5 Recap of Hilbert and Banach spaces

Some references for functional analysis are:

- Reed–Simon, Methods of Modern Mathematical Physics I, II.
- Halmos, Introduction to Hilbert space and the theory of spectral multiplicity.
- Arveson, A short course on spectral theory.
- Conway, A course in functional analysis.
- Kreyszig, Introductory functional analysis with applications.
- Berezin–Shubin, The Schrödinger equation.

There are many other references, especially if you want to study PDEs. Reed–Simon has quantum mechanics and Schrödinger operators in mind, and suffices for most of our purposes.

5.1 Inner products, norms, metrics

Let V be a metric space with distance function $d(\cdot,\cdot)$. Recall that a sequence $v_1, v_2, \ldots$ in V is *Cauchy* if for any $\epsilon > 0$, there exists N such that

$$d(v_m, v_n) < \epsilon, \qquad \forall\, m, n > N.$$

Although elements of a Cauchy sequence get arbitrarily close to one another, the putative limit may not exist in V. A metric space (V, d) is said to be *complete* if every Cauchy sequence in V converges in V.

On a vector space V over $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$, a (positive-definite) *norm* is a map

$$||\cdot|| : V \to [0, \infty)$$

such that

- $||v|| = 0$ iff $v = 0$;
- $||\lambda v|| = |\lambda| \cdot ||v||$ for all $\lambda \in \mathbb{F}, v \in V$;
- $||v + v'|| \leq ||v|| + ||v'||$ for all $v, v' \in v$.

A norm turns the vector space V into a metric space, via

$$d(v, v') = ||v' - v||.$$

As with any metric space, the balls $B_r(v)$ of points with distance $< r$ from $v \in V$ can be defined, and these balls generate the open sets of a topology on V.

If V has a (positive-definite) inner product, $\langle \cdot | \cdot \rangle$, then a norm, thus also a metric, is obtained by taking

$$||v|| = \sqrt{\langle v|v \rangle}.$$

This follows from the Cauchy–Schwarz inequality,

$$|\langle v|v' \rangle|^2 \leq \langle v|v \rangle \langle v'|v' \rangle.$$

(Note: For the $\mathbb{F} = \mathbb{C}$ case, we adopt the convention that an inner product is antilinear in the first argument and linear in the second argument.)

If v, v' are orthogonal, i.e., $\langle v|v' \rangle = 0$, then the *Pythagoraean identity* holds,

$$||v + v'||^2 = ||v||^2 + ||v'||^2.$$

A norm which arises from an inner product will satisfy the *parallelogram identity*,

$$||v + v'|| + ||v - v'|| = 2\left(||v|| + ||v'||\right).$$

Conversely, if a norm satisfies the parallelogram identity, it comes from an inner product via the *polarization* identity (Exercise),

$$\langle v|v' \rangle = \frac{1}{4}\left(||v + v'||^2 - ||v - v'||^2\right) + \frac{i}{4}\left(||v - iv'||^2 - ||v + iv'||^2\right). \tag{5.1}$$

5.2 Hilbert and Banach spaces

Definition 11. A *Banach space* V (respectively, *Hilbert space* $\mathcal{H}$) is a normed vector space (respectively, inner product space) over $\mathbb{C}$, for which the metric is complete.

5.2.1 ℓ^p spaces

Let c_{00} be the vector space of sequences $x : \mathbb{N} \to \mathbb{C}$ with finitely many non-zero entries. For each $1 \leq p < \infty$, we may define the ℓ^p norm

$$||x||_p = \left(\sum_{i \in \mathbb{N}} |x_i|^p \right)^{1/p}. \tag{5.2}$$

For $p = \infty$, the ℓ^∞ norm is

$$||x||_\infty = \sup_{i \in \mathbb{N}} |x_i|.$$

For $p = 2$, an inner product can be defined by

$$\langle x | x' \rangle = \sum_{i \in \mathbb{N}} \overline{x_i} x'_i, \tag{5.3}$$

and this induces the ℓ^2 norm.

