$\langle x\xi,\xi\rangle = \langle y^*y\xi,\xi\rangle = \langle y\xi,y\xi\rangle = \|y\xi\|^2 \qquad \forall \xi\in H. \qquad \square$$

Theorem 2.2.3 (Fuglede). *Suppose $x, y \in B(H)$ such that $xy = yx$. If x is normal, then $x^*y = yx^*$.*

Proof due to Rosenblum. Since $xy = yx$, $ye^{i\overline{\lambda}x} = e^{i\overline{\lambda}x}y$, so $x = e^{i\overline{\lambda}x}ye^{-i\overline{\lambda}x}$ for all $\lambda\in\mathbb{C}$. We define $f\colon \mathbb{C}\to B(H)$ by

$$f(\lambda) := e^{i\lambda x^*}ye^{-i\lambda x^*} = e^{i\lambda x^*}e^{i\overline{\lambda}x}ye^{-i\overline{\lambda}x}e^{-i\lambda x^*} = e^{i(\lambda x^*+\overline{\lambda}x)}ye^{-i(\lambda x^*+\overline{\lambda}x)}.$$

Since $\lambda x^* + \overline{\lambda}x$ is self-adjoint, $e^{i(\lambda x^*+\overline{\lambda}x)}$ is unitary. Hence $f\colon\mathbb{C}\to B(H)$ is a bounded $B(H)$-valued entire function, and thus constant by Liouville. Thus

$$0 = -i\cdot \frac{d}{d\lambda}\bigg|_{\lambda=0} f(\lambda) = x^*y - yx^*.$$

(Take the power series expansion to first order.) $\square$

Exercise 2.2.4. Where is normality of x used in the proof of Fuglede's Theorem 2.2.3?

Corollary 2.2.5. *If $x\in B(H)$ is normal and $xy = yx$, then $yf(x) = f(x)y$ for all $f\in C(\operatorname{sp}(x))$.*

Proof. By Fuglede's Theorem 2.2.3, the result holds for all polynomials in x and x^*. The result now follows by density of these polynomials in $C(\operatorname{sp}(x))$ by Stone-Weierstrass. $\square$

Remark 2.2.6. The results in this section also hold for operators in a unital C*-algebra, not just $B(H)$.

2.3 Projections and partial isometries

Example 2.3.1. Let $x \in B(H)$. The *support projection* of x is $\operatorname{supp}(x) := 1 - p_{\ker(x)} = p_{\ker(x)^\perp}$. The *range projection* of x is $\operatorname{Range}(x) := p_{\overline{xH}}$.

Observe that $x = \operatorname{Range}(x) \cdot x \cdot \operatorname{supp}(x)$. By (B1), $\operatorname{Range}(x) = \operatorname{supp}(x^*)$. If x is normal, then since $\ker(x) = \ker(x^*x) = \ker(xx^*) = \ker(x^*)$, $\operatorname{supp}(x) = \operatorname{Range}(x)$.

Lemma 2.3.2. *The map $p \mapsto pH$ is a bijective correspondence between projections and closed subspaces of H.*

Proof. It is clear that $pH \subseteq H$ is a closed subspace as p is continuous and $p = p^2$. Moreover, since $p = p^*$, $pH^\perp = \ker(p^*) = \ker(p) = (1-p)H$.
Conversely, every closed subspace $K \subseteq H$ has an orthogonal complement $K^\perp$, $H = K \oplus K^\perp$, and projection p_K onto K is an idempotent. We claim it is self-adjoint. Indeed, $\ker(p_K^*) = p_K H^\perp = K^\perp = \ker(p_K)$, which implies $p_K^*(1-p_K) = 0$, and thus $p_K^* p_K = p_K^*$. But $p_K^* p_K$ is self-adjoint, and thus $p_K = p_K^*$.
One checks these two constructions are mutually inverse. □

Lemma 2.3.3. *For $p, q \in P(M)$, the following are equivalent.*

1. *$p \leq q$ ($q - p \geq 0$),*
2. *$pH \subseteq qH$, and*
3. *$p = pq$.*

Proof.

(1) ⇒ (2): We show $(1-q)H \subseteq (1-p)H$, and the result follows by taking orthogonal complements. Suppose $\xi \in (1-q)H$ so $q\xi = 0$. Then since $0 \leq q - p$,

$$0 \leq \langle (q-p)\xi, \xi \rangle = \underbrace{\langle q\xi, \xi \rangle}_{=0} - \langle p\xi, \xi \rangle = -\langle p\xi, \xi \rangle = -\|p\xi\|^2.$$

Thus $p\xi = 0$, so $\xi \in (1-p)H$.

(2) ⇒ (3): If $pH \subseteq qH$, then projecting to qH and then to pH is the same as just projecting to pH.

(3) ⇒ (1): If $p = pq$, then $p = p^* = qp$. Thus $q - p = q - qpq = q(1-p)q \geq 0$. □

Exercise 2.3.4. We say projections p, q are mutually orthogonal, denoted $p \perp q$, if $pH \perp qH$. Show that $p \perp q$ if and only if $pq = 0$.

Exercise 2.3.5. For projections p, q, we define $p \wedge q$ to be the projection onto $pH \cap qH$ and $p \vee q$ to be the projection onto $\overline{pH + qH}$. Prove that $p \vee q = 1 - (1-p) \wedge (1-q)$.

Exercise 2.3.6. Prove the following statements about projections and invariant subspaces.

1. $K \subseteq H$ is x-invariant if and only if $p_K x p_K = x p_K$.
2. $K \subseteq H$ is x-invariant if and only if $K^\perp$ is x^*-invariant.
3. $K \subseteq H$ is x and x^*-invariant if and only if $x p_K = p_K x$.

Exercise 2.3.7. The following are equivalent for a $u \in B(H \to K)$.

1. u is a partial isometry.
2. $u = uu^*u$.
3. u^* is a partial isometry.
4. $u^* = u^*uu^*$.

Hint: Use the C^**-identity.*

Remark 2.3.8. By the exercise, a partial isometry $u \in B(H \to K)$ is a unitary from u^*uH onto uu^*K.

Exercise 2.3.9. Suppose $u, v \in B(H)$ are partial isometries with $uu^* \perp vv^*$ and $u^*u \perp v^*v$. Show that $u + v$ is again a partial isometry.

Proposition 2.3.10 (Polar decomposition)**.** *For each $x \in B(H \to K)$, there is a unique positive $|x| \in B(H)$ such that $|x|^2 = x^*x$ and $\|x\xi\| = \||x|\xi\|$ for all $\xi \in H$. Moreover, there is a unique partial isometry $u \in B(H \to K)$ such that $u|x| = x$ and* $\ker(u) = \ker(x) = \ker(|x|)$*. In particular, $u^*x = |x|$.*

Proof. If $y \geq 0$ such that $\|y\xi\| = \|x\xi\|$ for all $\xi \in H$, then

$$\langle x^*x\xi, \xi\rangle = \|x\xi\|^2 = \|y\xi\|^2 = \langle y^2\xi, \xi\rangle$$

so $x^*x = y^2$ by (B2), and thus $y = \sqrt{x^*x}$ by the uniqueness of the positive square root. Now define $u\colon |x|H \to K$ by $u|x|\xi := x\xi$, and note

$$\|u|x|\xi\| = \|x\xi\| = \||x|\xi\| \qquad \forall\, \xi \in H.$$

So u is an isometry on $|x|H$, and is thus well-defined. We can extend u to $\overline{|x|H}$ by continuity, and define $u = 0$ on $(|x|H)^\perp = \ker(|x|)$ by (B1), and $\ker(|x|) = \ker(x)$ by

construction. We will call this extension u again by a slight abuse of notation. Then u is a partial isometry and $u|x| = x$.
If $v \in B(H)$ is another partial isometry with $\ker(v) = \ker(x) = \ker(u)$ and $v|x| = x$, then $u|x|\xi = v|x|\xi$ for all $\xi \in H$, so $u = v$ on $\overline{|x|H}$. But $u = v = 0$ on $(|x|H)^{\perp}$, so $u = v$.
Finally, u^*u is the projection onto $\overline{|x|H}$, so $u^*x\xi = u^*u|x|\xi = |x|\xi$ for all $\xi \in H$. □

Exercise 2.3.11. Suppose $x = u|x|$ is the polar decomposition. Prove that $x = |x^*|u$ and the polar decomposition of x^* is given by $u^*|x^*|$.

