Chapter 3

Operators on Hilbert Space

3.1 Topics to be covered

- Operators on Hilbert spaces

  Special families of operators: adjoints, projections, Hermitian, unitaries, partial isometries, polar decomposition

  Density matrices and trace class operators

  B(H) as dual of trace class

- Spectral Theory

  Spectrum and resolvent

  Spectrum versus point spectrum; approximate points spectrum

  Neumann series

  $\sigma(T)$ is non-empty and compact

  Unbounded operators and self-adjointness problems

- The Riesz functional calculus

  The functional calculus for normal operators

- Compact operators

  Singular values and Schmidt's theorem

  Schmidt decomposition of tensors

  The Schatten classes

- Unbounded operators

  Hellinger-Toeplitz theorem

  Closable and non-closable operators

  Self-adjointness problems

  Stone's theorem

3.2 Operators on Hilbert Space

Definition 3.1. Let $\mathcal{H}$ be a Hilbert space. We write $h \perp k \iff \langle h\ k\rangle = 0$. Let $S \subseteq \mathcal{H}$ be a subset, then we set

$$S^{\perp} = \{h : \langle h|k\rangle = 0, \forall k \in S\}.$$

It is easy to see that $S^{\perp}$ is a closed subspace of $\mathcal{H}$.

Proposition 3.2. *Let $\mathcal{M} \subseteq \mathcal{H}$ be a closed subspace. Then every vector $x \in \mathcal{H}$ has a decomposition as $x = h + k$ with $h \in \mathcal{M}$ and $k \in \mathcal{M}^{\perp}$ and this decomposition is unique.*

Proof. First uniqueness. Suppose that $x = h_1 + k_1 = h_2 + k_2$ with $h_1, h_2 \in \mathcal{M}$ and $k_1, k_2 \in \mathcal{M}^{\perp}$. Subtracting we have $h_1 - h_2 = k_2 - k_1 := v$ and since these are both subspaces, we have that $v \in \mathcal{M} \cap \mathcal{M}^{\perp}$. But this implies that $\langle v|v\rangle = 0$ and so $v = 0$ which implies the uniqueness.

Now we show existence. Since $\mathcal{M}$ is itself a HIlbert space, it has an o.n.b. Let $\{e_a; a \in A\}$ be an o.n.b. for $\mathcal{M}$ and set

$$P(x) = \sum_{a \in A} \langle e_a|x\rangle e_a.$$

By Bessel's inequality $\sum_{a \in A} |\langle e_a|x\rangle|^2$ converges and this is enough to show that the sum for $P(x)$ converges and that $\|P(x)\| \leq \|x\|$. Hence, the formula defines a bounded linear operator.

Next note that $\langle x - P(x)|e_a\rangle = 0$ for every $a \in A$. Hence, $x - P(x) \in \mathcal{M}^{\perp}$. So that $x = P(x) + (x - P(x))$ is the decomposition. □

Corollary 3.3. *Let $\{e_a : a \in A\}$ and $\{f_b : b \in B\}$ be two o.n.b.'s for $\mathcal{M}$. Then for every $x \in \mathcal{H}$, we have that $\sum_{a \in A} \langle e_a|x\rangle e_a = \sum_{b \in B} \langle f_b|x\rangle f_b$.*

Proof. This follows from the uniqueness of the decomposition. □

Definition 3.4. We call the map P given above the **orthogonal projection of $\mathcal{H}$ onto $\mathcal{M}$**.

Corollary 3.5. *Let $\mathcal{H}$ be a HIlbert space and $\mathcal{M} \subseteq \mathcal{H}$ a subspace. If $\{e_a : a \in A\}$ is an o.n.b. for $\mathcal{M}$ then there exists an orhtonormal set $\{f_b : b \in B\}$ such that $\{e_a : a \in A\} \cup \{f_b : b \in B\}$ is an o.n.b. for $\mathcal{H}$.*

Proof. Let $\{f_b : b \in B\}$ be an o.n.b. for $\mathcal{M}^\perp$. □

The above result is often summarized by saying that every o.n.b. for $\mathcal{M}$ can be **extended** to an o.n.b. for $\mathcal{H}$.

3.2.1 The dual of a Hilbert Space

Let $\mathcal{H}$ be a Hilbert space and let $h \in \mathcal{H}$ then the map $f_h : \mathcal{H} \to \mathbb{F}$ defined by

$$f_h(k) = \langle h\ k \rangle,$$

is linear and $|f_h(k)| \leq \|h\|\|k\|$ so it is a bounded linear functional with $\|f_h\| \leq \|h\|$. Since

$$\|h\|^2 = f_h(h) \leq \|f_h\|\|h,$$

we see that $\|f_h\| = \|h\|$. We now show that every bounded linear functional is of this form.

Theorem 3.6. *Let $\mathcal{H}$ be a HIlbert space and let $f : \mathcal{H} \to \mathbb{F}$ be a bounded linear functional. Then there exists a vector $h \in \mathcal{H}$ such that $f = f_h$. Consequently, the map*

$$L : \mathcal{H} \to \mathcal{H}^*,\ L(h) = f_h,$$

is a one-to-one isometric map onto the dual space that is linear in the $\mathbb{R}$ case and conjugate linear in the $\mathbb{C}$ case.

Proof. The case of the 0 functional is trivia, so assume tht $f \neq 0$. Let $\mathcal{K} = ker(f)$. Since f is continuous this is a closed subspace of $\mathcal{H}$ and is not all of $\mathcal{H}$. Hence, $\mathcal{K}^\perp \neq (0)$. Pick $k_0 \in \mathcal{K}^\perp$ with $f(k_0) = 1$. Then for every $h \in \mathcal{H}$, we have that $f(h - f(h)k_0) = 0$ so that $h - f(h)k_0 \in \mathcal{K}$.

Hence, this vector is orthogonal to k_0 and

$$0 = \langle k_0 | h - f(h)k_0 \rangle = \langle k_0 | h \rangle - f(h)\|k_0\|^2.$$

Thus, $f(h) = \langle \frac{k_0}{\|k_0\|^2} | h \rangle$ and so $f = f_{k_1}$ with $k_1 = \frac{k_0}{\|k_0\|^2}$.

Finally, note that $L(h_1 + h_2) = f_{h_1+h_2} = f_{h_1} + f_{h_2} = L(h_1) + L(h_2)$, while $L(\lambda h) = f_{\lambda h} = \overline{\lambda} f_h = \overline{\lambda} L(h)$. □

3.2.2 The Hilbert Space Adjoint

Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces and let $T : \mathcal{H} \to \mathcal{K}$ be a bounded linear map. Then for each $k \in \mathcal{K}$ the map from $\mathcal{H}$ to $\mathbb{F}$ given by

$$h \to g_k(h) = \langle k|T(h)\rangle_{\mathcal{K}},$$

is easily seen to be linear and it is bounded since,

$$|\langle k|T(h)\rangle| \leq \|k\|\|T(h)\| \leq (\|k\|\|T\|)\|h\|.$$

Thus, it is a bounded linear functional and hence there is a unique vector, denoted $T^*(k) \in \mathcal{H}$ such that

$$\langle T^*(k)|h\rangle_{\mathcal{H}} = \langle k|T(h)\rangle_{\mathcal{K}}.$$

This last equation is known as the **adjoint equation**.

It is easy to check that the map $T^* : \mathcal{K} \to \mathcal{H}$ is linear. Also, $\|T^*(k)\| = \|g_k\| \leq (\|k\|\|T\|)$ so that T^* is bounded with $\|T^*\| \leq \|T\|$. Equality follows since

$$\|T^*\| = \sup\{|\langle T^*(k)|h\rangle| : \|k\| \leq 1, \|h\| \leq 1\} = \sup\{|\langle k|T(h)\rangle| : \|k\| \leq 1, \|h\| \leq 1\} = \|T\|.$$

Here are some basic properties:

- $(T_1 + T_2)^* = T_1^* + T_2^*$,
- $(\lambda T)^* = \bar{\lambda} T^*$,
- $(T_1 T_2)^* = T_2^* T_1^*$,
- $(T^*)^* = T$,
- if $T = (t_{i,j}) : \mathbb{C}^m \to \mathbb{C}^n$ is a $n \times m$ matrix and we regard $\mathbb{C}^k$ as a Hilbert space in the usual way, then $T^* = (b_{i,j})$ where $b_{i,j} = \bar{t}_{j,i}$, i.e., T^* is the conjugate transpose.

If $T : \mathcal{H} \to \mathcal{K}$, we let

$$\mathcal{R}(T) = \{T(x) : x \in \mathcal{H}\} \subseteq \mathcal{K}$$

denote the **range of** T, and let

$$\mathcal{N}(T) = \{x \in \mathcal{H} : T(x) = 0\} \subseteq \mathcal{H},$$

denote the **kernel of T** also called the **nullspace of T**.

Proposition 3.7. $\mathcal{R}(T)^\perp = \mathcal{N}(T^*)$ *and* $\mathcal{R}(T)^- = \mathcal{N}(T^*)^\perp$.

3.2.3 The Banach Space Adjoint

Operators on normed spaces also have adjoints but there is a bit of a difference. Given X, Y normed spaces and $T : X \to Y$ a bounded linear map, for each $f \in Y^*$–the dual space, we have that $f \circ T : X \to \mathbb{F}$ is a bounded linear functional, i.e., $f \circ T \in X^*$.

We define $T^* : X^* \to Y^*$ by $T^*(f) = T \circ f$.

We have that:

- $(T_1 + T_2)^* = T_1^* + T_2^*$,
- $(\lambda T)^* = \lambda T^*$,
- $(T_1T_2)^* = T_2^*T_1^*$,
- if $T = (t_{i,j}) : \mathbb{F}^m \to \mathbb{F}^n$ is a matrix, both spaces have a norm, and we identify the dual spaces as $\mathbb{F}^m$ and $\mathbb{F}^n$ with the inner product pairing and a dual norm, then $T^* = (b_{i,j})$ where $b_{i,j} = t_{j,i}$, i.e., T^* is the transpose.

Finally, if $T : X \to Y$, then $T^{**} : X^{**} \to Y^{**}$ and generally $X^{**} \neq X$ and $Y^{**} \neq Y$ so $T^{**} \neq T$. However, we do have that $T^{**}(\hat{x}) = \widehat{T(x)}$ where $\hat{x} \in X^{**}$ is the functional defined by $\hat{x}(f) = f(x)$.

3.3 Some Important Classes of Operators

In this section we introduce some of the important types of operators that we will encounter later. One common theme is to find spatial characterizations of operators that are given by an algebraic characterization.

Definition 3.8. A bounded linear operator $P : \mathcal{H} \to \mathcal{H}$ is called an **idempotent** if $P^2 = P$.