Example 5.1. The sequence space c_{00} equipped with the ℓ^p norms are not Banach spaces (Exercise).

As a general abstract fact, we may densely embed a non-complete metric space V into a complete metric space $\overline{V}$, uniquely up to isometry. But this is usually not good enough. We want $\overline{V}$ to be directly defined with the same procedures as V is.

Example 5.2. For $1 \leq p < \infty$, the space ℓ^p is defined as the subset of sequences $x : \mathbb{N} \to \mathbb{C}$ such that

$$\sum_{i \in \mathbb{N}} |x_i|^p < \infty.$$

Then Eq. (5.2) defines a norm on ℓ^p, making it a *Banach space* (Exercise). Similarly, ℓ^2 is a Hilbert space, under the inner product (5.3).

For sequences indexed by $\mathbb{Z}$ instead of $\mathbb{N}$, we write $\ell^p(\mathbb{Z})$.

5.2.2 L^p spaces

The theory of measure spaces and Lebesgue integration are dealt with in detail in standard analysis textbooks. We give a summary here.

Recall that the general definition of a σ-algebra on a set X is collection σ of subsets of X such that

- $X \in \Sigma$,
- $A \in \Sigma$ iff $A^c = X \setminus A \in \Sigma$,
- If $A_1, A_2, \ldots \in \Sigma$, then $A = \cup_{i \in \mathbb{N}} A_i \in \Sigma$.

A σ-algebra determines which subsets of X are *measurable*. A function $f : X \to Y$ between spaces with σ-algebras is said to be *measurable*, if the preimages of measurable subsets are measurable.

Let X be equipped with a σ-algebra Σ. A *measure* on (X, Σ) is an assignment

$$\mu : \Sigma \to [0, \infty]$$

such that

- $\mu(\emptyset) = 0$,
- $\mu\big(\bigsqcup_{i \in \mathbb{N}} A_i\big) = \sum_{i \in \mathbb{N}} \mu(A_i)$, for disjoint $A_i \in \Sigma$.

The triple (X, Σ, μ) is called a *measure space*.

It is often convenient to admit all subsets of μ-measure-0 sets as measurable sets. Write $\tilde{\Sigma}$ for the σ-algebra generated by Σ and these extra sets. Elements of $\tilde{\Sigma}$ have the form $A \cup N$ where $A \in \Sigma$ and N is one of the extra sets. Then μ is *completed* to a measure $\tilde{\mu}$ by taking $\tilde{\mu}(A \cup N) = \mu(A)$.

Definition 12. If X is a topological space, then the *Borel σ-algebra* is the smallest σ-algebra containing all the open sets of X. A *Borel measure* is a measure on a Borel σ-algebra. If X is a metric space, *metric measure* is a Borel measure for the topology induced by the metric.

Example 5.3. Consider $\mathbb{R}$ with the standard metric, topology, and Borel σ-algebra. We may assign $\mu(a, b] = b - a$ for any $a, b \in \mathbb{R}$, and extend this to a measure on the Borel σ-algebra. The completion is the *Lebesgue* measure on $\mathbb{R}$. On $\mathbb{R}^d$, the Lebesgue measure is the completion of the n-fold product of the Lebesgue measure on $\mathbb{R}$.

Example 5.4. If X is a (path-connected) Riemannian manifold, we can regard it as a metric space, and there is a canonical way to get a metric measure. This will be explained in lectures on Riemannian volume forms/densities.

The *Lebesgue integral* is a generalization of the Riemann integral. Furthermore, it can deal with functions on general measure spaces. We refer to standard textbooks, e.g. Rudin's Real and Complex Analysis, Folland's Real Analysis, for full details of how $\int_X f\,d\mu$ is defined. Here is a brief summary.