Corollary 2.3.12. *If $x = u|x|$ is the polar decomposition, then $u^*u = \operatorname{supp}(x)$ and $uu^* = \operatorname{Range}(x)$.*

Proof. Since $\ker(u) = \ker(x)$, $\operatorname{supp}(x) = p_{\ker(x)^{\perp}} = p_{\ker(u)^{\perp}} = u^*u$. Since $x^* = u^*|x^*|$ is the polar decomposition of x^*, we have $\operatorname{Range}(x) = \operatorname{supp}(x^*) = uu^*$. □

Remark 2.3.13. If x is invertible, then so are x^* and x^*x, and by the CFC for x^*x, so is $|x|$. If $x = u|x|$ is the polar decomposition, then $u = x|x|^{-1} \in C^*(x)$ is a unitary. Hence if A is a unital C*-algebra and $a \in A$ is invertible, then a has a unique polar decomposition in A.

2.4 Compact operators

Recall $x \in B(H \to K)$ is called compact if it maps bounded subsets of H to precompact subsets (subset with compact closure) of K. We write $K(H \to K)$ for the subset of compact operators in $B(H \to K)$, and we write $K(H)$ for the compact operators in $B(H)$. Recall that $K(H)$ is a closed 2-sided ideal in $B(H)$.

Fact 2.4.1 (Spectra of compact operators)**.** Suppose $x \in K(H)$. The non-zero points of $\operatorname{sp}(x)$ are isolated eigenvalues, and all correspondonding eigenspaces are finite dimensional. There are only countably many of them, and zero is the only possible accumulation point.

Exercise 2.4.2. An operator $x \in B(H)$ is called finite rank if xH is finite dimensional.

1. Show that every finite rank operator is compact.
2. Show that the finite rank operators form a $*$-closed 2-sided ideal in $B(H)$.

Fact 2.4.3. Every $*$-closed 2-sided ideal $J \subseteq B(H)$ is spanned by its positive operators. First, note that every self-adjoint $x \in J$ can be written as $x = x_+ - x_-$ with $x_{\pm} \geq 0$ and $x_+x_- = 0$ by setting $x_+ := \chi_{[0,\infty)}(x)x$ and $x_- := \chi_{(-\infty,0]}(x)x$. Clearly $x_{\pm} \in J$, so every self-adjoint in J is in the span of the positives of J. Second, every $x = \Re(x) + i\Im(x)$ with $\Re(x) = (x + x^*)/2$ and $\Im(x) = (x - x^*)/(2i)$. Since J is $*$-closed, $\Re(x)$ and $\Im(x)$ are in J. Thus $\Re(x)_{\pm}, \Im(x)_{\pm} \in J$, and x is a linear combination of these 4 positives.

Lemma 2.4.4. *There is a net (p_i) of finite rank projections such that $p_i\xi \to \xi$ for all $\xi \in H$. In other words, $p_i \to 1$ in the strong operator topology (the topology of pointwise convergence).*

Proof. Let $(e_i)_{i\in I}$ be an ONB of H. Let $\mathcal{F}$ be the subset of finite subsets of I, ordered by inclusion. For $F \in \mathcal{F}$, define p_F to be the projection onto the finite dimensional (and thus closed) subspace span $\{e_i | i \in F\}$. By Parseval's identity, $\|p_F\xi - \xi\| \to 0$ for all $\xi \in H$. □

Theorem 2.4.5. *The following are equivalent for $x \in B(H)$. Below, B denotes the norm-closed unit ball in H.*

(K1) x is compact.

(K2) x is in the norm closure of the finite rank operators in $B(H)$.

(K3) $x|_B$ is weak-norm continuous $B \to H$

(K4) xB is compact in H.

Proof.

(1) ⇒ (2): Let $x \in K(H)$ and let (p_i) be a net as in Lemma 2.4.4. We claim that $p_i x \to x$ in norm. Otherwise, there is a $\varepsilon > 0$ such that (passing to a subnet if necessary) for all i, there is a $\xi_i \in H$ with $\|\xi_i\| = 1$ and $\varepsilon \leq \|(1 - p_i)x\xi_i\|$ and $x\xi_i \to \eta$ in H (by compactness of x). Then

$$\varepsilon \leq \|(1-p_i)x\xi_i\| \leq \|(1-p_i)(x\xi_i - \eta)\| + \|(1-p_i)\eta\| \leq \|x\xi_i - \eta\| + \|(1-p_i)\eta\| \to 0,$$

a contradiction.

(2) ⇒ (3): Suppose x is a norm limit of finite rank operators and (ξ_i) is a net of vectors in B converging weakly to $\xi \in B$. Let $\varepsilon > 0$. Choose a finite rank $y \in B(H)$ such that $\|x - y\| < \varepsilon$. We claim that $y\xi_i \to y\xi$. Indeed, choosing an ONB $\{e_1, \dots, e_n\}$ for the finite dimensional Hilbert space yH,

$$\|y(\xi_i - \xi)\|^2 = \sum_{k=1}^{n} |\langle y(\xi_i - \xi), e_k\rangle|^2 = \sum_{k=1}^{n} |\langle \xi_i - \xi, y^* e_k\rangle|^2 \longrightarrow 0.$$

Now pick j so that $i > j$ implies $\|y\xi_i - y\xi\| < \varepsilon$. For all $i > j$,

$$\|x\xi_i - x\xi\| \leq \|x\xi_i - y\xi_i\| + \|y\xi_i - y\xi\| + \|x\xi - y\xi\| < 3\varepsilon.$$

The result follows.

(3) ⇒ (4): Since B is weakly compact by Banach-Alaoglu, xB is the continuous image of a compact set which is thus compact.

(4) ⇒ (1): If $S \subset H$ is bounded, then $S \subset B_r = B_r(0_H)$ for some $r > 0$. Then $xB_r = rxB$ is compact, so the closure of xS is compact. □

Exercise 2.4.6. Prove that if $x \in B(H)$ is finite rank, then so is x^*. Deduce that $K(H)$ is $*$-closed.

Notation 2.4.7. We write $\langle \eta | \xi \rangle := \langle \xi, \eta \rangle$, which is linear on the right, and conjugate linear on the left. For $\eta \in H$, we write $\langle \eta | \in H^*$ for $\xi \mapsto \langle \eta | \xi \rangle$, and we can also denote $\xi \in H$ by $|\xi\rangle$. This allows us to define the rank one operator $|\eta\rangle\langle\xi| \in B(H)$ by $\zeta \mapsto |\eta\rangle\langle\xi|\zeta\rangle = \langle \zeta, \xi \rangle \eta$.

Exercise 2.4.8. Prove the following statements about rank one operators.

1. $|\eta\rangle\langle\xi|^* = |\xi\rangle\langle\eta|$
2. $|\eta_1\rangle\langle\eta_2| \cdot |\xi_1\rangle\langle\xi_2| = \langle\eta_2|\xi_1\rangle \cdot |\eta_1\rangle\langle\xi_2|$
3. If $\|\xi\| = 1$, then $|\xi\rangle\langle\xi|$ is the rank one projection onto $\mathbb{C}\xi$.

Definition 2.4.9. An operator $x \in B(H)$ is orthogonally diagonalizable if there is an ONB (e_i) of eigenvectors for x.

Exercise 2.4.10. Show that if $x \in B(H)$ is orthogonally diagonalizable, then the eigenvalues (λ_i) for (e_i) are in $\ell^\infty(I)$, where I is given counting measure.