Proposition 3.9. *If P is an idempotent then $\mathcal{R}(P)$ is a closed subspace and every $h \in \mathcal{H}$ decomposes uniquely as $h = h_1 + h_2$ with $h_1 \in \mathcal{R}(P)$ and $Ph_2 = 0$.*

Proof. If $Ph_n \in \mathcal{R}(P)$ and $P(h_n) \to k$, then $P^2(h_n) \to P(k)$, but $P^2(h_n) = P(h_n)$ so $k = P(k)$, so k is also in the range of P.

Now $h = Ph + (I - P)h$ with $h_1 = Ph \in \mathcal{R}(P)$ and $P(I - P)h = (P - P^2)h = 0$ so $h_2 = (I - P)h \in \mathcal{N}(P)$.

To see uniqueness, if $h = h_1 + h_2 = k_1 + k_2$ is two ways to decompose h as a sum, then $h_1 - k_1 = k_2 - h_2$ with $h_1 - k_1 \in \mathcal{R}(P)$ and $k_2 - h_2 \in \mathcal{N}(P)$. This implies that

$$h_1 - k_1 = P(h_1 - k_1) = P(k_2 - h_2) = 0,$$

so that $h_1 = k_1$ and similarly, $h_2 = k_2$. □

In general, we need not have that $\mathcal{R}(P)$ and $\mathcal{N}(P)$ are orthogonal when P is idempotent.

A good example is to consider $P = \begin{pmatrix} 1 & t \\ 0 & 0 \end{pmatrix}$ then $\mathcal{R}(P) = span\{e_1\}$ while $\mathcal{N}(P) = span\{-te_1 + e_2\}$ so that these two subspaces are orthogonal iff $t = 0$.

Proposition 3.10. *If $P \in B(\mathcal{H})$ and $P = P^2 = P^*$ then P is the orthogonal projection onto its range.*

Proof. By the above every vector decomposes uniquely as $h = h_1 + h_2$ with $h_1 \in \mathcal{R}(P)$ and $h_2 \in \mathcal{N}(P)$. But

$$\langle h_1 | h_2 \rangle = \langle P h_1 | h_2 \rangle = \langle h_1 | P^* h_2 \rangle = \langle h_1 | P h_2 \rangle = 0,$$

since $h_2 \in \mathcal{N}(P)$. Hence, $h_1 \perp h_2$. □

By a **projection** we will always mean an operator P such that $P = P^2 = P^*$, i.e., an orthogonal projection.

Definition 3.11. Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. A map $V : \mathcal{H} \to \mathcal{K}$ is called an **isometry** iff $\|Vh\| = \|h\|, \forall h \in \mathcal{H}$. A map $U : \mathcal{H} \to \mathcal{K}$ is called a **unitary** if it is an onto isometry.

Note that an isometry is always one-to-one since $h \in \mathcal{N}(V) \iff \|Vh\| = 0 \iff \|h\| = 0 \iff h = 0$. Also the range of an isometry is easily seen to be closed.

The **unilateral shift** $S : \ell^2 \to \ell^2$ given by $S((x_1, x_2, \ldots)) = (0, x_1, x_2, \ldots)$ is easily seen to be an isometry and is not onto.

Proposition 3.12. *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces and let $V : \mathcal{H} \to \mathcal{K}$ be linear. The following are equivalent:*

1. *V is an isometry,*
2. *$\forall h_1, h_2 \in \mathcal{H}$, $\langle V h_1 | V h_2 \rangle_{\mathcal{K}} = \langle h_1 | h_2 \rangle_{\mathcal{H}}$,(this property is called* ***inner product preserving****)*

3. $V^*V = I_{\mathcal{H}}$.

Problem 3.13. Prove the above proposition.

Proposition 3.14. *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces and let $U : \mathcal{H} \to \mathcal{K}$. The following are equivalent:*

1. *U is a unitary,*
2. *U is an invertible siometry,*
3. *U is invertible and $U^{-1} = U^*$,*
4. *U and U^* are both isometries.*

Definition 3.15. Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. A map $W : \mathcal{H} \to \mathcal{K}$ is called a **partial isometry** provided that $\|Wh\| = \|h\|, \forall h \in \mathcal{N}(W)^\perp$. We shall call $\mathcal{N}(W)^\perp$ the **initial space of W** and $\mathcal{R}(W)$ the **final space of W**.

Proposition 3.16. *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces and let $W \in B(\mathcal{H}, \mathcal{K})$. The following are equivalent:*

1. *W is a partial isometry,*
2. *W^*W is a projection,*
3. *WW^* is a projection,*
4. *W^* is a partial isometry.*

Moreover, we have that W^*W is the projection onto the initial space of W and WW^* is the projection onto the final space of W.

Definition 3.17. Let $\mathcal{H}$ be a Hilbert space. A bounded linear map $H : \mathcal{H} \to \mathcal{H}$ is called **self-adjoint** or **Hermitian** if $H = H^*$.

Proposition 3.18 (Cartesian Decomposition)**.** *Let $\mathcal{H}$ be a Hilbert space and let $T \in B(\mathcal{H})$. Then there are unique self-adjoint operators, H, K such that $T = H + iK$.*

Proof. Let
$$H = \frac{T + T^*}{2} \text{ and } K = \frac{T - T^*}{2i},$$
then $H = H^*, K = K^*$ and $T = H + iK$. If we also had $T = H_1 + iK_1$, then $H_1 = \frac{T+T^*}{2} = H$ and $K_1 = \frac{T-T^*}{2i}$ and so uniqueness follows. □

The notation $Re(T) = \frac{T+T^*}{2}$ and $Im(T) = \frac{T-T^*}{2i}$ is often used.

Definition 3.19. Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. A bounded linear map $T : \mathcal{H} \to \mathcal{K}$ is called **finite rank** provided that $dim(\mathcal{R}(T))$ is finite.

Proposition 3.20. *T is finite rank if and only if for some n there are vectors $h_1, ..., h_n \in \mathcal{H}$ and $k_1, ..., k_n \in \mathcal{K}$ such that*

$$T(h) = \sum_{i=1}^{n} \langle h_i | h \rangle k_i.$$

Proof. If T has this fomr then $\mathcal{R}(T) \subseteq span\{k_1, ..., k_n\}$ so that T is finite rank.

Conversely, assume that T is finite rank and choose an orthonormal basis $\{k_1, ..., k_n\}$ for $\mathcal{R}(T)$. By Parseval,

$$Th = \sum_{i=1}^{n} \langle k_i | Th \rangle k_i = \sum_{i=1}^{n} \langle T^* k_i | h \rangle k_i,$$

so we have the desired form with $h_i = T^* k_i$. □

In physics the above operator is often denoted

$$T = \sum_{i=1}^{n} |k_i\rangle \langle h_i|.$$

Definition 3.21. Let $\mathcal{H}$ and $\mathcal{K}$ be HiIbert spaces. A bounded linear map $K : \mathcal{H} \to \mathcal{K}$ is called **compact** if it is in the closure of the finite rank operators. We let $\mathbb{K}(\mathcal{H}, \mathcal{K})$ denote the set of compact operators.

It is not hard to see that $\mathbb{K}(\mathcal{H}, \mathcal{K})$ is a closed linear subspace of $B(\mathcal{H}, \mathcal{K})$. We will return to the compact operators later.

3.4 Spectral Theory

For a finite square matrix, its eigenvalues, which are characterized as the roots of the **characteristic polynomial**, $p_A(t) = det(tI - A)$, play a special role in understanding the matrix. This polynomial need not factor over $\mathbb{R}$ which is why we generally assume that our spaces are complex when studying eigenvalues.

The spectrum is the set that plays a similar role in infinite dimensions. For similar reasons we will from now on assume that our Hilbert space is over $\mathbb{C}$.

Definition 3.22. Let $T \in B(\mathcal{H})$. Then the **spectrum of T** is the set

$$\sigma(T) = \{z \in \mathbb{C} : (zI - T) \text{ is not invertible}\},$$

the complement of this set is called the **resolvent of T** and is denoted $\rho(T)$.

Theorem 3.23 (Neumann Series). *Let $\mathcal{H}$ be a Hilbert space and let $A, B, X, T \in B(\mathcal{H})$.*

1. *If $\|X\| < 1$, then $I - X$ is invertible and the series $\sum_{n=0}^{\infty} X^n$ converges in norm to $(I - X)^{-1}$.*
2. *If A is invertible and $\|A - B\| < \|A^{-1}\|^{-1}$, then B is invertible.*
3. *If $|z| > \|T\|$, then $(zI - T)$ is invertible.*

Theorem 3.24. *Let $T \in B(\mathcal{H})$, then $\sigma(T)$ is a non-empty compact set.*

Proof. By (3), $\sigma(T) \subseteq \{z : |z| \leq \|T\|\}$, so $\sigma(T)$ is bounded. By (2), the set of invertible operators is open, so $\rho(T)$ is open and so $\sigma(T)$ is closed. Thus, $\sigma(T)$ is compact.

The deepest part is that $\sigma(T)$ is non-empty. This uses some non-trivial complex analysis. Suppose that $\sigma(T)$ is empty. Fix $h, k \in \mathcal{H}$ and define $f_{h,k} : \mathbb{C} \to \mathbb{C}$, by

$$f_{h,k}(z) = \langle h | (zI - T)^{-1} k \rangle.$$

One shows that this function is analytic on all of $\mathbb{C}$ and uses the Neumann series to show that as $|z| \to +\infty$ we have $|f_{h,k}(z)| \to 0$. This in turn implies that $f_{h,k}$ is bounded. A result in complex analysis says that any bounded entire function is constant. Since it tends to 0 it must be the 0 function. Setting $z = 0$ we see that $\langle h \; T^{-1} k \rangle = 0$ for every pair of vectors. Choosing $h = T^{-1}k$ yields that $\|T^{-1}k\| = 0$ so that $T^{-1} = 0$, a contradiction.

Hence, $\sigma(T)$ must be non-empty. □

Definition 3.25. We let $\sigma_p(T)$ denote the set of eigenvalues of T, i.e., the $\lambda \in \mathbb{C}$ such that there eixsts $h \neq 0$ with $Th = \lambda h$.

If $\lambda \in \sigma_p(T)$ then $\mathcal{N}(\lambda I - T) \neq (0)$ so is not invertible. Thus, $\sigma_p(T) \subseteq \sigma(T)$. In finite dimensions we have that $\sigma(T) = \sigma_p(T)$. Also in finite dimensions, if λ is an eigenvalue of T then $\overline{\lambda}$ is an eigenvalue of T^*. The following example shows that both of these are far from the case in infinite dimensions.

First a little result.

Proposition 3.26. $\sigma(T) = \{\overline{\lambda} : \lambda \in \sigma(T)\}$.