A measurable function s is *simple* if its range is just a finite set of points $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$. If we set $A_i := \{t \in X : s(t) = \alpha_i\}$ and write χ_{A_i} for the characteristic function of A_i, then

$$s = \sum_{i=1}^{n} \alpha_i \chi_{A_i}.$$

When s is $\mathbb{R}_{\geq 0}$-valued, we define

$$\int_X s\,d\mu := \sum_{i=1}^{n} \alpha_i \mu(A_i).$$

For a general $[0, \infty]$-valued measurable function f, we define

$$\int_X f\,d\mu := \sup_{0 \leq \text{simple}\, s \leq f} \int_X s\,d\mu.$$

A complex-valued measurable function f is *integrable* with respect to μ if $\int_X |f|\,d\mu < \infty$. In that case, the integral $\int_X f\,d\mu \in \mathbb{C}$ is obtained by integrating separately the positive/negative parts of the real/imaginary parts of f, then assembling the four non-negative results back into a single complex number.

An important point is that many properties of functions are stated as μ-*almost everywhere* (μ-a.e.) properties. For example, measurable functions f, g are equal μ-a.e. if the subset on which they differ has μ-measure zero. This is another place where strict pointwise values of functions lose their importance.

Definition 13. On a space X with measure μ, and for $1 \leq p < \infty$, the space $L^p(X) = L^p(X, \mu)$ is defined to be the μ-a.e. equivalence classes of measurable functions $f : X \to \mathbb{C}$, such that

$$||f||_p = \left(\int_X |f|^p\,d\mu \right)^{1/p} < \infty.$$

For $p = 2$, we define the inner product,

$$\langle f|g\rangle = \int_X \overline{f} g \, d\mu,$$

which induces the norm $||\cdot||_2$. The space $L^\infty(X)$ is the set of μ-a.e. equivalence classes of μ-a.e. bounded functions $f : X \to \mathbb{C}$, and we define

$$||f||_\infty = \operatorname{ess\,sup}|f| = \inf\{\nu \in \mathbb{R}_{\geq 0} \,:\, |f(x)| \leq \nu \text{ holds } \mu\text{-a.e.}\}.$$

Using the basic monotone/dominated convergence theorems for the Lebesgue integral, one shows that the $L^p(X)$ are Banach spaces, and that $L^2(X)$ is a Hilbert space. If we used the Riemann integral (for $X = \mathbb{R}^d$), we would not have enough p-integrable functions to achieve completeness.

5.3 Basic Hilbert space structures

5.3.1 Orthogonal projection and complement

If V is an inner product space, and $W \subset V$ is a linear subspace, the *orthogonal complement* of W is

$$W^\perp := \{v \in V \,:\, \langle v|w\rangle = 0 \;\; \forall w \in W\}.$$

Clearly $W^\perp$ is a subspace, and $W \cap W^\perp = \{0\}$.

Exercise 5.1. For a subspace $W \subset V$, check the following properties of the orthogonal complement (Exercise):

- $W \subset W^{\perp\perp}$.
- If $W \subset W'$, then $W'^\perp \subset W^\perp$.
- $W^\perp = W^{\perp\perp\perp}$.

Suppose V is finite-dimensional. Then given any $v \in V$ and any subspace $W \subset V$, there is a unique $w_v \in W$ which is nearest to v. Furthermore, the vector $v - w_v$ lies in $W^\perp$. Thus we have an orthogonal decomposition,

$$V = W \oplus W^\perp, \qquad v = w_v + \big(v - w_v\big). \tag{5.4}$$

If V is infinite-dimensional, the above generally fails, because "the point in W nearest to v" may not exist. This issue is remedied when V is a Hilbert space $\mathcal{H}$.

Lemma 5.1. *Let $W \subset \mathcal{H}$ be a closed subspace of a Hilbert space, and $v \in \mathcal{H}$. There exists a unique $w_v \in W$ such that*

$$||v - w_v|| = \mathrm{dist}(v, W) \equiv \inf_{w \in W} ||v - w||.$$

Furthermore, w_v is the unique vector in W such that $v - w_v \in W^\perp$.