Lemma 2.4.11. *An orthogonally diagonalizable operator $x \in B(H)$ is compact if and only if the eigenvalues (λ_i) for (e_i) is in $c_0(I)$, where I has the discrete topology, and $x = \sum_i \lambda_i |e_i\rangle\langle e_i|$, where the sum converges in norm.*

Proof. By Fact 2.4.1, since $\operatorname{sp}(x) \subseteq \{\lambda_i | i \in I\} \cup \{0\}$, we must have $(\lambda_i) \in c_0(I)$. Conversely, if $(\lambda_i) \in c_0(I)$, then $\sum \lambda_i |e_i\rangle\langle e_i|$ converges in operator norm to x. Indeed, if we define $x_F := \sum_{i \in F} \lambda_i |e_i\rangle\langle e_i|$ for each finite $F \subset I$, then picking $F \subset I$ so that $|\lambda_i| < \varepsilon$ for all $i \in F^c$, we have

$$\|(x - x_F)\xi\|^2 = \left\| \sum_{i \notin F} \lambda_i |e_i\rangle\langle e_i|\xi\rangle \right\|^2 = \sum_{i \notin F} |\lambda_i|^2 |\langle \xi, e_i \rangle|^2 < \varepsilon^2 \|\xi\|^2,$$

so $x_F \to x$ in norm. □

Theorem 2.4.12 (Spectral theorem for compact normal operators)**.** *Compact normal operators are diagonalizable.*

Proof. Suppose $x \in K(H)$ is normal. It suffices to prove H is the orthogonal direct sum of eigenspaces of x. We may assume $\dim(H) = \infty$. Using Fact 2.4.1, let (λ_n) be the non-zero eigenvalues of x, which is either a finite list or $\lambda_n \searrow 0$. Let E_n be the corresponding eigenspaces. Then E_n is an eigenspace for x^* with eigenvalue $\overline{\lambda}$ by (N1), and $E_n \perp E_k$ for all $1 \leq k < n$. Since each E_n is x and x^*-invariant, so is $\bigoplus_{n\geq 1} E_n$. Setting $E_0 := (\bigoplus_{n\geq 1} E_n)^\perp$, we have E_0 is x and x^*-invariant by Exercise 2.3.6. Then $x|_{E_0}$ is compact and has no non-zero eigenvalues, and so $x|_{E_0} = 0$. We conclude that $H = \bigoplus_{n\geq 0} E_n$ is the desired direct sum decomposition into eigenspaces. □

Remark 2.4.13. Using the Borel functional calculus and Theorem 2.4.12, one can show that a positive operator $x \in B(H)$ is compact if and only if for all $\varepsilon > 0$, the spectral projection $\chi_{(\varepsilon,\infty)}(x)$ is finite rank.

Corollary 2.4.14. *If $x \in B(H \to K)$ such that x^*x is compact, then x is compact.*

Proof. Writing $x^*x = \sum \lambda_n |e_n\rangle\langle e_n|$ with $\lambda_n \searrow 0$ by Theorem 2.4.12, we have $|x| = \sum \sqrt{\lambda_n}|e_n\rangle\langle e_n|$ with $\sqrt{\lambda_n} \searrow 0$. Thus $|x|$ is compact by Lemma 2.4.11, and so is $x = u|x|$ using polar decomposition 2.3.10. □

Definition 2.4.15. Suppose $x \in K(H)$, so $|x| = (x^*x)^{1/2}$ is compact. Enumerate the eigenvalues of $|x|$ by

$$\lambda_0 \geq \lambda_1 \geq \lambda_2 \geq \cdots$$

with multiplicity as necessary. Note that $\lambda_0 = \|x\|$.

We define $s_n(x) := \lambda_n$, called the n-th *singular value* of x.

Now pick orthonormal vectors (f_n) such that $|x|f_n = \lambda_n f_n$ and $|x| = \sum \lambda_n |f_n\rangle\langle f_n|$, which converges in operator norm. Set $e_n := uf_n$ where $x = u|x|$ is the polar decomposition 2.3.10. Then (e_n) is an orthonormal set, and $x = u|x| = u\sum \lambda_n |f_n\rangle\langle f_n| = \sum \lambda_n |e_n\rangle\langle f_n|$, where the sum converges in operator norm. This is called a *Schmidt representation* of x.

Warning 2.4.16. We warn the reader that a Schmidt decomposition of $x \in K(H)$ is not unique, but the singular values are well-defined. The usefulness of a Schmidt decomposition is that x is realized as an explicit norm-limit of finite rank operators.

For a unique representation, we can define $p_n = p_{E_n}$ to be the (finite rank) orthogonal projection with range E_n, the eigenspace of $|x|$ corresponding to $s_n(x)$. Then $|x| = \sum s_n(x)p_n$ and $x = \sum s_n(x)up_n$.

Here are some elementary properties about singular values.

(SV1) $s_n(x) = s_n(x^*)$ for all n.

Proof. Let $x = \sum s_n(x)|e_n\rangle\langle f_n|$ be a Schmidt decomposition for x. Using

Exercise 2.3.11, one can see that

$$x^* = \sum s_n(x)|f_n\rangle\langle e_n| = u^* \sum s_n(x)|e_n\rangle\langle e_n|$$

is a Schmidt decomposition for x^*, and thus $s_n(x^*) = s_n(x)$. Alternatively, we see that $xx^* = \sum s_n(x)^2|e_n\rangle\langle e_n|$ converges in norm, so $|x^*| = \sum s_n(x)|e_n\rangle\langle e_n|$, which also implies $s_n(x^*) = s_n(x)$. □

(SV2) (Minimax) Suppose $x \in K(H)$ is positive and non-zero. Then for all $n \geq 0$ such that $n \leq \dim(H)$,

$$s_n(x) = \min_{\substack{E \subseteq H \\ \operatorname{codim}(E)=n}} \max_{\substack{\xi \in E \\ \|\xi\|=1}} \langle x\xi, \xi\rangle. \tag{2.4.17}$$

Proof. First, we prove that $\max\{\langle x\xi, \xi\rangle | \xi \in E \text{ and } \|\xi\| = 1\}$ exists. By (K4), x is weak-norm continuous on B_E. Second, $\langle\,\cdot\,,\,\cdot\,\rangle : H \times H \to \mathbb{C}$ is *jointly* continuous on norm bounded sets in the product topology where the first factor has the norm topology and the second factor has the weak topology. Indeed, if $\eta_i \to \eta$ in norm and $\xi_i \to \xi$ weakly, we can find j in our index set so that $i > j$ implies $\|\eta_i - \eta\| < \varepsilon/M$ where M is a bound for the norm of all ξ_i and ξ. Then

$$|\langle \eta_i, \xi_i\rangle - \langle \eta, \xi\rangle| \leq \underbrace{|\langle \eta_i - \eta, \xi_i\rangle|}_{\leq \|\eta_i - \eta\| \cdot \|\xi_i\| < \varepsilon} + \underbrace{|\langle \eta, \xi_i - \xi\rangle|}_{\to 0}.$$

Hence the map $\xi \mapsto (x\xi, \xi) \mapsto \langle x\xi, \xi\rangle$ is continuous on B_E equipped with the weak topology. Since B_E is weakly compact by Banach-Alaoglu, the max exists. Now denote the right hand side of (2.4.17) by m_n. We know the case $n = 0$ holds. Assume $n > 0$ and let (f_k) be an orthonormal subset such that $x = \sum s_k(x)|f_k\rangle\langle f_k|$ with $\lambda_k \searrow 0$. For $E = \operatorname{span}\{f_0, \dots, f_{n-1}\}^\perp$, we have $f_n \in E$ and $\langle xf_n, f_n\rangle = s_n(x)$, so $m_n \leq \lambda_n$.
Conversely, if $\operatorname{codim}(E) = n$, then there is a $\xi \in E \cap \operatorname{span}\{f_0, \dots, f_n\}$ with $\|\xi\| = 1$. Then writing $\xi = \sum_{i=0}^n \alpha_i f_i$ with $\alpha_i = \langle \xi, f_i\rangle$ and $\sum |\alpha_i|^2 = 1$, we have

$$\langle x\xi, \xi\rangle = \sum_{i=0}^n s_i(x)|\alpha_i|^2 \geq s_n(x).$$

Hence $s_n(x) \leq m_n$. □

(SV3) If $x \in K(H)$, then

$$s_n(x) = \min_{\substack{E \subseteq H \\ \operatorname{codim}(E)=n}} \max_{\substack{\xi \in E \\ \|\xi\|=1}} \|x\xi\|. \tag{2.4.18}$$

Proof. Observe that $s_n(x) = \sqrt{s_n(x^*x)}$ and $\langle x^*x\xi, \xi\rangle = \|x\xi\|^2$. Apply Minimax (SV2) for x^*x and take square roots. ☐

(SV4) If $x \in K(H)$ and $y \in B(H)$, then both $s_n(xy), s_n(yx) \le \|y\| s_n(x)$.