Proof. If $(\lambda I - T)^{-1} = R$ then check that R^* is the inverse of $(\overline{\lambda} I - T^*)$. This means that if $\lambda \in \rho(T)$, then $\overline{\lambda} \in \rho(T^*)$ and the result follows. □

Example 3.27. Let $S : \ell^2 \to \ell^2$ denote the unilateral shift. Since this is an isometry, $Sx = 0 \implies x = 0$ so $0 \notin \sigma_p(S)$. Now let $\lambda \neq 0$ and suppose that $Sx = \lambda x$ with $x = (x_1, x_2, ...)$. We have that $(\lambda x_1, \lambda x_2, ...) = (0, x_1, x_2, ...)$ so that $\lambda x_1 = 0 \implies x_1 = 0$ but then, $\lambda x_2 = x_1 \implies x_2 = 0$ and, inductively, $x_n = 0$ for all n. Hence $\lambda \notin \sigma_p(S)$.

Recall $S^*((x_1, x_2, ...)) = (x_2, x_3, ...)$. So $S^* x = \mu x$ if and only if $\mu x_n = x_{n+1}$. Now if $x_1 = 0$ then all entries would be 0, so we can assume that $x_1 \neq 0$ and that after scaling $x_1 = 1$. Then shows that then $x_n = \mu^{n-1}$. So that the eigenvector would need to have the form

$$(1, \mu, \mu^2, ...).$$

Now it is not hard to see that this vector is in ℓ^2 iff $|\mu| < 1$. Thus,

$$\sigma_p(S^*) = \{\mu : |\mu| < 1\} \subseteq \sigma(S^*).$$

Since $\|S^*$
$+ 1$ we have that $\sigma(S^*) \subseteq \{\mu; |\mu| \leq 1\}$, so using the fact that it is a closed set, we have that

$$\sigma(S^*) = \{\mu : |\mu| \leq 1\},$$

and by the last result,

$$\sigma(S) = \{\lambda : |\lambda| \leq 1\}.$$

Example 3.28. Let $\mathcal{H} = L^2([0,1], \mathcal{M}, m)$ be the Hilbert space of equivalence classes of square-integrable measurable functions with respect to Lebesgue measure. Recall that two functions are equivalent iff they are equal almost everywhere(a.e.). We shall write $[f]$ for the equivalence class of a function. We define M_t to be the operator of multiplication by the variable t, so that $[f] = [tf(t)]$. It is easy to check that this is a bounded linear operator and is self-adjoint.

In finite dimensions we are often use that every self-adjoint matrix an orthonormal basis of eigenvectors with real eigenvalues. We show that M_t has no eigenvalues. To see this suppose that $M_t[f] = \lambda[f]$ then we would have that $tf(t) = \lambda f(t)$ a.e. so that $(t - \lambda)f(t) = 0$ a.e. But this implies that $f(t) = 0$ a.e. and so $[f] = 0$. Thus, $\sigma_p(M_t) = \emptyset$. To find the spectrum we will need another new idea.

Definition 3.29. An operator $T \in B(\mathcal{H}, \mathcal{K})$ is **bounded below** if there exist $C > 0$ such that $\|Th\| \geq C\|h\|$ for all $h \in \mathcal{H}$.

Proposition 3.30. *If $T \in B(\mathcal{H})$ is invertible then T is bounded below. In fact, $\|Th\| \geq \|T^{-1}\|^{-1}\|h\|$.*

Proof. This follows from, $\|h\| = \|T^{-1}(Th)\| \leq \|T^{-1}\|\|Th\|$. □

Definition 3.31. Let $T \in B(\mathcal{H})$ we say that $\lambda \in \mathbb{C}$ is an **approximate eigenvalue** of T provided that there exists a sequence $\|h_n\| = 1$ such that $\lim_n \|(Th_n - \lambda h_n\| = 0$. We let $\sigma_{ap}(T)$ denote the set of approximate eigenvalues of T.

Note that if $(T - \lambda I)h_n \to 0$ then $(T - \lambda I)$ is not bounded below and so not invertible. Thus,

$$\sigma_{ap}(T) \subseteq \sigma(T).$$

We now show that $[0, 1] \subseteq \sigma_{ap}(M_t)$. To this end let $0 < \lambda < 1$, and for n sufficiently large, set

$$f_n(t) = \begin{cases} 0, & 0 \leq t < \lambda - \frac{1}{n} \\ \sqrt{n/2}, & \lambda - \frac{1}{n} \leq t \leq \lambda + \frac{1}{n} \\ 0, & \lambda + \frac{1}{n} < t \leq 1 \end{cases}.$$

It is easily checked that $f_n \in L^2$ and that $\|f_n\| = 1$. Also,

$$\|(M_t f_n - \lambda f_n\|^2 = \int_{\lambda - \frac{1}{n}}^{\lambda + \frac{1}{n}} (t - \lambda)^2 \frac{n}{2} dt = \frac{1}{3n^2} \to 0.$$

Thus, $\lambda \in \sigma_{ap}(M_t)$. The cases of $\lambda = 0$ and $\lambda = 1$ are similar.

Finally, if $\lambda \notin [0, 1]$, then the function $(t - \lambda)^{-1}$ is bounded on $[0, 1]$. Hence the operator $M_{(t-\lambda)^{-1}}$ of multiplication by this function defines a bounded operator and it is easily seen that it is the inverse of the operator $M_t - \lambda I$. Hence, $\lambda \notin \sigma(M_t)$ and we have that

$$\sigma(M_t) = [0, 1].$$

3.5 Numerical Ranges

DK: classical and higher rank numerical ranges; can introduce Pauli operators; maybe some baby QEC here including stabilizer formalism basics.

3.6 Majorization and Quantum State Convertibility

DK: spectral majorization, LOCC and state convertibility, Nielsen's theorem.

3.7 Spectral Mapping Theorems and Functional Calculi

In the study of eigenvalues of matrices a special role is played by being able to take polynomials of the matrix. Similarly, if $T \in B(\mathcal{H})$, and $p(z) = a_0 + a_1 z + \cdots + a_n z^n$ then we set $p(T) := a_0 I + a_1 T + \cdots + a_n T^n$.

Note that if we let $\mathcal{P}$ denote the algebra of polynomials, then the mapping $\pi_T : \mathcal{P} \to B(\mathcal{H})$ defined by $\pi_T(p) = p(T)$ satisfies:

- $\pi_T(p+q) = \pi_T(p) = \pi_T(q)$,
- $\pi_T(\lambda p) = \lambda \pi_T(p)$,
- $\pi_T(pq) = \pi_T(p)\pi_T(q)$.

That is it is linear and multiplicative. Such a map between two algebras is called a **homomorphism**. This mapping is sometimes called the **polynomial functional calculus**.

We now wish to see how spectrum behaves under such maps.

Lemma 3.32. *Let $A, B, T \in B(\mathcal{H})$ with $AT = I$ and $TB = I$, then $A = B$ and hence T is invertible.*

Theorem 3.33 (Spectral Mapping Theorem for Polynomials)**.** *Let $T \in B(\mathcal{H})$ and let p be a polynomial. Then*

$$\sigma(p(T)) = \{p(\lambda) : \lambda \in \sigma(T)\}.$$

Proof. For a polynomial p we always assume that $a_n \neq 0$. Given $\lambda \in \sigma(T)$ set $q(z) = p(z) - p(\lambda)$. Since $q(\lambda) = 0$ this means that $(z - \lambda)$ divides q and we may write $q(z) = (z - \lambda)q_1(z)$ for some polynomial $q_1(z)$. Hence, $p(T) - p(\lambda)I = (T - \lambda I)q_1(T) = q_1(T)(T - \lambda I)$. Suppose that $p(T) - p(\lambda)I$ is invertible. Then if B is the inverse then,

$$I = \big(Bq_1(T)\big)(T - \lambda I) = (T - \lambda I)\big(q_1(T)B\big),$$

and so $T-\lambda I$ is invertible by the lemma. Hence, $p(T)-p(\lambda)I$ is not invertible and we have that $\{p(\lambda) : \lambda \in \sigma(T)\} \subseteq \sigma(p(T))$.

On the other hand, if $\mu \in \sigma(p(T))$,(we assume that $a_n \neq 0$) factor

$$p(z) - \mu = a_n(z - \mu_1) \cdots (z - \mu_n).$$

If $(T - \mu_i I)$ is invertible for all i, then $p(T) - \mu$ is a product of invertible operators and so is invertible. Thus, there exists an i so that $\mu_i \in \sigma(T)$. But $p(\mu_i) - \mu = 0$ so that $\mu = p(\lambda)$ for some $\lambda \in \sigma(T)$. □

Definition 3.34. Let $T \in B(\mathcal{H})$, then the **spectral radius** of T is the number

$$r(T) := \sup\{|\lambda| : \lambda \in \sigma(T)\}.$$

Theorem 3.35 (Spectral Radius Theorem)**.** *Let $T \in B(\mathcal{H})$. Then*

- $r(T) \leq \|T^n\|^{1/n}$, $\forall n$,
- $\lim_n \|T^n\|^{1/n}$ *converges,*
- $r(T) = \lim_n \|T^n\|^{1/n}$.

Proof. We sketch the key ideas. First we know that $r(X) \leq \|X\|$ for any X. But by the spectral mapping theorem, $r(T^n) = r(T)^n$, so $r(T) = r(T^n)^{1/n} \leq \|T^n\|^{1/n}$.

From this it follows that $r(T) \leq \liminf_n \|T^n\|^{1/n}$. The proof is completed by showing that $\limsup_n \|T^n\|^{1/n} \leq r(T)$. This is done by setting $w = z^{-1}$ and observing that $zI - T$ is invertible iff $I - wT$ is invertible.

But $(I - wT)^{-1} = \sum_n (wT)^n$ for w small enough(hence z large enough). But we know that this power series is analytic for $|z| > r(T)$ and hence for $|w| < r(T)^{-1}$and w cannot be any larger. This guarantees that the radius of this power series is $R = r(T)^{-1}$. But even for power series with into Banach spaces we have that, $R^{-1} = \limsup \|T^n\|^{1/n}$. We have that $r(T) = \limsup_n \|T^n\|^{1/n}$. □

3.8 The Riesz Functional Calculus

In this section we will assume that the reader is familiar with complex analysis.

Let $K \subseteq \mathbb{C}$ be a compact set. We let $Hol(K)$ denote the set of functions that are analytic on some open set containing K. Sums and products of such functions are again in this set. Recall from complex analysis that given a simple closed curve or set of closed curves, $\Gamma = \{\Gamma_1, ..., \Gamma_N\}$ then each point that is not on the curves has a **winding number**. For a family of curves the winding number of a point is the sum of its winding numbers. Let $\gamma_j : [a_j, b_j] \to \mathbb{C}$ be a smooth parametrization of each curve Γ_j. If $g(z)$ is continuous on Γ then we set

$$\int_\Gamma g(z)dz = \sum_{j=1}^n \int_{a_j}^{b_j} g(\gamma_j(t))\gamma_j'(t)dt.$$

Note that since everything is continuous these integrals are defined as limits of Riemann sums.