Proof. Write $\delta = \inf_{w \in W} ||v - w||$. Let $\{w_n\}$ be a sequence in W with $||v - w_n|| \to \delta$. By the parallelogram identity,

$$\begin{aligned}
||w_m - w_n||^2 &= ||(v - w_n) - (v - w_m)||^2 \\
&= -||(v - w_n) + (v - w_m)||^2 + 2(||v - w_n||^2 + ||v - w_m||^2) \\
&= -4||v - \underbrace{\tfrac{1}{2}(w_m + w_n)}_{\in W}||^2 + 2(||v - w_n||^2 + ||v - w_m||^2) \\
&\leq -4\delta^2 + 2(||v - w_n||^2 + ||v - w_m||^2) \\
&\underset{m,n\to\infty}{\to} -4\delta^2 + 2(\delta^2 + \delta^2) = 0.
\end{aligned}$$

Thus $\{w_n\}$ is a Cauchy sequence. Since $\mathcal{H}$ is a Hilbert space and W is closed, $w_n \to w_v$ for some $w_v \in W$. The norm is continuous, so $||v - w_v|| = \lim_{n\to\infty} ||v - w_n|| = \delta$.

For the second part, note that

$$||v - w_v||^2 = \delta^2 \leq ||v - w_v + \lambda w||^2, \qquad \forall \lambda \in \mathbb{C}, w \in W.$$

So

$$\begin{aligned}
0 &\leq ||v - w_v + \lambda w||^2 - ||v - w_v||^2 \\
&= \lambda\langle v - w_v | w\rangle + \bar{\lambda}\langle w | v - w_v\rangle + |\lambda|^2||w||^2, \qquad \forall \lambda \in \mathbb{C}, w \in W.
\end{aligned}$$

In particular, fix w and allow for $\lambda = \mu\langle w | v - w_v\rangle$ with real μ, so

$$0 \leq 2\mu|\langle v - w_v | w\rangle|^2 + \mu^2|\langle v - w_v | w\rangle|^2||w||^2, \qquad \forall \mu \in \mathbb{R}.$$

For small negative μ, the first term dominates, but it is non-positive, so we must have $\langle v - w_v | w\rangle = 0$. Since $w \in W$ is arbitrary, $v - w_v \in W^\perp$.

The uniqueness parts are omitted. $\square$

Theorem 5.2. *Let W be a* closed *subspace of a Hilbert space $\mathcal{H}$. Then $\mathcal{H} = W \oplus W^\perp$.*

Proof. Lemma 5.1 shows that $\mathcal{H} = W + W^\perp$, and the latter is a direct sum since $W \cap W^\perp = \{0\}$. $\square$

Theorem 5.2 says that orthogonal decomposition, (5.4), holds in Hilbert spaces. In this sense, Hilbert spaces have a geometry much like Euclidean spaces.

Exercise 5.2. Let W be a subspace of a Hilbert space $\mathcal{H}$.

- Check that $W^\perp$ is a closed subspace.
- Show that $W^\perp = \overline{W}^\perp$.
- Use Lemma 5.1 to deduce that

$$\overline{W} = W^{\perp\perp}.$$

Exercise 5.3. Check that c_{00} is a dense subspace of ℓ^2, but it is not closed. Similarly for the continuous functions $C[0,1]$ inside $L^2[0,1]$.

5.3.2 Orthonormal bases

Definition 14. Let $\mathcal{H}$ be a Hilbert space. A family $\{e_i\}_{i\in I}$ is *orthonormal* if

$$\langle e_i | e_j \rangle = \delta_{ij}, \qquad \forall\, i, j, \in I.$$

It is called an *orthonormal basis* for $\mathcal{H}$ if the closure of the linear span of the e_i is $\mathcal{H}$.