Proof. Using Minimax (2.4.18), we have[a]

$$s_n(yx) = \min_{\substack{E\subseteq H\\ \operatorname{codim}(E)=n}} \max_{\substack{\xi\in E\\ \|\xi\|=1}} \|yx\xi\| \le \min_{\substack{E\subseteq H\\ \operatorname{codim}(E)=n}} \max_{\substack{\xi\in E\\ \|\xi\|=1}} \|y\|\cdot\|x\xi\| = \|y\|\cdot s_n(x).$$

Observe now that

$$s_n(xy) = s_n(y^*x^*) \le \|y^*\|\cdot s_n(x^*) = \|y\|\cdot s_n(x).$$ ☐

[a]Starting with $\|yx\xi\| \le \|y\|\cdot\|x\xi\|$, add max on the right then the left, and then add min on the left then the right.

(SV5) For $x \in K(H)$, $s_n(x) = \operatorname{dist}(x, F_n := \{\text{rank} \le n \text{ operators}\})$.

Proof. Write $x = \sum_i \lambda_i |e_i\rangle\langle f_i|$ in Schmidt representation. The operator $y := \sum_{i=0}^{n-1} \lambda_i |e_i\rangle\langle f_i|$ is in F_n and $x - y = \sum_{i\ge n} \lambda_i |e_i\rangle\langle f_i|$ has norm λ_n. Hence $\operatorname{dist}(x, F_n) \le \lambda_n$. Now for all $y \in F_n$, $\dim\operatorname{span}\{f_0, \dots, f_n\} = n+1$, so there is a $\xi \in F_n$ with $\|\xi\| = 1$ and $y\xi = 0$. Then

$$\|x - y\| \ge \|(x-y)\xi\| = \|x\xi\| \ge \lambda_n.$$ ☐

(SV6) If $x, y \in K(H)$, then $s_{m+n}(x+y) \le s_m(x) + s_n(y)$.

Proof. Let $\varepsilon > 0$. Using (SV5), take $z_1 \in F_m$ such that $\|x - z_1\| < s_m(x) + \varepsilon$ and take $z_2 \in F_n$ such that $\|y - z_2\| < s_n(y) + \varepsilon$. Then $z_1 + z_2 \in F_{m+n}$ and thus

$$\begin{aligned} s_{m+n}(x+y) = \operatorname{dist}(x+y, F_{m+n}) &\le \|x + y - (z_1 + z_2)\| \\ &\le \|x - z_1\| + \|y - z_2\| < s_m(x) + s_n(y) + 2\varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ was arbitrary, the result follows. ☐

2.5 The trace and the Schatten p-classes

Let (e_i) be an orthonormal basis of H. Define $\operatorname{Tr}\colon B(H)_+ \to [0,\infty]$ by $\operatorname{Tr}(x) := \sum_i \langle xe_i, e_i\rangle$.

Here are some basic properties about the trace.

(Tr1) Tr is positive-linear, i.e., $\operatorname{Tr}(\lambda x + y) = \lambda \operatorname{Tr}(x) + \operatorname{Tr}(y)$ for all $\lambda > 0$ and $x, y \in B(H)_+$.

(Tr2) Tr is lower semicontinuous on $B(H)_+$.

Proof. This follows immediately from the fact that each functional $x \mapsto \langle xe_i, e_i\rangle$ is continuous and $[0,\infty)$-valued together with the following exercise.

Exercise 2.5.1. Let X be a topological space and (f_n) a sequence of lower semicontinuous $[0,\infty)$-valued functions. Prove that $\sum f_n : X \to [0,\infty)$ defined by $(\sum f_n)(x) = \sum f_n(x)$ is again lower semicontinuous. □

(Tr3) $\operatorname{Tr}(x^*x) = \operatorname{Tr}(xx^*)$ for all $x \in B(H)$.

Proof. Since the sum of positive numbers is independent of ordering,

$$\begin{aligned}\sum_i \langle x^*xe_i, e_i\rangle &= \sum_i \langle xe_i, xe_i\rangle = \sum_{i,j} \langle \langle xe_i, e_j\rangle e_j, xe_i\rangle = \sum_{i,j} \langle xe_i, e_j\rangle \langle e_j, xe_i\rangle \\ &= \sum_{i,j} \langle x^*e_j, e_i\rangle \langle e_i, x^*e_j\rangle = \sum_{j,i} \langle x^*e_j, e_i\rangle \langle e_i, x^*e_j\rangle \\ &= \sum_{j,i} \langle \langle x^*e_j, e_i\rangle e_i, x^*e_j\rangle = \sum_j \langle x^*e_j, x^*e_j\rangle = \sum_j \langle xx^*e_j, e_j\rangle. \quad \square\end{aligned}$$

(Tr4) $\operatorname{Tr}(x) = \operatorname{Tr}(u^*xu)$ for all unitaries $u \in B(H)$ and $x \geq 0$. Hence if (f_i) is another orthonormal basis of H, then $\operatorname{Tr}(x) = \sum_i \langle xf_i, f_i\rangle$.

Proof. Write $x = \sqrt{x}^2$ so that by (Tr3),

$$\operatorname{Tr}(u^*xu) = \operatorname{Tr}((\sqrt{x}u)^*(\sqrt{x}u)) = \operatorname{Tr}((\sqrt{x}u)(\sqrt{x}u)^*) = \operatorname{Tr}(\sqrt{x}^2) = \operatorname{Tr}(x).$$

Now if (f_i) is another ONB, then define a unitary $v \in B(H)$ by $e_i \mapsto f_i$. Then

$$\operatorname{Tr}(x) = \operatorname{Tr}(u^*xu) = \sum_i \langle u^*xue_i, e_i\rangle = \sum_i \langle xue_i, ue_i\rangle = \sum_i \langle xf_i, f_i\rangle. \quad \square$$

(Tr5) If $x \geq 0$, then $\operatorname{Tr}(x) \geq \|x\|$.

Proof. If $x \geq 0$, then by (N5), there is a unit vector $\xi \in H$ such that $\langle x\xi, \xi\rangle = \max\{\lambda | \lambda \in \operatorname{sp}(x)\} = \|x\|$. Extend $\{\xi\}$ to an ONB $\{\xi\} \amalg (f_i)$, and observe that

$$\operatorname{Tr}(x) = \langle x\xi, \xi\rangle + \sum_i \langle xf_i, f_i\rangle \geq \langle x\xi, \xi\rangle = \|x\|. \qquad \square$$

Lemma 2.5.2.

1. *If $x \in K(H)$, then $\operatorname{Tr}(|x|^p) = \sum s_n(x)^p$.*
2. *If $\operatorname{Tr}(|x|^p) < \infty$ for some $p > 0$, then x is compact.*

Proof.