Cauchy's integral formula says that if $g \in Hol(K)$ and Γ has the property that each point in K has winding number 1, then

$$g(w) = \frac{1}{2\pi i}\int_\Gamma g(z)(z-w)^{-1}dz.$$

The **Riesz functional calculus** is defined as follows. Given $T \in B(\mathcal{H})$, pick a system of curves Γ so that each point in $\sigma(T)$ has winding number 1. We define

$$f(T) := \frac{1}{2\pi i}\int_\Gamma f(z)(zI-T)^{-1}dz = \sum_{j=1}^n \frac{1}{2\pi i}\int_{a_j}^{b_j} f(\gamma(t))(\gamma(t)I-T)^{-1}\gamma_j'(t)dt.$$

Again since everything is continuous in the norm topology, these integrals can be defined as limits of Riemann sums, alternatively, for each pair of vectors h, k we can set

$$\langle h|f(T)k\rangle := \sum_{j=1}^n \frac{1}{2\pi i}\int_{a_j}^{b_j} f(\gamma(t))\langle h|(\gamma(t)I-T)^{-1}k\rangle\gamma_j'(t)dt,$$

and check that this defines a bounded operator.

The Riesz functional calculus refers to a group of results that explain how these integral formulas behave. Here is a brief summary of the key results.

Theorem 3.36 (Riesz Functional Calculus)**.** *Let $T \in B(\mathcal{H})$, let $f \in Hol(\sigma(T))$ and let Γ and $\widetilde{\Gamma}$ be two systems of curves that have winding number one about $\sigma(T)$.*

- *$\int_\Gamma f(z)(zI-T)^{-1}dz = \int_{\widetilde{\Gamma}} f(z)(zI-T)^{-1}dz$, so that the definition of $f(T)$ is independent of the particular curves used.*

- *if* $f(z) = \sum_{n=0}^{\infty} a_n z^n$ *is a power series with radius of convergence* $R, r(T) < R$, *then*
$$f(T) = \sum_{n=0}^{\infty} a_n T^n.$$
- *the map* $\pi_T : Hol(\sigma(T)) \to B(\mathcal{H})$ *given by* $\pi_T(f) = f(T)$ *is a homomorphism.*
- $\sigma(f(T)) = \{f(\lambda) : \lambda \in \sigma(T)\}$.

3.8.1 The Spectrum of a Hermitian

It is easy to see that any eigenvalue of a Hermitian matrix must be real, but Hermitian operators need not have any eigenvalues. We illustrate how to use the functional calculus to prove that the spectrum of a Hermitian operator is contained in the reals.

Proposition 3.37. *If* T *is invertible, then* $\sigma(T^{-1}) = \{\lambda^{-1}; \lambda \in \sigma(T)\}$.

Proof. First, since T^{-1} is invertible, $0 \notin \sigma(T)$. Now let $\mu \neq 0$, then

$$(\mu I - T^{-1}) = (\mu I)(T - \mu^{-1} I)T^{-1}.$$

Since the first and third terms are invertible, the left hand side is invertible iff $(T - \mu^{-1} I)$ is invertible and the result follows. □

Proposition 3.38. *Let* $U \in B(\mathcal{H})$ *be unitary. Then* $\sigma(U) \subseteq \{\lambda : |\lambda| = 1\}$.

Proof. If $\lambda \in \sigma(U)$ then $|\lambda| \leq \|U\| = 1$. Since $U^{-1} = U^*$, $\|U^{-1}\| \leq !$. Hence, $\lambda \in \sigma(U) \implies \lambda^{-1} \in \sigma(U^{-1}) \implies |\lambda^{-1}| \leq \|U^{-1}\| = 1$. Thus, $|\lambda| \geq 1$ and the result follows. □

We will often use $\mathbb{T} := \{\lambda : |\lambda| = 1\}$ to denote the unit circle in the complex plane.

Proposition 3.39. *Let* $H \in B(\mathcal{H})$, $H = H^*$. *Then* $\sigma(H) \subseteq \mathbb{R}$.

Proof. We have that $e^{iz} = \sum_{n=0}^{\infty} \frac{i^n z^n}{n!}$ has infinite radius of convergence. Set $U = e^{iH}$, then $U^* = \sum_{n=0}^{\infty} \frac{(-i)^n H^{*n}}{n!} = e^{-iH}$.

Hence, $UU^* = U^*U = I$ so U is unitary. Hence,

$$\{e^{iz} : z \in \sigma(H)\} = \sigma(U) \subseteq \mathbb{T}.$$

But it is easily checked that $e^{iz} \in \mathbb{T}$ iff $z \in \mathbb{R}$. □

3.9 Normal Operators

Definition 3.40. An operator $N \in B(\mathcal{H})$ is **normal** provided that $NN^* = N^*N$.

We see that Hermitian and unitary operators are normal.

Proposition 3.41. *Let $N \in B(\mathcal{H})$ and let $N = H + iK$ be its Cartesian decomposition. Then N is normal iff $HK = KH$.*

Given a polynomial, $p(z, \overline{z})) = \sum_{i,j} a_{i,j} z^i \overline{z}^j$ in z and $\overline{z}$ we can evaluate it for any normal operator, by setting $p(N, N^*) = \sum_{i,j} a_{i,j} N^i N^{*j}$. If we let $\mathcal{P}(z, \overline{z})$ denote the set of such polynomials, then it is again an algebra and we see that when N is normal, the map $\pi_N : \mathcal{P}(z, \overline{z}) \to B(\mathcal{H})$ is a homomorphism. In fact, if it is a homomorphism, then necessarily N is normal, since in the polynomials, $z\overline{z} = \overline{z}z$.

Proposition 3.42 (Gelfand-Naimark)**.** *Let N be normal and let $p(z, \overline{z})$ be a polynomial. Then*

$$\|p(N, N^*)\| = \{|p(\lambda, \overline{\lambda})| : \lambda \in \sigma(N)\}.$$

The right hand side of the above formula is the supremum norm $\| \cdot \|_\infty$ of the polynomial over $\sigma(N)$. By the Stone-Weierstrass theorem, the polynomials in z and $\overline{z}$ are dense in the continuous functions on $\sigma(N)$, Thus, given any continuous function $f \in C(\sigma(N))$ there is a Cauchy sequence of polynomials $\{p_n(z, \overline{z})\}$ that converge to it in the supremum norm. By the above result, $\|p_n(N, N^*) - p_m(N, N^*)\| = \|p_n - p_m\|_\infty$ and so this sequence of operators is also Cauchy in norm and hence converges to an operator that we shall denote by $f(N)$. It is not hard to check that the operator $f(N)$ is independent of the particular Cauchy sequence chosen.

This gives us a **continuous functional calculus for normal operators**.

Theorem 3.43 (Gelfand-Naimark)**.** *Let $N \in B(\mathcal{H})$ be normal and let $\pi_N : C(\sigma(N)) \to B(\mathcal{H})$ be defined by $\pi_N(f) = f(N)$. Then*

- *π_N is a homomorphism,*
- *$\pi_N(\overline{f}) = f(N)^*$ (maps satisfying these two properties are called* ***-homomorphisms***),*
- *$\|f(N)\| = \sup\{|f(\lambda)| : \lambda \in \sigma(N)\}$,*
- *$\sigma(f(N)) = \{f(\lambda) : \lambda \in \sigma(N)\}$.*

3.9.1 Positive Operators

Definition 3.44. An operator $P \in B(\mathcal{H})$ is called **positive**, denoted $P \geq 0$ provided that $\langle h|Ph\rangle \geq 0$ for every $h \in \mathcal{H}$.

CAUTION: In the theory of matrices these are often called **positive semidefinite** but the same notation $P \geq 0$ is used and positive is reserved for matrices such that $\langle h|Ph\rangle > 0$ for all $h \neq 0$, for which they use the notation $P > 0$. So the terminology is different for the two fields.

By one of our earlier results, $P \geq 0 \implies P = P^*$, since $\langle h|Ph\rangle \in \mathbb{R}$. So positive operators are self-adjoint.

Note that if $T \in B(\mathcal{H}, \mathcal{K})$ then $T^*T \geq 0$.

Problem 3.45. Let $H = H^*$. Prove that $\sigma_{ap}(H) = \sigma(H)$. (HINT: For $\lambda \in \mathbb{R}$, $\mathcal{N}(H - \lambda I) = \mathcal{R}(H - \lambda I)^\perp$, so that if λ is not an eigenvalue, then $\mathcal{R}(H - \lambda I)$ is dense. Show that if $H - \lambda I$ is bounded below and has dense range, then it is invertible.)

Problem 3.46. Prove that the following are equivalent:

1. $P \geq 0$,
2. $P = P^*$ and $\sigma(P) \subseteq [0, +\infty)$,
3. there exists $B \in B(\mathcal{H})$ such that $P = B^*B$.

By the above results, since P is normal and $\sigma(P) \subseteq [0, +\infty)$ the function $\sqrt{t}$ is continuous on $\sigma(P)$ and so by the continuous functional calculus, there is an operator $\sqrt{P}$ defined.

Definition 3.47. Let $T \in B(\mathcal{H}, \mathcal{K})$ we define the **absolute value of T** to be the operator,

$$|T| := \sqrt{T^*T}.$$

CAUTION: Unlike numbers, $|T|$ is not always equal to $|T^*|$. In fact, if $T \in B(\mathcal{H}, \mathcal{K})$, then $|T| \in B(\mathcal{H})$ while $|T^*| \in B(\mathcal{K})$.

Here is another application of the continuous functional calculus. Let $H = H^*$. Define

$$f^+(t) = \begin{cases} 0, & t \leq 0 \\ t, & t > 0 \end{cases} \text{ and } f^-(t) = \begin{cases} -t, & t \leq 0 \\ 0, & t > 0 \end{cases}.$$

These are both continuous functions and they satisfy

- $f^+(t) + f^-(t) = |t| = \sqrt{t^2}$,

- $f^+(t) - f^-(t) = t$,
- $f^+(t)f^-(t) = 0$.

By the functional calculus, we may set $H^+ = f^+(H)$ and $H^- = f^-(H)$. Then we have that

- $H^+ + H^- = |H| = \sqrt{H^2}$,
- $H^+ - H^- = H$,
- $H^+H^- = H^-H^+ = 0$.

Note that if H was just a matrix in diagonal form, then H^+ would be the matrix gotten by leaving the positive eigenvalues alone and setting the negative eigenvalues equal to 0, with H^- similarly defined. These two matrices would satisfy the above properties. Thus, the functional calculus allows us to abstractly carry out such operations even when we have no eigenvalues.

3.9.2 Polar Decomposition

Every complex number z can be written as $z = e^(it)|z|$ for some unique "rotation". The following is the analogue for operators.