Here, "linear span" is an algebraic notion — only finite linear combinations are allowed. When we take the closure of the linear span, convergent infinite linear combinations are allowed. So the linear span of an orthonormal basis is *dense* in the Hilbert space.

Theorem 5.3. *Every Hilbert space has an orthonormal basis, and the cardinalities of any two such bases are the same. A separable Hilbert space has a countable orthonormal basis.*

Here, a separable space is one which has a countable dense subset. *We will assume separability of Hilbert spaces unless otherwise stated.*

Proof. This basically follows from a Zorn's Lemma argument (details omitted). □

Theorem 5.4 (Parseval–Plancherel)**.** *Let $\{e_i\}_{i\in\mathbb{N}}$ be an orthonormal basis for a (separable, infinite-dimensional) Hilbert space $\mathcal{H}$. The map*

$$\begin{aligned} f : \mathcal{H} &\to \ell^2 \\ \psi &\mapsto (\langle e_i|\psi\rangle)_{i\in\mathbb{N}} \end{aligned}$$

is an isomorphism of Hilbert spaces. If $\varphi, \psi \in \mathcal{H}$, their inner product may be written as

$$\langle\varphi|\psi\rangle = \sum_{i\in\mathbb{N}} \langle\varphi|e_i\rangle\langle e_i|\psi\rangle. \tag{5.5}$$

In particular,

$$||\psi||^2 = \sum_{i\in\mathbb{N}} |\langle e_i|\psi\rangle|^2.$$

Proof. Exercise. □

Eq. (5.5) in Theorem 5.4 is symbolically written in physics as the insertion of a "resolution of identity", or "completeness relation",

$$\langle\varphi|\psi\rangle = \langle\varphi| \left(\sum_{i\in\mathbb{N}} |e_i\rangle\langle e_i|\right) |\psi\rangle, \qquad \left(\sum_{i\in\mathbb{N}} |e_i\rangle\langle e_i|\right) = 1_{\mathcal{H}}.$$

The precise meaning of the right side is actually not so straightforward, since it involves an infinite sum of rank-1 projection *operators*, whose convergence requires an understanding of operator topologies.

5.3.3 Duality

If V is a normed vector space (over $\mathbb{C}$), the space of *continuous* linear functionals $V \to \mathbb{C}$ is called the *dual space*, and is denoted V^*. We may define a norm on V^* via

$$||f|| := \sup_{||v||=1} |f(v)|, \qquad f \in V^*.$$

In fact, V^* becomes a Banach space in this way (Exercise).

In particular, for Hilbert spaces, each vector $\varphi \in \mathcal{H}$ is associated to a linear functional,

$$\langle\varphi| : \psi \mapsto \langle\varphi|\psi\rangle, \qquad \psi \in \mathcal{H}, \tag{5.6}$$

and $\langle\varphi|$ is continuous (why?). Remarkably, the converse is true as well; more precisely, there is the following duality between $\mathcal{H}$ and $\mathcal{H}^*$.

Theorem 5.5 (Riesz representation theorem)**.** *The assignment*

$$\begin{aligned} g : \mathcal{H} &\to \mathcal{H}^* \\ \varphi &\mapsto \langle\varphi| \end{aligned}$$

is antilinear, norm-preserving, and bijective. Thus $\mathcal{H}^*$ is a Hilbert space under the inner product

$$\langle f|f'\rangle_{\mathcal{H}^*} = \langle g^{-1}(f')|g^{-1}(f)\rangle_{\mathcal{H}}$$

and it is canonically anti-isomorphic to $\mathcal{H}$.