(1) Write $|x| = \sum \lambda_n |e_n\rangle\langle e_n|$ with $\lambda_n \searrow 0$ by Theorem 2.4.12 so that $|x|^p = \sum \lambda_n^p |e_n\rangle\langle e_n|$. Extending (e_n) to an ONB (e_i), we see

$$\operatorname{Tr}(x) = \sum_i \langle xe_i, e_i\rangle = \sum_n \lambda_n^p = \sum_n s_n(x)^p.$$

(2) Let (e_i) be an ONB and suppose $\varepsilon > 0$. There is a finite subset $F \subset I$ such that $\sum_{i \notin F} \langle |x|^p e_i, e_i\rangle < \varepsilon$. Let p_F denote the projection onto $\operatorname{span}\{e_i | i \in F\}$, and observe that

$$\begin{aligned}\||x|^{p/2}(1-p_F)\|^2 = \|(1-p_F)|x|^p(1-p_F)\| &\leq \operatorname{Tr}((1-p_F)|x|^p(1-p_F)) \\ &= \sum_{i\notin F} \langle |x|^p e_i, e_i\rangle < \varepsilon.\end{aligned}$$

Thus we may approximate $|x|^{p/2}$ by finite rank operators, so $|x|^{p/2}$ is compact, and thus so is $|x|^p$. Using the Spectral Theorem for compact normal operators 2.4.12, we can write $|x|^p = \sum \lambda_n |e_n\rangle\langle e_n|$ with $\lambda_n \searrow 0$. But then $|x| = \sum \lambda_n^{1/p} |e_n\rangle\langle e_n|$ and $\lambda_n^{1/p} \searrow 0$, so $|x|$ is compact by Lemma 2.4.11. Hence $x = u|x|$ is compact. $\square$

Definition 2.5.3. The Schatten p-class/p-ideal is the set

$$\mathcal{L}^p(H) := \left\{ x \in B(H) \middle| \operatorname{Tr}(|x|^p) = \sum s_n(x)^p < \infty \right\}.$$

We call $\mathcal{L}^1(H)$ the trace class operators and $\mathcal{L}^2(H)$ the Hilbert-Schmidt operators. Observe that $\mathcal{L}^p(H) \subset K(H)$ by Lemma 2.5.2.

Remark 2.5.4. Recall that when $1 \leq q \leq p$, $\ell^q \subseteq \ell^p$ with $\|\cdot\|_q \geq \|\cdot\|_p$. Since $\operatorname{Tr}(|x|^p) = \|(s_n(x))\|_{\ell^p}$, $\mathcal{L}^q(H) \subseteq \mathcal{L}^p(H)$ with $\|\cdot\|_q \geq \|\cdot\|_p$.

Lemma 2.5.5. *The Schatten p-class $\mathcal{L}^p(H)$ is a $*$-closed 2-sided ideal of $B(H)$ which is algebraically spanned by its positive operators.*

Proof.

<u>$*$-closed:</u> $s_n(x) = s_n(x^*)$ for all $n \geq 0$.

<u>$+$-closed:</u> $s_{2n}(x+y) \leq s_n(x)+s_n(y)$, so $(s_n(x)), (s_n(y)) \in \ell^p$ implies $(s_{2n}(x+y)) \in \ell^p$. Similarly, $s_{2n+1}(x+y) \leq s_n(x) + s_{n+1}(y)$, so $(s_n(x)), (s_n(y)) \in \ell^p$ implies $(s_{2n+1}(x+y)) \in \ell^p$. Thus $(s_n(x+y)) \in \ell^p$.

<u>ideal:</u> For all $x \in B(H)$ and $y \in \mathcal{L}^p(H)$, $s_n(xy), s_n(yx) \leq s_0(x)s_n(y) = \|x\| s_n(y)$, so $xy, yx \in \mathcal{L}^p(H)$.

<u>positive spanning:</u> Immediate by Fact 2.4.3. □

Corollary 2.5.6. $\mathcal{L}^1(H) = \operatorname{span}\{x \geq 0 | \operatorname{Tr}(x) < \infty\}$.

Proposition 2.5.7. Tr *extends to a linear map* $\mathcal{L}^1(H) \to \mathbb{C}$ *satisfying:*

- $x \leq y$ *implies* $\operatorname{Tr}(x) \leq \operatorname{Tr}(y)$ *(when x, y are self-adjoint) and*
- $|\operatorname{Tr}(x)| \leq \operatorname{Tr}(|x|)$.

Proof. For $x \in \mathcal{L}^1(H)$, we can write $x = \sum_{k=0}^3 i^k x_k$ with each $x_k \in \mathcal{L}^1(H)_+$. Define $\operatorname{Tr}(x) = \sum_{k=0}^3 i^k \operatorname{Tr}(x_k)$. This formula is clearly linear as long as it is well-defined. First, suppose x is self-adjoint. Since $\mathfrak{R}(x) = x_0 - x_2$ and $\mathfrak{I}(x) = x_1 - x_3 = 0$, we must have $x_1 = x_3$, so $x = x_0 - x_2$. If $x = y_0 - y_2$ for $y_0, y_2 \in \mathcal{L}^1(H)_+$, then

$$x_0 - x_2 = x = y_0 - y_2 \qquad \Longleftrightarrow \qquad x_0 + y_2 = y_0 + x_2.$$

Thus $\operatorname{Tr}(x_0) + \operatorname{Tr}(y_2) = \operatorname{Tr}(y_0) + \operatorname{Tr}(x_2)$, and since these numbers are finite, $\operatorname{Tr}(x_0) - \operatorname{Tr}(x_2) = \operatorname{Tr}(y_0) - \operatorname{Tr}(y_2)$. Now when x is arbitrary, if we can also write $x = \sum_{k=0}^3 i^k y_k$ with each $y_k \in \mathcal{L}^1(H)_+$, then $\mathfrak{R}(x) = y_0 - y_2$ and $\mathfrak{I}(x) = y_1 - y_3$. Hence $\sum_{k=0}^3 i^k \operatorname{Tr}(y_k) = \operatorname{Tr}(\mathfrak{R}(x)) - i\operatorname{Tr}(\mathfrak{I}(x))$ which is independent of the $y_k \geq 0$. Now suppose $x \leq y$ in $\mathcal{L}^1(H)$. Then $y - x \geq 0$, so $0 \leq \operatorname{Tr}(y-x) = \operatorname{Tr}(y) - \operatorname{Tr}(x)$. To prove the last relation, take a Schmidt decomposition $x = \sum_n s_n(x)|e_n\rangle\langle f_n|$ with (e_n) and (f_n) orthonormal. Then

$$\begin{aligned} |\operatorname{Tr}(x)| &= \left|\sum_i \left\langle \sum_n s_n(x)|e_n\rangle\langle f_n|f_i, f_i\right\rangle\right| = \left|\sum_n s_n(x)\langle e_n, f_n\rangle\right| \\ &\leq \sum_n s_n(x)|\langle e_n, f_n\rangle| = \sum_n s_n(x) = \operatorname{Tr}(|x|). \end{aligned}$$

□

Proposition 2.5.8. *For $x, y \in \mathcal{L}^2(H)$, $x^*y \in \mathcal{L}^1(H)$. The space $\mathcal{L}^2(H)$ is a Hilbert space with inner product $\langle x, y\rangle_{\mathcal{L}^2} := \operatorname{Tr}(y^*x)$.*

Proof. First, if $x \in \mathcal{L}^2(H)$ if and only if $x^*x \in \mathcal{L}^1(H)$ as $\operatorname{Tr}(|x|^2) = \operatorname{Tr}(x^*x)$. By polarization,

$$y^*x = \frac{1}{4}\sum_{k=0}^{3} i^k \underbrace{(x+i^k y)^* \underbrace{(x+i^k y)}_{\in\mathcal{L}^2(H)}}_{\in\mathcal{L}^1(H)}.$$

It is clear that $\langle\,\cdot\,,\,\cdot\,\rangle_{\mathcal{L}^2(H)}$ is a positive sesquilinear form. Definiteness follows from the estimate

$$\|x\|_2^2 := \operatorname{Tr}(x^*x) \underset{\text{(Tr5)}}{\geq} \|x^*x\| = \|x\|^2.$$