Theorem 3.48 (Polar Decomposition)**.** *Let $T \in B(\mathcal{H}, \mathcal{K})$. Then there exists a partial isometry W such that $T = W|T|$. Moreover, W is unique if we require that*

- *W^*W is the projection onto $\mathcal{N}(T)^\perp$,*
- *WW^* is the projection onto $\mathcal{R}(T)^-$.*

Because $\mathcal{R}(|T|)^- = \mathcal{N}(|T|)^\perp = \mathcal{N}(T)^\perp$ we have that $W^*T = W^*W|T| = |T|$.

3.10 Density Operators and Mixed States

DK: spectral decompositions and how can relate (special case of cp map case), partial trace operation, purification of mixed states (special case of Stinespring later) with eg's.

3.11 Compact Operators

Theorem 3.49. *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces, let $\mathbb{B}_1 = \{h \in \mathcal{H} : \|h\| \le 1\}$ and let $T \in B(\mathcal{H}, \mathcal{K})$. The following are equivalent:*

- *T is a norm limit of finite rank operators,*
- *$T(\mathbb{B}_1)^-$ is compact,*
- *$T(\mathbb{B}_1)$ is compact.*

Such operators are called **compact** and the set of all compact operators in $B(\mathcal{H}, \mathcal{K})$ is denoted $\mathbb{K}(\mathcal{H}, \mathcal{K})$. When $\mathcal{H} = \mathcal{K}$ we write more simply, $\mathbb{K}(\mathcal{H})$.

Theorem 3.50. *$\mathbb{K}(\mathcal{H}, \mathcal{K})$ is a norm closed subspace of $B(\mathcal{H}, \mathcal{K})$. If $A \in B(\mathcal{H})$ and $B \in B(\mathcal{K})$ and $K \in \mathbb{K}(\mathcal{H}, \mathcal{K})$, then $BKA \in \mathbb{K}(\mathcal{H}, \mathcal{K})$.*

Theorem 3.51. *Let $H \in \mathbb{K}(\mathcal{H})$ with $H = H^*$. Then*

1. *every $0 \neq \lambda \in \sigma(H)$ is an eigenvalue,*
2. *H has at most countably many non-zero eigenvalues, and each eigenvalue has a finite dimensional eigenspace,*
3. *there exists a countable orthonormal set of eigenvectors $\{\phi_n : n \in A\}$ with corresponding eigenvalues $\{\lambda_n : n \in A\}$ such that*

$$H = \sum_{n \in A} \lambda_n |\phi_n\rangle\langle\phi_n|,$$

Moreover, if $H = \sum_j \beta_j |\gamma_j\rangle\langle\gamma_j|$ is another such representation with orthonormal vectors, then $\{\beta_j\} = \{\lambda_n\}$ including multplicities.

Proof. First if $0 \neq \lambda \in \sigma(H)$, then we know that there exists a sequence of vectors $\{h_n\}$ such that $Hh_n - \lambda h_n\| \to 0$. Since H is compact, there is a subsequence $\{h_{n_k}\}$ such that $\{Hh_{n_k}\}$ converges. Set $v = \lim_k Hh_{n_k} = \lim_k \lambda h_{n_k}$, since $\|Hh_{n_k} - \lambda h_{n_k}\| \to 0$. But this implies that $\lim_k h_{n_k} = \lambda^{-1} v := w$. Since each h_{n_k} is a unit vector, $\|w\| =!$ and in particular $w \neq 0$. Now $Hw = \lim_k Hh_{n_k} = v = \lambda w$, so we have that λ is an eigenvalue.

Now suppose that for some $0 \neq \lambda \in \sigma(H)$ the eigenspace was infinite dimensional. Then by choosing an orthonormal basis for the eigenspace, we could find countably many orthonormal vectors $\{h_n\}$ with $Hh_n = \lambda h_n$ Since for $n \neq m$ we have that $\|\lambda h_n - \lambda h_m\| = |\lambda|\sqrt{2}$, we see that no subsequence could converge, contradicting (2). Thus, each eigenspace is at most finite dimensional.

We prove that the set of non-zero eigenvalues is countable by contradiction. Now suppose that there was an uncountable number of non-zero points in $\sigma(H)$. For each n, consider $S_n = \{\lambda \in \sigma(H) : |\lambda| \geq 1/n\}$. If this set was finite for all n then

$$\sigma(H)\backslash\{0\} = \cup_n S_n,$$

would be a countable union of finite sets and hence countable. Thus, the assumption implies that for some n_0 the set S_{n_0} is infinite. This allows us to choose a countably infinite subset $\{\lambda_k : k \in \mathbb{N}\} \subset S_{n_0}$ of distinct points. For each k there is a unit eigenvector, v_k with $Hv_k = \lambda_k v_k$. But then for $k \neq j$,

$$\lambda_k \langle v_k | v_j \rangle = \langle H v_k | v_j \rangle = \langle v_k | H v_j \rangle = \langle v_k | v_j \rangle \lambda_j.$$

Since $\lambda_k \neq \lambda_j$ this equation is possible iff $\langle v_k | v_j \rangle = 0$. Thus, the eigenvectors are orthogonal. This implies that $\|Hv_k - Hv_j\| = \|\lambda_k v_k - \lambda_j v_j\| \geq \frac{\sqrt{2}}{n_0}$ which means that no subsequence of these vectors could converge. This contradicts that H is compact.

The third statement follows by choosing an orthonormal basis for each finite dimensional eigenspace and taking the union of all these vectors to for $\{\phi_n\}$.

Finally, the last statement follows from the fact that both sets must be the eigenvalues with the appropriate multiplicities. □

Definition 3.52. If $K \in \mathbb{K}(\mathcal{H}, \mathcal{K})$. Then $|K|$ is a positive, compact operator and hence by the above theorem it has countably many distinct non-zero eigenvalues, every non-zero eigenvalue is positive and each non-zero eigenvalue has at most a finite dimensional eigenspace. We let $\lambda_1 \geq \lambda_2 \geq \ldots$, denote these eigenvalues with the convention that each eigenvalue is repeated a finite number of times corresponding to the dimension of its eigenspace. The number

$$s_j(K) := \lambda_j,$$

is called the **j-th singular value of K**.

Theorem 3.53 (Schmidt)**.** *Let $K \in \mathbb{K}(\mathcal{H}, \mathcal{K})$. Then there exist countable collections of orthonormal vectors $\{\phi_n\} \subseteq \mathcal{H}$ and $\{\psi_n\} \subseteq \mathcal{K}$, such that*

$$K(h) = \sum_n s_n(K) \langle \phi_n | h \rangle \psi_n,$$

i.e., $K = \sum_n s_n(K)|\psi_n\rangle\langle\phi_n|$. *Moreover, if* $K = \sum_j \beta_j\langle|\gamma_j\rangle\langle\delta_j|$ *is another representation with* $\beta_j > 0$ *and* $\{\gamma_j\}$ *and* $\{\delta_j\}$ *orthonormal families, then* $\{\beta_j\} = \{s_n(K)\}$*, including multiplicities.*

Proof. Use the polar decomposition to write $K = W|K|$ with W a partial isometry. Since $|K|$ is compact and positive with eigenvalues $s_j(K)$, we can write $|K| = \sum_j s_j(K)|\phi_j\rangle\langle\phi_j|$, with $\{\phi_j\}$ a countable orthonormal set. Since the vectors $\phi_j \in \mathcal{R}(|K|)$ and W acts as an isometry on that space the vectors $\psi_j = W\phi_j$ are also orthonormal, by the inner product preserving property of isometries.

For any h we have that $K(h) = W\big(\sum_j s_j(K)\langle\phi_j|h\rangle\phi_j\big) = \sum_j s_j(K)\langle\phi_j|h\rangle\psi_j$ and so $K = \sum_j s_j(K)|\phi_j\rangle\langle\psi_j|$.

Finally for the last statement, given such a representation, we have that $K^* = \sum_j \beta_j|\delta_j\rangle\langle\gamma_j|$. Hence, $K^*K = \sum_j \beta_j^2|\gamma_j\rangle\langle\gamma_j|$. Thus, $|K| = \sum_j \beta_j|\gamma_j\rangle\langle\gamma_j|$. From the earlier result we conclude that, $\{\beta_j\} = \{s_n(K)\}$, including multiplicities. □

3.11.1 The Schatten Classes

Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. For $1 < p < +\infty$ the **Schatten p-class** is defined to be

$$\mathcal{C}_p(\mathcal{H},\mathcal{K}) = \{K \in \mathbb{K}(\mathcal{H},\mathcal{K}) : \sum_j s_j(K)^p < +\infty\}.$$

Note also that since $\{s_j(K)\}$ are the eigenvalues of $|K|$ including multiplicities, that if we use a basis for $\mathcal{H}$ that includes the eigenvectors for $|K|$, then we see that

$$Tr(|K|^p) = \sum_j s_j(K)^p,$$

so that $\mathcal{C}_p$ is exactly the set of operators for which $|K|^p$ is in the trace-class.

For $K \in \mathcal{C}_p$ we set $\|K\|_p = \big(\sum_j s_j(K)^p\big)^{1/p}$ and this is called the **p-norm** of the operator.

The space $\mathcal{C}_1$ is called the **trace class operators** and the space $\mathcal{C}_2$ is called the **Hilbert-Schmidt** operators.

Theorem 3.54. *We have that:*

1. $\mathcal{C}_p(\mathcal{H},\mathcal{K})$ *is a Banach space in the p-norm,*
2. *for* $A \in B(\mathcal{H}), B \in B(\mathcal{K})$ *and* $K \in \mathcal{C}_p(\mathcal{H},\mathcal{K})$ *we have that* $BKA \in \mathcal{C}_p(\mathcal{H},\mathcal{K})$ *and* $\|BKA\|_p \le \|B\|\|K\|_p\|A\|$.

Theorem 3.55. *Let $K \in \mathcal{C}_1(\mathcal{H})$ and let $\{e_a : a \in A\}$ be any orthonormal basis for $\mathcal{H}$. Then $\sum_a \langle e_a|Ke_a\rangle$ converges and its value is independent of the particular orthonormal basis.*

Definition 3.56. Let $K \in \mathcal{C}_1(\mathcal{H})$, then we set

$$Tr(K) = \sum_a \langle e_a|Ke_a\rangle,$$

where $\{e_a : a \in A\}$ is any orthonormal basis.