Proof. Let $f \neq 0 \in \mathcal{H}^*$. Since $\ker(f)$ is a proper closed subspace of $\mathcal{H}$, Theorem 5.2 says that $\ker(f)^\perp$ is non-zero, and we may pick some nonzero $\varphi \in \ker(f)^\perp$. We shall verify that

$$\tilde{\varphi} = \overline{\frac{f(\varphi)}{||\varphi||^2}}\varphi$$

represents f. To see this, first observe that given any $\psi \in \mathcal{H}$, we have

$$\psi - \tfrac{f(\psi)}{f(\varphi)}\varphi \in \ker(f) \;\Rightarrow\; \langle\varphi|\psi - \tfrac{f(\psi)}{f(\varphi)}\varphi\rangle = 0.$$

It follows that

$$\begin{aligned} \langle\tilde{\varphi}|\psi\rangle = \frac{f(\varphi)}{||f(\varphi)||^2}\langle\varphi|\psi\rangle &= \frac{f(\varphi)}{||f(\varphi)||^2}\langle\varphi|\psi - \tfrac{f(\psi)}{f(\varphi)}\varphi + \tfrac{f(\psi)}{f(\varphi)}\varphi\rangle \\ &= \frac{f(\varphi)}{||f(\varphi)||^2}\langle\varphi|\tfrac{f(\psi)}{f(\varphi)}\varphi\rangle = f(\psi) \qquad\qquad \forall\psi \in \mathcal{H}, \end{aligned}$$

and therefore $f = \langle\tilde{\varphi}|$. That this representative is unique is straightforward. The remaining statements are left as an exercise. □

Remark. The notation $\langle\varphi|$ for the linear function (5.6) is due to Dirac, who called it a "bra vector". He also wrote $|\psi\rangle \equiv \psi$ and called it a "ket vetor". The result of applying $\langle\varphi|$ to a vector ψ is the numerical "bra-ket" $\langle\varphi|\psi\rangle \in \mathbb{C}$.

Remark 2. It is useful to think of $\psi = |\psi\rangle \in \mathcal{H}$ as a linear map

$$|\psi\rangle : \mathbb{C} \to \mathcal{H}, \qquad \lambda \mapsto \lambda\psi,$$

and compare it to the associated linear functional

$$\langle\psi| : \mathcal{H} \to \mathbb{C}.$$

Indeed, we shall see that the duality between $|\psi\rangle$ and $\langle\psi|$ is a special case of the adjoint operation.

6 Bounded linear operators and spectrum

Lots of linear operators on normed vector spaces are discontinuous. For example, consider the Hilbert space $\mathcal{H} = L^2([0, 2\pi])$. The functions $e_n(x) = \frac{1}{\sqrt{2\pi}} e^{inx}, n \in \mathbb{Z}$ are orthonormal. We can apply the momentum (derivative) operator to each e_n,

$$-i\frac{d}{dx} : e_n \mapsto ne_n,$$

and extend this to finite linear combinations without issue.

But suppose we have an infinite sum, such as

$$\psi = \sum_{n\in\mathbb{N}} \frac{1}{n} e_n,$$

which has square-summable coefficients, thus $\psi \in \mathcal{H}$. If we tried to commute $\sum_{n\in\mathbb{N}}$ past $-i\frac{d}{dx}$, we would get

$$-i\frac{d}{dx} \sum_{n\in\mathbb{N}} \psi \stackrel{?}{=} \sum_{n\in\mathbb{N}} -i\frac{d}{dx}\left(\frac{1}{n} e_n\right) \stackrel{?}{=} \sum_{n\in\mathbb{N}} e_n, \tag{6.1}$$

but the right side does not converge in $\mathcal{H}$. The problem is that there is no uniform bound on how $-i\frac{d}{dx}$ scales the Hilbert space norm. This prevents us from extending the definition of $-i\frac{d}{dx}$ to all of $\mathcal{H}$ by "continuity", as attempted in (6.1). Of course, we could extend it *dis*continuously to all of $\mathcal{H}$, but this is seldom useful, and is not a canonical procedure.