This also shows every $\|\cdot\|_2$-Cauchy sequence is $\|\cdot\|$-Cauchy. To see $\mathcal{L}^2(H)$ is complete with respect to $\|\cdot\|_2$, it suffices to prove that if (x_n) is $\|\cdot\|_2$-Cauchy with $x_n \to x$ in $\|\cdot\|$, then $x_n \to x$ in $\|\cdot\|_2$. First, $x \in K(H)$ as $K(H)$ is closed. Next, for all finite rank projections p,

$$\begin{aligned}\|(x-x_n)p\|_2^2 = \operatorname{Tr}(p(x-x_n)^*(x-x_n)p) &\overset{(!)}{=} \lim_m \operatorname{Tr}(p(x_m-x_n)^*(x_m-x_n)p)\\ &= \lim_m \operatorname{Tr}((x_m-x_n)p(x_m-x_n)^*) \leq \limsup_m \operatorname{Tr}((x_m-x_n)(x_m-x_n)^*)\\ &= \limsup_m \operatorname{Tr}((x_m-x_n)^*(x_m-x_n)) = \limsup_m \|x_m-x_n\|_2^2.\end{aligned}$$

In the equality marked (!) above, we are using the fact that there is only one trace on $B(pH) \cong M_k(\mathbb{C})$, where pH is a finite dimensional Hilbert space with dimension k. Thus $x_m \to x$ in norm implies $p(x_m-x_n)^*(x_m-x_n)p \to p(x-x_n)^*(x-x_n)p$ in norm, and we know the trace on $B(pH)$ is continuous.
Since p was arbitrary, we conclude that

$$\|x-x_n\|_2^2 \leq \limsup_m \|x_m-x_n\|_2^2,$$

which implies both $x \in \mathcal{L}^2(H)$ and $x_n \to x$ in $\|\cdot\|_2$. $\square$

Exercise 2.5.9. Suppose H is a Hilbert space (which you may assume is separable) with ONBs (e_i) and (f_i).

1. Show that for every $x \in \mathcal{L}^2(H)$, $\sum_{i,j} |\langle xe_j, f_i\rangle|^2 = \sum_n |s_n(x)|^2 = \sum_n \|xe_n\|^2$.

2. Show that for each $a = (a_{ij}) \in \ell^2(\mathbb{N}^2)$, there is an $a \in \mathcal{L}^2(H)$ such that $a_{ij} = \langle ae_j, f_i\rangle$.

3. Construct a unitary isomorphism $\mathcal{L}^2(H) \to \ell^2(\mathbb{N}^2)$.

4. Construct a canonical isomorphism $\mathcal{L}^2(H) \cong H \otimes H^*$.

Corollary 2.5.10. *For all $x \in \mathcal{L}^1(H)$ and $y \in \mathcal{B}(H)$, $|\operatorname{Tr}(xy)|, |\operatorname{Tr}(yx)| \le \|y\| \cdot \operatorname{Tr}(|x|)$.*

Proof. Since $xy \in \mathcal{L}^1(H)$, $|\operatorname{Tr}(xy)| \le \operatorname{Tr}(|xy|)$. Since $s_n(|xy|) \le \|y\| \cdot s_n(x)$ by (SV4),

$$\operatorname{Tr}(|xy|) = \sum s_n(|xy|) \le \sum \|y\| s_n(x) = \|y\| \sum s_n(x) = \|y\| \operatorname{Tr}(|x|).$$

Similarly, $\operatorname{Tr}(|yx|) \le \|y\| \operatorname{Tr}(|x|)$. □

Lemma 2.5.11. *For $x, y \in \mathcal{L}^2(H)$, $\operatorname{Tr}(xy) = \operatorname{Tr}(yx)$. The conclusion also holds for $x \in \mathcal{L}^1(H)$ and $y \in \mathcal{B}(H)$.*

Proof. As $(x, y) \mapsto \operatorname{Tr}(x^*y)$ and $(y, x) \mapsto \operatorname{Tr}(yx^*)$ are both sesquilinear forms on $\mathcal{L}^2(H)$, by polarization, they agree if and only if they agree on the diagonal. But $\operatorname{Tr}(x^*x) = \operatorname{Tr}(xx^*)$, so $\operatorname{Tr}(x^*y) = \operatorname{Tr}(yx^*)$ for all $x, y \in \mathcal{L}^2(H)$.
For the second part, by linearity in x, we may assume $x \in \mathcal{L}^1(H)_+$ so that $\sqrt{x} \in \mathcal{L}^2(H)_+$. We then calculate

$$\operatorname{Tr}(xy) = \operatorname{Tr}(\sqrt{x}(\sqrt{x}y)) = \operatorname{Tr}((\sqrt{x}y)\sqrt{x}) = \operatorname{Tr}(\sqrt{x}(y\sqrt{x})) = \operatorname{Tr}((y\sqrt{x})\sqrt{x}) = \operatorname{Tr}(yx).$$

□

Proposition 2.5.12. *$\mathcal{L}^1(H)$ is a Banach $*$-algebra with $\|x\|_1 := \operatorname{Tr}(|x|) = \sum s_n(x)$.*

Proof. We show $\|\cdot\|_1$ has the required properties.

Homogeneous: $\|\lambda x\|_1 = \operatorname{Tr}(|\lambda x|) = \operatorname{Tr}(|\lambda| \cdot |x|) = |\lambda| \operatorname{Tr}(|x|) = |\lambda| \cdot \|x\|_1$

Definite: $\|x\|_1 = \operatorname{Tr}(|x|) = 0$ implies $|x| = 0$, so $x = 0$.

Subadditive: Let $x+y = u|x+y|$ be the polar decomposition so that $|x+y| = u^*x+u^*y$. Since $u^*x, u^*y \in \mathcal{L}^1(H)$,

$$\begin{aligned}\|x+y\|_1 &= \operatorname{Tr}(|x+y|) = \operatorname{Tr}(u^*x + u^*y) = \operatorname{Tr}(u^*x) + \operatorname{Tr}(u^*y) \\ &\le |\operatorname{Tr}(u^*x)| + |\operatorname{Tr}(u^*y)| \le \|u^*\| \operatorname{Tr}(|x|) + \|u^*\| \operatorname{Tr}(|y|) \\ &\le \operatorname{Tr}(|x|) + \operatorname{Tr}(|y|) = \|x\|_1 + \|y\|_1.\end{aligned}$$

Submultiplicative: Let $xy = u|xy|$ be the polar decomposition so that $|xy| = u^*xy$. Then

$$\operatorname{Tr}(|xy|) = \operatorname{Tr}(u^*xy) \underset{\text{(Cor. 2.5.10)}}{\le} \underbrace{\|u^*x\|}_{=\||x|\|} \operatorname{Tr}(|y|) \underset{\text{(Tr5)}}{\le} \operatorname{Tr}(|x|) \operatorname{Tr}(|y|) = \|x\|_1 \cdot \|y\|_1.$$

$*$-isometric: $\|x\|_1 = \operatorname{Tr}(|x|) = \sum s_n(x) = \sum s_n(x^*) = \operatorname{Tr}(|x^*|) = \|x^*\|_1$.

Complete: Suppose (x_n) is $\|\cdot\|_1$-Cauchy. By (Tr5),

$$\|x_m - x_n\|_1 = \operatorname{Tr}(|x_m - x_n|) \geq \||x_m - x_n|\| = \|x_m - x_n\|,$$

so (x_n) is $\|\cdot\|$-Cauchy. Since $K(H)$ is closed, there is an $x \in K(H)$ with $x_n \to x$ in norm. Consider the polar decomposition $x - x_n = u_n|x - x_n|$. For all finite rank projections p,

$$\begin{aligned}\operatorname{Tr}(p|x - x_n|) &= \operatorname{Tr}(pu_n^*(x - x_n)p) = |\operatorname{Tr}(pu_n^*(x - x_n)p)| \\ &= \lim_m |\operatorname{Tr}(pu_n^*(x_m - x_n)p)| \underset{\text{(Cor. 2.5.10)}}{\leq} \limsup_m \|x_m - x_n\|_1.\end{aligned}$$

This implies $x \in \mathcal{L}^1(H)$ and $x_n \to x$ in $\|\cdot\|_1$. ☐

Proposition 2.5.13. *For all $1 < p < \infty$, $\mathcal{L}^p(H)$ is a Banach space with $\|x\|_p^p := \operatorname{Tr}(|x|^p) = \|(s_n(x))\|_{\ell^p}$.*

We omit the proof which is similar to those for $\mathcal{L}^2(H)$ and $\mathcal{L}^1(H)$. ☐

Theorem 2.5.14. *Suppose $1 < q, p < \infty$ with $1/p + 1/q = 1$. For all $x \in \mathcal{L}^p(H)$ and $y \in \mathcal{L}^q(H)$, $xy \in \mathcal{L}^1(H)$ and $|\operatorname{Tr}(xy)| \leq \|x\|_p \cdot \|y\|_q$.*

Proof. Without loss of generality, $2 \leq p$. We proceed via the following steps.