Theorem 3.57. *Let $\mathcal{H}$ be a Hilbert space.*

1. *for $A \in B(\mathcal{H})$ and $K \in \mathcal{C}_1(\mathcal{H})$, $Tr(AK) = Tr(KA)$.*
2. *For each $A \in B(\mathcal{H})$ define*

   $$f_A : \mathcal{C}_1(\mathcal{H}) \to \mathbb{C},$$

   by $f_A(K) = Tr(AK)$, then f_A is a bounded linear functional and $\|f_A\| = \|A\|$.
3. *If $f : \mathcal{C}_1(\mathcal{H}) \to \mathbb{C}$ is any bounded, linear functional, then there exists a unique $A \in B(\mathcal{H})$ such that $f = f_A$. Consequently, the map $A \to f_A$ is an isometric isomorphism of $B(\mathcal{H})$ onto the Banach space dual $\mathcal{C}_1(\mathcal{H})^*$.*
4. *If $K \in \mathcal{C}_1(\mathcal{H})$ then the map*

   $$f_K : \mathbb{K}(\mathcal{H}) \to \mathbb{C},$$

   defined by $f_K(T) = Tr(KT)$ is a bounded linear functional and $\|f_K\| = \|K\|_1$.
5. *If $f : \mathbb{K}(\mathcal{H}) \to \mathbb{C}$ is any bounded linear functional, then there exists a unique $K \in \mathcal{C}_1(\mathcal{H})$ such that $f = f_K$. Consequently, the map $K \to f_K$ is an isometric isomorphism of $\mathcal{C}_1(\mathcal{H})$ onto the Banach space dual $\mathbb{K}(\mathcal{H})^*$.*

Theorem 3.58. *Let $1 < p < +\infty$ and let $\frac{1}{p} + \frac{1}{q} = 1$. Then*

1. *for $T \in \mathcal{C}_q(\mathcal{H})$ and $R \in \mathcal{C}_p(\mathcal{H})$, $TR \in \mathcal{C}_1(\mathcal{H})$, $Tr(RT) = Tr(TR)$ and $|Tr(TR)| \leq \|R\|_p \|T\|_q$.*
2. *the map $f_T : \mathcal{C}_p(\mathcal{H}) \to \mathbb{C}$ given by $f_T(R) = Tr(TR)$ is a bounded linear functional with $\|f_T\| = \|T\|_q$.*
3. *if $f : \mathcal{C}_p(\mathcal{H}) \to \mathbb{C}$ is any bounded linear functional, then there exists a unique $T \in \mathcal{C}_q(\mathcal{H})$ such that $f = f_T$.*

Consequently, the map $T \to f_T$ is an isometric isomorphism of $\mathcal{C}_q(\mathcal{H})$ onto the Banach space dual $\mathcal{C}_p(\mathcal{H})^*$.

3.11.2 Hilbert-Schmidt and Tensor Products

Given orthonormal bases $\{e_a : a \in A\}$ for $\mathcal{H}$ and $\{f_b : b \in B\}$ for $\mathcal{K}$ we have a matrix representation for any $T \in B(\mathcal{H}, \mathcal{K})$, namely,

$$T_{mat} = (t_{b,a}), \text{ with } t_{b,a} = \langle f_b | T e_a \rangle.$$

Proposition 3.59. *$T \in \mathcal{C}_2(\mathcal{H}, \mathcal{K})$ if and only if $\sum_{a\in A, b\in B} |t_{b,a}|^2 < +\infty$. Moreover,*

$$\|T\|_2^2 = \sum_{a\in A, b\in B} |t_{b,a}|^2.$$

Proof. $T \in \mathcal{C}_2$ if and only if $\|T\|_2^2 = Tr(T^*T) < +\infty$. But

$$Tr(T^*T) = \sum_{a\in A} \langle e_a | T^*T e_a \rangle = \sum_{a\in A} \|Te_a\|^2 = \sum_{a\in A}\sum_{b\in B} |\langle f_b, Te_a\rangle|^2 = \sum_{a\in A, b\in B} |t_{b,a}|^2,$$

and the result follows. □

Let $\phi \in \mathcal{H}\overline{\otimes}\mathcal{K}$ and let $\{e_a : a \in A\}$ and $\{f_b : b \in B\}$ be orthonormal bases for $\mathcal{H}$ and $\mathcal{K}$, respectively. Expanding ϕ with respect to this basis we have that $\phi = \sum_{a,b} t_{b,a} e_a \otimes f_b$ with $\|\phi\|^2 = \sum_{a,b} |t_{b,a}|^2$. Hence, $(t_{b,a})$ is the matrix of a Hilbert-Schmidt operator T from $\mathcal{H}$ to $\mathcal{K}$, with $\|T\|_2 = \|\phi\|$. Thus, if we fix bases, then we see that there is an isometric linear map, $\Gamma : \mathcal{H}\overline{\otimes}\mathcal{K} \to \mathcal{C}_2(\mathcal{H}, \mathcal{K})$, that is onto.

In the case of Hilbert spaces over $\mathbb{R}$ it is possible to define this map in a basis free manner. If $\phi = \sum_i h_i \otimes k_i$ is a finite sum, then define an operator T_ϕ by

$$T_\phi(h) = \sum_i \langle h_i | h \rangle k_i,$$

i.e., $T_\phi = \sum_i |k_i\rangle\langle h_i|$. It is not hard to check that T_ϕ is Hilbert-Schmidt and that the map, $\Delta : \mathcal{H} \otimes \mathcal{K} \to \mathcal{C}_2(\mathcal{H}, \mathcal{K})$ given by $\Delta(\phi) = T_\phi$ is linear and isometric. Moreover, if we pick bases for $\mathcal{H}$ and $\mathcal{K}$ as above, then the matrix of the linear map T_ϕ is $\Gamma(\phi)$. Thus, in the real case the map Γ is the continuous extension of the map Δ after picking bases.

Unfortunately, in the complex case, we can no longer define the map Δ map, since the identification of the vector h_i with the functional $|h_i\rangle$ is conjugate linear, and so we are forced to pick bases and deal with the map Γ, which in this case really is a basis dependent map.

Here is one application of these ideas.

Proposition 3.60 (Schmidt Decomposition of Vectors)**.** *Let $\phi \in \mathcal{H}\overline{\otimes}\mathcal{K}$. Then there are countable orthonormal families of vectors $\{u_n\} \subseteq \mathcal{H}$ and $\{v_n\} \subseteq \mathcal{K}$ and constants $\beta_n > 0$ such that*

$$\phi = \sum_n \beta_n u_n \otimes v_n.$$

Moreover, if $\phi = \sum_m \alpha_m w_m \otimes y_m$ is another such representation, then $\{\alpha_m\} = \{\beta_n\}$, including multiplicities. Moreover, $\|\phi\|^2 = \sum_n \beta_n^2$.

Proof. We first do the case of real Hilbert spaces. Consider the Hilbert-Schmidt operator $T_\phi : \mathcal{H} \to \mathcal{K}$. By the above results we may write

$$T_\phi = \sum_n s_n(T_\phi)|v_n\rangle\langle u_n|,$$

where $\{u_n\}$ and $\{v_n\}$ are orthonormal. Consequently, $\phi = \sum_n s_n(T_\phi)u_n \otimes v_n$, as desired.

If we have a representation of $\phi = \sum_m \alpha_m w_m \otimes y_m$, then $T_\phi = \sum_m \alpha_m |y_m\rangle\langle w_m|$ and by our earlier results this forces, $\{\alpha_m\} = \{s_n(T_\phi)\}$ along with multiplicities.

For the complex case, we fix a basis to define the matrix and Hilbert-Schmidt operator $T = (t_{b,a})$. This yields the representation of the vector with coefficients $s_n(T)$. Now one shows that if one chooses a different basis, then the operator defined by the new matrix, say R is unitarily to T and so $s_n(R) = s_n(T)$ for all n. Finally, given any representation, one extends the sets $\{w_m\}$ and $\{y_m\}$ to orthonormal bases, and observes that then $\{\alpha_m\}$ is then the singular values of the matrix obtained from ϕ for this basis. □

3.12 Unbounded Operators

Many naturally defined maps in mathematical physics, such as position and momentum operators, are often unbounded. So are many natural differential operators. Unbounded operators and self-adjointness questions play an important role in these areas. We provide a brief glimpse into this theory.

Definition 3.61. Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces. A map $B : \mathcal{K} \times \mathcal{H} \to \mathbb{C}$ such that

- $B(k, h_1 + \lambda h_2) = B(k, h_1) + \lambda B(k, h_2)$,
- $B(k_1 + \lambda k_2, h) = B(k_1, h) + \overline{\lambda} B(k_2, h)$

is called a **sesquilinear form**. A sesquilinear form is called **bounded** if there is a constant $C \geq 0$ such that

$$|B(k,h)| \leq C\|k\|\|h\|, \forall h \in \mathcal{H}, k \in \mathcal{K}.$$

In this case, the least constant C satisfying this inequality is called the **norm** of the sesquilinear form and is denoted $\|B\|$.

Proposition 3.62. *If $T \in B(\mathcal{H},\mathcal{K})$, then $B(k,h) = \langle k|Th\rangle$ is a bounded sesquilinear form with $\|B\| = \|T\|$. Conversely, if B is a bounded sesquilinear form, then there exists a unique $T \in B(\mathcal{H},\mathcal{K})$ such that $B(k,h) = \langle k|Th\rangle$.*

Problem 3.63. Prove the above proposition.

Theorem 3.64 (Hellinger-Toeplitz)**.** *Let $\mathcal{H}$ and $\mathcal{K}$ be Hilbert spaces and let $T : \mathcal{H} \to \mathcal{K}$ be a linear map. If there exists a linear map $R : \mathcal{K} \to \mathcal{H}$ such that*

$$\langle Rk|h\rangle = \langle k|Th\rangle, \forall h \in \mathcal{H},\ k \in \mathcal{K},$$

then T is bounded.

Problem 3.65. Prove the Hellinger-Toeplitz theorem.

By the above theorem, if one wants to study notions of adjoints in the unbounded setting, then necessarily the domain of the map will not be all of the Hilbert space. We will think of a linear map as a pair consisting of the domain of the map and the map itself. To this end let $\mathcal{H},\mathcal{K}$ be Hilbert spaces, let $\mathcal{D}(T) \subseteq \mathcal{H}$ be a vector subspace, not necessarily closed, and let $T : \mathcal{D}(T) \to \mathcal{K}$ be a linear map. By the **graph of T** we mean the set

$$\Gamma(T) = \{(h,Th) \in \mathcal{H} \oplus \mathcal{K} : h \in \mathcal{D}(T)\}.$$

It is not hard to see that $\Gamma(T)$ is a subspace.

We say that $T_1 : \mathcal{D}(T_1) \to \mathcal{K}$ **extends T** provided that $\mathcal{D}(T) \subseteq \mathcal{D}(T_1) \subseteq \mathcal{H}$ and for every $h \in \mathcal{D}(T)$ we have that $T_1h = Th$. In this case we write $T \subseteq T_1$.

It is not hard to see that $T \subseteq T_1 \iff \Gamma(T) \subseteq \Gamma(T_1)$. The following problem clarifies some of these connections.