Step 1: If $x \in \mathcal{L}^p(H)_+$ with $p \geq 2$ and $\xi \in H$ with $\|\xi\| = 1$, then $\langle x^2\xi, \xi\rangle^{p/2} \leq \langle x^p\xi, \xi\rangle$.

Proof. Let (e_n) be an ONB with $x = \sum \lambda_n |e_n\rangle\langle e_n|$. For all $\xi \in \operatorname{span}\{e_1, \dots, e_k\}$,

$$\begin{aligned}\langle x^2\xi, \xi\rangle &= \sum_{i,j=1}^k \langle\langle \xi, e_i\rangle x^2 e_i, \langle \xi, e_j\rangle e_j\rangle \\ &= \sum_{i,j=1}^k \langle \xi, e_i\rangle \overline{\langle \xi, e_j\rangle} \langle x^2 e_i, e_j\rangle = \sum_{i=1}^k |\langle \xi, e_i\rangle|^2 \lambda_i^2.\end{aligned}$$

Since the function $r \mapsto r^{p/2}$ is convex and $\sum_{i=1}^k |\langle \xi, e_i\rangle|^2 = \|\xi\|^2 = 1$, we have

$$\langle x^2\xi, \xi\rangle^{p/2} = \left(\sum_{i=1}^k |\langle \xi, e_i\rangle|^2 \lambda_i^2\right)^{p/2} \leq \sum_{i=1}^k |\langle \xi, e_i\rangle|^2 \lambda_i^p = \langle x^p\xi, \xi\rangle.$$

Hence the desired inequality holds on the algebraic span of the e_i, which is dense in H. Since the continuous function $\xi \mapsto \langle x^p \xi, \xi\rangle - \langle x^2\xi, \xi\rangle^{p/2}$ is non-negative on a dense subspace, the result follows. $\square$

Step 2: If $x \in \mathcal{L}^p(H)_+$ with $p \geq 2$ and $y \in \mathcal{L}^q(H)_+$ with $1/p + 1/q = 1$, then $xy \in \mathcal{L}^1(H)$ and $\mathrm{Tr}(|xy|) \leq \|x\|_p \cdot \|y\|_q$.

Proof. Pick an ONB (f_n) such that $y = \sum \mu_n |f_n\rangle\langle f_n|$. For every $n \in \mathbb{N}$,

$$\begin{aligned} |\langle |xy| f_n, f_n\rangle|^2 &\underset{\text{(CS)}}{\leq} \||xy| f_n\|^2 \cdot \underbrace{\|f_n\|}_{=1} = |\langle |xy|^2 f_n, f_n\rangle| = |\langle y^* x^* xy f_n, f_n\rangle| \\ &= |\langle x^* xy f_n, y f_n\rangle| = \mu_n^2 |\langle |x|^2 f_n, f_n\rangle|. \end{aligned}$$

Hence by Step 1, we have

$$\langle |xy| f_n, f_n\rangle \leq \mu_n \langle |x|^2 f_n, f_n\rangle^{1/2} \underset{\text{(Step 1)}}{\leq} \mu_n \langle |x|^p f_n, f_n\rangle^{1/p}.$$

Now setting $a_n = \langle |x|^p f_n, f_n\rangle^{1/p}$, $(a_n) \in \ell^p$ as $x \in \mathcal{L}^p(H)$:

$$\|(a_n)\|_p^p = \sum_n \langle |x|^p f_n, f_n\rangle = \mathrm{Tr}(|x|^p) < \infty.$$

Also, $(\mu_n) \in \ell^q$ as $\sum_n \mu_n^q = \mathrm{Tr}(|y|^q) < \infty$ since $y \in \mathcal{L}^q(H)$. By Hölder's Inequality,

$$\begin{aligned} \mathrm{Tr}(|xy|) &= \sum_n \langle |xy| f_n, f_n\rangle \leq \sum_n \mu_n \langle |x|^p f_n, f_n\rangle^{1/p} \\ &\leq \|(a_n)\|_p \cdot \|(\mu_n)\|_q = \|x\|_p \cdot \|y\|_q. \end{aligned}$$

$\square$

Step 3: For arbitrary $x \in \mathcal{L}^p(H)$ with $p \geq 2$ and $y \in \mathcal{L}^q(H)$ with $1/p + 1/q = 1$, $xy \in \mathcal{L}^1(H)$ and $|\mathrm{Tr}(xy)| \leq \|x\|_p \cdot \|y\|_q$.

Proof. Consider the polar decompositions $x = u|x|$ and $y^* = v|y^*|$ and note that $|x|, |y^*| \geq 0$, $|x| = u^* x \in \mathcal{L}^p(H)$, and $|y^*| = v^* y^* \in \mathcal{L}^q(H)$. By Step 2, we have $|x| \cdot |y^*| \in \mathcal{L}^1(H)$ and

$$\mathrm{Tr}\left(\big| |x| \cdot |y^*| \big|\right) \leq \|x\|_p \cdot \|y\|_q.$$

It follows immediately that

$$xy = x(y^*)^* = u|x|(v|y^*|)^* = u|x||y^*|v^* \in \mathcal{L}^1(H).$$

and

$$|\operatorname{Tr}(xy)| = |\operatorname{Tr}(u|x||y^*|v^*)| \underset{\text{(Cor. 2.5.10)}}{\leq} \|u\| \cdot \|v^*\| \cdot \operatorname{Tr}\left(\big|\,|x| \cdot |y^*|\,\big|\right)$$

$$\underset{\text{(Step 2)}}{\leq} \||x|\|_p \cdot \||y^*|\|_q = \|x\|_p \cdot \|y\|_q. \qquad \square$$

Exercise 2.5.15. Show that the pairing $(x, y) \mapsto \operatorname{Tr}(xy)$ implements a duality exhibiting an isometric isomorphisms $K(H)^* \cong \mathcal{L}^1(H)$ and $\mathcal{L}^1(H)^* \cong B(H)$. Explain how one can view this as an analogy of the facts that $c_0^* \cong \ell^1$ and $(\ell^1)^* \cong \ell^\infty$.

Theorem 2.5.16. *Suppose $1 < p, q < \infty$ with $1/p + 1/q = 1$. The bilinear form $(x, y) := \operatorname{Tr}(xy)$ implements a duality exhibiting $\mathcal{L}^p(H)$ and $\mathcal{L}^q(H)$ as isometrically isomorphic to each other's dual spaces.*

Proof. First, note that if $(x_n) \in \ell^q$, then $(|x_n|^{q-1}) \in \ell^p$ and

$$\|x_n\|_q^q = \sum |x_n|^q = \sum |x_n|^{(q-1)p} = \|(|x_n|^{q-1})\|_p^p \qquad \text{and}$$
$$\|x_n\|_q^q = \left(\sum |x_n|^q\right)^{1/p+1/q} = \left(\sum |x_n|^{(q-1)p}\right)^{1/p} \left(\sum |x_n|^q\right)^{1/q} = \|(|x_n|^{q-1})\|_p \cdot \|(x_n)\|_q.$$

We now proceed via the following steps.

Step 1: The map $y \mapsto \operatorname{Tr}(\,\cdot\, y)$ is an isometry $\mathcal{L}^q(H) \to \mathcal{L}^p(H)^*$.