Problem 3.66. Let $\mathcal{H},\mathcal{K}$ be Hilbert spaces and let $\Gamma \subseteq \mathcal{H} \oplus \mathcal{K}$ be a vector subspace. Prove that Γ is the graph of a linear map if and only if $(0,k) \in \Gamma \implies k = 0$.

Such a subspace is called a **graph.**

Definition 3.67. We say that $T : \mathcal{D}(T) \to \mathcal{K}$ is **closed** provided that $\Gamma(T)$ is a closed subspace of $\mathcal{H} \oplus \mathcal{K}$. We say that T is **closalbe** if there exists T_1 with $T \subseteq T_1$ such that T_1 is closed. If T is closable, the smallest closed extension of T is called the **closure of T** and is denoted $\bar{T}$.

Proposition 3.68. *T is closable if and only if the closure of $\Gamma(T)$ is a graph. In this case, $\Gamma(T)^- = \Gamma(\bar{T})$.*

Thus, by the last problem we see that T is closable if and only if $(0, k) \in \Gamma(T)^-$ implies that $k = 0$.

For the following examples, we assume some measure theory. It is useful to note that if f, g are continuous on $[0, 1]$ and $f = g$ a.e., then $f = g$. Thus, we can think of $C([0, 1])$ as a subspace of $L^2([0, 1])$, where this latter space denotes the Hilbert space of equivalence classes of square-integrable functions with respect to Lebesgue measure. We let $\mathcal{C}^1([0, 1]) \subseteq C([0, 1])$ denote the vector space of continuously differentiable functions.

Example 3.69. Let $\mathcal{H} = \mathcal{K} = L^2([0, 1])$, let $\mathcal{D}(T_1) = \mathcal{C}^1([0, 1])$, let $\mathcal{D}(T_2) = \{f \in \mathcal{C}^1([0, 1]) : f(0) = f(1) = 0\}$ and define $T_1(f) = T_2(f) = f'$. Note that $f_\lambda(t) = e^{\lambda t} \in \mathcal{D}(T_1)$ but is never in $\mathcal{D}(T_2)$ and $T_1(f_\lambda) = \lambda f_\lambda$. Thus, every complex number is an eigenvalue of T_1. From calculus we know that this is the unique solution to $f' = \lambda f$, so that T_2 has no eigenvalues. Hence, we have that

$$\sigma_p(T_1) = \mathbb{C}, \;\; \sigma_p(T_2) = \emptyset.$$

Are these operators closed or closable? For closed, we need to decide that if $\{f_n\} \subseteq \mathcal{D}(T_i)$ such that $\|f_n - f\|_2 \to 0$ and $\|f_n' - g\|_2 \to 0$, then do we necessarily have that $f \in \mathcal{D}(T_i)$ with $g = f'$? The answer to this is no. For example, the functions,

$$f(t) = |t - 1/2|, \, g(t) = \begin{cases} 1, & 0 \le t < 1/2 \\ -1, & 1/2 < t \le 1 \end{cases}$$

can be approximated as above by just "smoothing f near 1/2" and f is in neither domain. For closable, we need to rule out $(0, g) \in \Gamma(T_i)^-$, i.e., if $\|f_n\|_2 \to 0$ and $\|f_n' - g\|_2 \to 0$ does this force $g = 0$ a.e. ? Trying one sees that this is hard to do with no tools. The next result gives a very useful tool.

Definition 3.70. We say that $T : \mathcal{D}(T) \to \mathcal{K}$ is **densely defined** if $\mathcal{D}(T) \subseteq \mathcal{H}$ is a dense subspace. When T is densely defined, we set

$$\mathcal{D}(T^*) = \{k \in \mathcal{K} : \exists h_1 \in \mathcal{H}, \langle k|Th\rangle = \langle h_1|h\rangle, \forall h \in \mathcal{D}(T)\}.$$

It follows readily, using the fact that $\mathcal{D}(T)$ is dense, that if such a vector h_1 exists, then it is unique and we define $T^* : \mathcal{D}(T^*) \to \mathcal{H}$ by $T^*(k) = h_1$. This operator(along with this domain) is called the **adjoint of T**.

Theorem 3.71. *1. T^* is closed.*

2. T is closable if and only if $\mathcal{D}(T^*)$ is a dense subset of $\mathcal{K}$. In this case $\overline{T} = T^{**}$.

3. If T is closable, then $(\overline{T})^* = T^*$.

Returning to the examples, we see that $g \in \mathcal{D}(T_i^*)$ if and only if there is a $h_1 \in L^2$ such that

$$\langle g|T_i(f)\rangle = \int_0^1 g(t)f'(t)dt = \int_0^1 h_1(t)f(t)dt = \langle h_1|f\rangle, \forall f \in \mathcal{D}(T_i).$$

This suggests integration by parts:

$$\int_0^1 gf'dt = -\int_0^1 g'(t)f(t)dt + (f(1)g(1) - f(0)g(0)).$$

Thus, if $g \in \mathcal{C}^1([0,1])$ and $g(1) = g(0) = 0$, then

$$\langle g|T_i(f)\rangle = \langle -g'|f\rangle,$$

so that in both cases every such g is in $\mathcal{D}(T_i)$. Because such g's are dense in L^2, we have that both operators are closable.

To understand the domains of either operator one needs more measure theory.

Definition 3.72. A function $f : [a,b] \to \mathbb{R}$ is **absolutely continuous(AC)** provided that for every $\epsilon > 0$ there exists $\delta > 0$ such that whenever we have finitely many $[a_i, b_i]$ that are non-intersecting subintervals with $\sum_i |b_i - a_i| < \delta$, then $\sum_i |f(b_i) - f(a_i)| < \epsilon$.

Not that if we only use one subinterval, then this is precisely uniform continuity.

Theorem 3.73. *Let $f : [a,b] \to \mathbb{R}$. Then f is AC if and only if*

1. *$f'(t)$ exists for almost all t,*
2. *f' is integrable,*
3. *for any $a \leq c < d \leq b$, $\int_c^d f'(t)dt = f(d) - f(c)$.*

For our two examples we have that

$$\mathcal{D}(T_1^*) = \{g : \text{g is AC, g(0) = g(1) =0}\},\ \mathcal{D}(T_2^*) = \{g :\ \text{g is AC }\},$$

and in both cases $T_1^*(g) = T_2^*(g) = -g'$.

We now turn to the definition of self-adjoint in the unbounded case.

Definition 3.74. A densely defined operator on $\mathcal{H}$ is **symmetric**(some times called **Hermitian**) if $T \subseteq T^*$. i.e., $\mathcal{D}(T) \subseteq \mathcal{D}(T^*)$ and $T^*h = Th, \forall h \in \mathcal{D}(T)$. A densely defined operator is called **self-adjoint** if $T = T^*$.

If T is self-adjoint, then $T = T^* = T^**= \bar{T}$.

Just as with bounded self-adjoint operators there is a nice functional calculus for unbounded self-adjoint operators.

Example 3.75. For the operator of differentiation we saw that a -1 was involved in the adjoint. For this reason we consider $\mathcal{H} = \mathcal{K} = L^2([0,1])$, let $\mathcal{D}(S_1) = \mathcal{C}^1([0,1])$, let $\mathcal{D}(S_2) = \{f \in \mathcal{C}^1([0,1]) : f(0) = f(1) = 0\}$ and define $S_1(f) = S_2(f) = if'$. Then we have that $S_1^*(g) = S_2^*(g) = -ig'$ when g is in $\mathcal{C}^1$ but a careful look at the boundary values shows that $S_2 \subseteq S_2^*$ while $\mathcal{D}(S_1^*)$ is not a subset of $\mathcal{D}(S_1)$. Thus, S_2 is symmetric while S_1 is not.

A natural question is what kinds of boundary values do give self-adjoint extensions?

Example 3.76. Let $\mathcal{H} = L^2([0,1])$, fix $|\alpha| = 1$ and let $\mathcal{D}(R_\alpha) = \{f \in AC([0,1]) : f(0) = \alpha f(1)\}$ and set $R_\alpha(f) = if'$. Now it is easily checked that $R_\alpha^* = R_\alpha$.

Notice that R_α is a one-dimensional extension of S_2.

Since self-adjoint operators must be closed a natural question is if we start with a symmetric operator, when will its closure be self-adjoint? The following results address this issue.

Definition 3.77. A symmetric operator is **essentially self-adjoint** if its closure is self-adjoint.

Theorem 3.78. *Let T be symmetric. The following are equivalent:*

1. *T is self-adjoint,*

2. *$T = \bar{T}$ and $\mathcal{N}(T^* \pm iI) = (0)$,*

3. *$\mathcal{R}(T \pm iI) = \mathcal{H}$.*

Theorem 3.79. *Let T be symmetric. The following are equivalent:*

1. *T is essentially self-adjoint,*

2. *$\mathcal{N}(T^* \pm iI) = (0)$,*

3. *$\mathcal{R}(T \pm iI) = \mathcal{H}$.*

Example 3.80. We apply these results to S_1 and S_2. We have that $S_1^*(g) = S_2^*(g) = -ig'$ for g in their domains and $\mathcal{D}(S_1^*) = \{g \in AC([0,1]) : g(0) = g(1) = 0\}$ while $\mathcal{D}(S_2^*) = AC([0,1])$. The unique solutions to $ig' = \pm ig$ are $g_\pm(x) = e^{\pm x}$. Both these functions belong to $\mathcal{D}(S_2^*)$ and so $\mathcal{N}(S_2^* \pm iI) \neq (0)$. Hence, S_2 is not essentially self-adjoint. However, neither of these functions belong to $\mathcal{D}(S_1^*)$. Tempting to think S_1 is essentially self-adjoint, but remember we already showed that it is not even symmetric! So some care must be taken not to fall into a trap!

Here is one reason that these concepts are important. If we imagine that our Hilbert space represents states of some system and that system evolves over time, then we would expect that there is a family of unitaries, $U(t)$ so that if we are in state h_0 at time 0, then at time t we are in state $U(t)h_0$,

Definition 3.81. A family of operators $\{U(t) : t \in \mathbb{R}\}$ on a Hilbert space $\mathcal{H}$ is called a **strongly continuous one-parameter unitary group** provided that:

1. $U(t)$ is a unitary operator for all $t \in \mathbb{R}$,

2. for each $h \in \mathcal{H}$ the function $t \to U(t)h$ is continuous,

3. $\forall t, s \in \mathbb{R}$, $(t)U(s) = U(t+s)$.

Theorem 3.82 (Stone's Theorem)**.** *Let A be a self-adjoint operator on $\mathcal{H}$ and use the functional calculus to define $U(t) = e^{iAt}$. Then $U(t)$ is a strongly continuous one-parameter unitary group and for $h \in \mathcal{D}(A)$, $\|\frac{U(t)h-h}{t} - iAh\| \to 0$ as $t \to 0$. Conversely, let $U(t)$ be a strongly continuous one-parameter unitary group, set*

$$\mathcal{D}(A) = \{h \in \mathcal{H} : \lim_{t \to 0} \frac{U(t)h - h}{t} exists\},$$

and set

$$Ah = -i \lim_{t \to 0} \frac{U(t)h - h}{t}.$$

Then A is self-adjoint and $U(t) = e^{iAt}$.

One instructive example is to consider $\mathcal{H} = L^2(\mathbb{R})$ and set $[U(t)f](x) = f(x+t)$. Then $f \in \mathcal{D}(A)$ precisely, when the difference quotients, $\frac{f(x+t)-f(x)}{t}$ converge in norm to a square-integrable function. That function must be equal to f' almost everywhere. Again measure theory shows that $\mathcal{D}(A) = \{f \in AC(\mathbb{R}) : f' \in L^2(\mathbb{R})\}$. Thus, $Af = -if'$.

3.13 Introductory Quantum Algorithms

DK: Not decided on whether to include this. Could be a brief intro to the basic tools in circuit-gate approach to quantum computing, and then give Deutsch algorithm as example, and possibly Grover search algorithm.

3.14 Von Neumann Algebras

DK: probably merge this into next chapter.

These were originally called "Algebras of Operators" by von Neumann but their name has since been changed in his honor. This is a difficult theory, but his motivation for studying these was quantum mechanics, so we start with that motivation.

In quantum mechanics the unit vectors in a Hilbert space represent the states of a system. Suppose that the system is in state $\psi \in \mathcal{H}$ and we want to perform a measurement that has K possible outcomes. Then the theory says that there will exist K measurement operators, $M_1, ...M_K \in B(\mathcal{H})$ such that:

- the probability of getting outcome k is $p_k = \|M_k \psi\|^2$,
- if outcome k is observed then the state of the system changes to $\frac{M_k \psi}{\|M_k \psi\|}$.

The fact that $1 = \sum_{k=1}^K p_k$ implies that $I = \sum_{k=1}^K M_k^* M_k$. Von Neumann argued that in certain settings the underlying state space could have a family of unitaries that acted upon it $\{U_a : a \in A\}$ such that if the system was in state $U_a \psi$ and we observed outcome k, then the state of the system should change to $\frac{U_a M_k \psi}{\|M_k U_a \psi\|}$. This implies that $U_a M_k = M_k U_a, \forall a \in A$.

This lead him to study sets of operators that commute with a set of unitary operators. He denoted such sets by $\mathcal{M}$ since this is what he thought sets of measurement operators should look like. For this reason von Neumann algebras are still generally denoted by the letter $\mathcal{M}$.

Definition 3.83. Given a set $\mathcal{S} \subseteq B(\mathcal{H})$ we call the set

$$\mathcal{S}' = \{T \in B(\mathcal{H}) : TS = ST, \forall S \in \mathcal{S}\},$$

the **commutant** of $\mathcal{S}$.

Note that

$$\mathcal{S}' = \mathcal{S}''',$$

and

$$\mathcal{S}'' = \mathcal{S}''''.$$

Note that if U is a unitary and $UT = TU$, then $TU^* = U^*(UT)U^* = U^*(TU)U^* = U^*T$, so T commutes with the adjoint too. Also T always commutes with I.

We briefly recall weak, strong and weak* convergence. A net of operators $\{T_\lambda\}_{\lambda \in D}$ converges to T in

- the weak-topology if $\langle k|T_\lambda h\rangle \to \langle k|Th\rangle$ for all $h, k \in \mathcal{H}$,
- the strong-topology if $\|T_\lambda h - Th\| \to 0$ for all $h \in \mathcal{H}$,
- the weak*-topology if $Tr(T_\lambda K) \to Tr(TK)$ for all $K \in \mathcal{C}_1(\mathcal{H})$.

We use $\mathcal{S}^{-w}$, $\mathcal{S}^{-s}$ and $\mathcal{S}^{-wk*}$, to denote the sets of operators that are limits of nets of operators from $\mathcal{S}$ in the weak, strong, and weak* sense.

Recall that a set $\mathcal{A}$ is called an algebra if it is a vector space and $X, Y \in \mathcal{A} \implies XY \in \mathcal{A}$.

Proposition 3.84. *Let $\mathcal{S} \subseteq B(\mathcal{H})$ be a set such that $I \in \mathcal{S}$ and $X \in \mathcal{S} \implies X^* \in \mathcal{S}$. Then $\mathcal{S}'$ is an algebra of operators that is closed in the weak, strong and weak* topologies.*

Theorem 3.85 (von Neumanns bicommutant Theorem)**.** *Let $\mathcal{A} \subseteq B(\mathcal{H})$ be an algebra of operators such that $I \in \mathcal{A}$ and $X \in \mathcal{A} \implies X^* \in \mathcal{A}$. Then $\mathcal{A}'' = \mathcal{A}^{-w} = \mathcal{A}^{-s} = \mathcal{A}^{-wk*}$.*

Thus, not only are operators that can be realized as limits in these three senses all equal, but something defined purely algebraically, $\mathcal{A}''$ is equal to something defined as a topological closure.

Corollary 3.86. *Let $\mathcal{M} \subseteq B(\mathcal{H})$ be an algebra of operators such that $I \in \mathcal{M}$ and $X \in \mathcal{M} \implies X^* \in \mathcal{M}$. The following are equivalent:*

- $\mathcal{M} = \mathcal{M}''$,
- $\mathcal{M} = \mathcal{M}^{-s}$,
- $\mathcal{M} = \mathcal{M}^{-w}$,
- $\mathcal{M} = \mathcal{M}^{-wk*}$.

Corollary 3.87. *Let $\mathcal{S} \subseteq B(\mathcal{H})$ satisfy $I \in \mathcal{S}$ and $X \in \mathcal{S} \implies X^* \in \mathcal{S}$. Then $\mathcal{S}''$ is the smallest weak* closed algebra containing $\mathcal{S}$.*

Definition 3.88. Any $\mathcal{M} \subseteq B(\mathcal{H})$ satisfying $I \in \mathcal{M}$, $X \in \mathcal{M} \implies X^* \in \mathcal{M}$ and $\mathcal{M} = \mathcal{M}''$ is called a **von Neumann algebra**.

Murray and von Neumann set about to classify all such algebras. This program goes on to this day and has an influence on quantum mechanics, which we will try to explain. First, we discuss the classification program.

Definition 3.89. A von Neumann algebra $\mathcal{M} \subseteq B(\mathcal{H})$ is called a **factor** if $\mathcal{M} \cap \mathcal{M}' = \mathbb{C} \cdot I$.

Just like all integers can be decomposed into products of primes, Von Neumann proved that every von Neumann algebra could be built from factors by something called he called **direct integration theory**. Hence, to understand all von Neumann algebras we only need to understand all factors.

The next step that Murray and von Neumann made was to break factors down into three types.

Definition 3.90. Let $\mathcal{M}$ be a von Neumann algebra and let $E, F \in \mathcal{M}$ be projections. We write $E \leq F$ if $\mathcal{R}(E) \subseteq \mathcal{R}(F)$ and $E < F$ when $E \leq F$ and $E \neq F$. We say that $F \neq 0$ is **minimal** if $E < F \implies E = 0$. We say that E and F are **Murray-von Neumann equivalent** and write $E \sim F$ if there exists a partial isometry $W \in \mathcal{M}$ such that $E = W^*W, F = WW^*$. We say that F is **finite** if there is no E such that $E \sim F$ and $E < F$.

Some examples are useful.

Example 3.91. Let $\mathcal{M} = B(\mathbb{C}^m)$. Then F is minimal if and only if F is a rank one projection, $E \sim F$ if and only if E and F are projections onto subspaces of the same dimension. Hence, every projection is finite.

Example 3.92. Let $\mathcal{M} = B(\mathcal{H})$, where $\mathcal{H}$ is infinite dimensional. Again a projection is minimal if and only if it is rank one. Since a partial isometry with $E = W^*W, F = WW^*$ is an isometry from $\mathcal{R}(E)$ to $\mathcal{R}(F)$ and isometries preserve inner products, they take an onb for one space to an onb to the other space. Hence, $E \sim F \iff dim_{HS}(\mathcal{R}(E)) = dim_{HS}(\mathcal{R}(F))$. On the other hand as soon as a set is infinite we can throw away one element and the subset has the same cardinality. Hence, as soon as $\mathcal{R}(F)$ is infinite dimensional, we can take an onb, throw away one element and that will be an onb for a subspace $\mathcal{R}(E) \subseteq \mathcal{R}(F)$ of the same dimension. We can now take a partial isometry that sends the onb for $\mathcal{R}(F)$ to the onb for $\mathcal{R}(E)$. This shows that $E \sim F$ with $E < F$. Hence, F is NOT finite in the M-vN sense.

Hence, the only finite projections are the projections onto finite dimensional subspaces.

Definition 3.93. A von Neumann factor is called **Type I** if it has a minimal projection. It is **Type II** if it has no minimal projections, but has a finite projection. It is called **Type** II_1 is it is Type II and the identity is a finite projection. If it is Type II but not Type Type II_1, then it is called **Type** II_∞. It is called **Type III** if it is not Type I or Type II.

Theorem 3.94 (Murray-von Neumann)**.** *A von Neumann algebra is Type I if and only if it is *-isomorphic to $B(\mathcal{H})$ for some $\mathcal{H}$.*

Definition 3.95. Let $\mathcal{M}$ be a von Neumann algebra. A map $\tau : \mathcal{M} \to \mathbb{C}$ is called a **state** is $\tau(I) = 1$ and $p \geq 0 \implies \tau(p) \geq 0$. A map is called a **tracial state** if it is a state and satisfies $\tau(XY) = \tau(YX)$. It is **faithful** if $\tau(X^*X) = 0 \implies X = 0$.

For $B(\mathbb{C}^n)$, there is only one faithful tracial state, namely $\tau(X) = \frac{1}{n}Tr(X)$. Note that in this setting the possible traces of projections are the numbers $\{\frac{k}{n} : 0 \leq k \leq n\}$, which represent the "fractional" dimension of the corresponding subspace.

Theorem 3.96 (Murray-von Neumann)**.** *Let $\mathcal{M}$ be a Type II_1 factor. Then:*

- *there exists a faithful tracial state, $\tau : \mathcal{M} \to \mathbb{C}$ that is also continuous in the weak*-topology,*
- *for every $0 \leq t \leq 1$ there exists a projection $P \in \mathcal{M}$ with $\tau(P) = t$,*
- *if $P, Q \in \mathcal{M}$ are projections, then $P \sim Q \iff \tau(P) = \tau(Q)$.*

This lead von Neumann to refer to "continuous geometries" since, unlike finite dimensions, in a Type II_1 setting there are subspaces of every (fractional) dimension $t, 0 \leq t \leq 1$.