Proof. First, note that the map $\mathcal{L}^q(H) \to \mathcal{L}^p(H)^*$ given by $y \mapsto \operatorname{Tr}(\,\cdot\, y)$ is well-defined and norm-decreasing by Theorem 2.5.14. We use polar decomposition to write $y = u|y|$ and note $|y| = u^*y \in \mathcal{L}^q(H)$.
We claim that

Claim. *For every $r > 0$, $s_n(|y|)^r = s_n(|y|^r) = s_n(u|y|^r) = s_n(|y|^r u^*)$.*

Proof of claim. If $|y| = \sum \lambda_n |f_n\rangle\langle f_n|$ is the Schmidt decomposition, then $s_n(|y|)^r = \lambda_n^r = s_n(|y|^r)$. Moreover, if $e_n = uf_n$ for all n, then

$$u|y| = \sum \lambda_n |e_n\rangle\langle f_n| \qquad \Longrightarrow \qquad u|y|^r = \sum \lambda_n^r |e_n\rangle\langle f_n|.$$

Then since $(u|y|^r)^* u|y|^r = |y|^r u^* u|y|^r = \sum \lambda_n^{2n} |f_n\rangle\langle f_n|$,

$$s_n(u|y|^r) = s_n(|y|^r u^* u|y|^r)^{1/2} = \lambda_n^r.$$

Since for any z, $s_n(z^*z)^{1/2} = s_n(z)$, we have $s_n(u|y|^r) = \lambda_n^r$. Finally, $s_n(u|y|^r) = s_n(|y|^r u^*)$ as the n-th singular value of adjoints agree, finishing the claim. $\square$

Now using the claim above, we have $a_n := s_n(|y|^{q-1}) = s_n(|y|)^{q-1}$, so $(a_n) \in \ell^p$ and $|y|^{q-1} \in \mathcal{L}^p(H)$. For $x := |y|^{q-1}u^* \in \mathcal{L}^p(H)$, setting $\mu_n = s_n(y)$,

$$\begin{aligned}\operatorname{Tr}(xy) &= \operatorname{Tr}(|y|^{q-1}u^*y) = \operatorname{Tr}(|y|^q) = \|y\|_q^q \\ &= \|(\mu_n)\|_q^q = \|(\mu_n^{q-1})\|_p \cdot \|(\mu_n)\|_q = \|x\|_p \cdot \|y\|_q \end{aligned} \qquad \square$$

Step 2: The map $y \mapsto \operatorname{Tr}(\,\cdot\, y)$ from Step 4 is surjective.

Proof. Since $1 < p$, $\mathcal{L}^1(H) \subseteq \mathcal{L}^p(H)$ with $\|\cdot\|_1 \geq \|\cdot\|_p$. Thus if $\varphi \in \mathcal{L}^p(H)^*$, $\varphi|_{\mathcal{L}^1(H)} \in \mathcal{L}^1(H)^* = B(H)$, so there is a $y \in B(H)$ such that $\varphi|_{\mathcal{L}^1(H)} = \operatorname{Tr}(\,\cdot\, y)$ by Exercise 2.5.15. It remains to prove $y \in \mathcal{L}^q(H)$ and $\varphi = \operatorname{Tr}(\,\cdot\, y)$ on $\mathcal{L}^p(H)$.

Claim. *$y \in K(H)$.*

Proof of Claim. By polar decomposition $y = u|y|$, we may assume $y \geq 0$ as $y \in K(H)$ iff $|y| \in K(H)$, and

$$|\operatorname{Tr}(x|y|)| = |\operatorname{Tr}(xu^*y)| \leq \|\varphi\| \cdot \|xu^*\|_p \underset{\text{(SV4)}}{\leq} \|\varphi\| \cdot \|x\|_p.$$

If $y \notin K(H)$, then by Remark 2.4.13, there is a $\varepsilon > 0$ such that $p := \chi_{(\varepsilon,\infty)}(y)$ has infinite dimensional image. Pick an orthonormal sequence $(f_n) \subset pH$, and note that $y \geq \varepsilon$ on pH, i.e., $\langle yf_n, f_n\rangle \geq \varepsilon$ for all n. Pick $(\mu_n) \in \ell^p \setminus \ell^1$ (we may assume $\mu_n \geq 0$ for all n) and set $x_k = \sum_{n=0}^k \mu_n |f_n\rangle\langle f_n|$ and $x = \lim x_k \in \mathcal{L}^p(H)$. Then $x_k \in \mathcal{L}^1(H)$ for all k, and

$$\varepsilon \sum_{n=0}^k \mu_n \leq \sum_{n=0}^k \mu_n \langle yf_n, f_n\rangle = \operatorname{Tr}(x_k y) = \varphi(x_k) \xrightarrow{k\to\infty} \varphi(x).$$

But $(\mu_n) \notin \ell^1$, so $\varepsilon \sum_{n=0}^k \mu_n \to \infty$, a contradiction. $\square$

Since $y \in K(H)$, we can take a Schmidt decomposition $|y| = \sum \lambda_n |f_n\rangle\langle f_n|$, and let $y = u|y|$ be the polar decomposition with $uf_n = e_n$ so that $y = \sum \lambda_n |e_n\rangle\langle f_n|$. For each k, let r_k be the orthogonal projection onto $\operatorname{span}\{f_0, f_1, \dots, f_k\}$, and observe that r_k commutes with $|y|^s$ for all $s > 0$. For each k, $x_k := |y|^{q-1} r_k u^*$ is finite rank and thus in $\mathcal{L}^2(H) \subseteq \mathcal{L}^p(H)$. Observe now that

$$x_k^* x_k = u|y|^{q-1} r_k |y|^{q-1} u^* = u r_k \left(\sum \lambda_n^{2q-2} |f_n\rangle\langle f_n|\right) u^* = \sum_{n=0}^k \lambda_n^{2q-2} |e_n\rangle\langle e_n|$$

which implies that

$$\|x_k\|_p^p = \mathrm{Tr}((x_k^* x_k)^{p/2}) = \sum_{n=0}^{k} (\lambda_n^{2q-2})^{p/2} = \sum_{n=0}^{k} \lambda_n^q = \mathrm{Tr}(|y|^q r_k).$$

But note that also

$$\varphi(x_k) = \mathrm{Tr}(x_k y) = \mathrm{Tr}(|y|^{q-1} r_k u^* y) = \mathrm{Tr}(|y|^{q-1} r_k |y|) = \mathrm{Tr}(|y|^q r_k).$$

This means

$$\mathrm{Tr}(|y|^q r_k) = |\varphi(x_k)| \leq \|\varphi\| \cdot \|x_k\|_p = \|\varphi\| \cdot \mathrm{Tr}(|y|^q r_k)^{1/p}$$

which implies that

$$\mathrm{Tr}(|y|^q r_k)^{1/q} = \mathrm{Tr}(|y|^q r_k)^{1-1/p} \leq \|\varphi\|.$$

Hence $\mathrm{Tr}(|y|^q r_k) \leq \|\varphi\|^q$ for all k, and so $y \in \mathcal{L}^q(H)$.
Finally, the finite rank operators are contained in $\mathcal{L}^2(H)$ and also dense in $\mathcal{L}^p(H)$. Indeed, if $x \in \mathcal{L}^p(H)^+$ has Schmidt decomposition $x = \sum \lambda_n |f_n\rangle\langle f_n|$, then $x_k := \sum_{n=0}^{k} \lambda_n |f_n\rangle\langle f_n|$ is finite rank, and

$$\|x - x_k\|_p^p = \left\| \sum_{n>k} \lambda_n |f_n\rangle\langle f_n| \right\|_p^p = \sum_{n>k} \lambda_n^p \xrightarrow{k\to\infty} 0.$$

Thus $\mathcal{L}^2(H)$ is dense in $\mathcal{L}^p(H)$, and so $\varphi = \mathrm{Tr}(\,\cdot\, y)$ on $\mathcal{L}^p(H)$. □

Since our proof above did not distinguish p and q, we also conclude $\mathcal{L}^p(H) \cong \mathcal{L}^q(H)^*$. □