Note that the Laplace operator is $(-i\frac{d}{dx})^2$, so we already have to confront discontinuous operators in the most elementary quantum mechanics Hamiltonian.

We could fix this problem by taking the "bounded transform" of $-i\frac{d}{dx}$, written

$$\left(-i\frac{d}{dx}\right)_{\text{bounded}} : e_n \mapsto \frac{n}{\sqrt{1+n^2}} e_n,$$

which is now a norm-decreasing map. Then we may extend this operator to arbitrary elements $\psi \in \mathcal{H}$ by continuity.

However, the passage from unbounded operators, such as $-i\frac{d}{dx}$, to bounded operators, is actually very subtle. For now, let us focus on the more well-behaved setting of bounded/continuous linear operators.

Definition 15. Let V_1, V_2 be normed vector spaces. A linear operator $T : V_1 \to V_2$ is *bounded*, or *continuous*, if its *operator norm* is finite,

$$||T||_{\mathrm{op}} := \sup_{||\psi||=1} ||T\psi|| < \infty.$$

We write $\mathcal{B}(V_1, V_2)$ for the normed linear space of bounded linear operators $V_1 \to V_2$. If $V_1 = V_2 = V$, we simply write $\mathcal{B}(V)$.

Notation: By default, $||T||$ refers to the operator norm above.

Exercise 6.1. Check that a bounded operator T according to Definition 15 is indeed continuous, that $||\cdot||_{\mathrm{op}}$ is indeed a norm, and that it is submultiplicative,

$$||T \circ T'||_{\mathrm{op}} \leq ||T||_{\mathrm{op}}\, ||T'||_{\mathrm{op}}, \qquad T' \in \mathcal{B}(V_1, V_2), T \in \mathcal{B}(V_2, V_3).$$

Exercise 6.2. Let V_1 be a normed vector space and let V_2 be a Banach space. Suppose $V_0 \subset V_1$ is a dense linear subspace, and $T : V_0 \to V_2$ is a linear operator such that

$$||T||_{\mathrm{op}} = \sup_{v \in V_0 : ||v||=1} ||Tv|| < \infty.$$

Then there is a unique extension of T to an operator $\tilde{T} \in \mathcal{B}(V_1, V_2)$, and $||\tilde{T}||_{\mathrm{op}} = ||T||_{\mathrm{op}}$.

For example, a Hilbert space operator $\mathcal{H}_1 \to \mathcal{H}_2$ is often initially defined by its action on some orthonormal basis of $\mathcal{H}_1$. Then one takes V_0 to be the linear span of that orthonormal basis, and applies Exercise 6.2 to get a bounded operator $\mathcal{H}_1 \to \mathcal{H}_2$.

Exercise 6.3. Check that if V_2 is a Banach space, then $\mathcal{B}(V_1, V_2)$ with the operator norm is a Banach space. (In particular, this applies to the space of bounded operators $V \to V$ on a Banach/Hilbert space V.)

6.1 Spectrum of bounded operators on Banach spaces

Definition 16. Let V be a Banach space. The *spectrum* of $T \in \mathcal{B}(V)$ is defined to be

$$\sigma(T) := \{\lambda \in \mathbb{C} \,:\, \lambda - T \text{ not invertible in } \mathcal{B}(V)\}, \tag{6.2}$$

and the complement

$$\rho(T) := \mathbb{C} \setminus \sigma(T)$$

is the *resolvent set* of T.

Remark. Usually, when we say that $T \in \mathcal{B}(V)$ is invertible, we mean that it has an inverse in $\mathcal{B}(V)$. Actually, if $T \in \mathcal{B}(V)$ is invertible set-theoretically, then the set-theoretic inverse (which is linear) is automatically bounded! This remarkable fact is called the *bounded inverse theorem*, which is one of the famous core theorems of functional analysis; see Chapter III.5 of Reed–Simon for a proof.

Recall that λ is an *eigenvalue* of T if $\lambda - T$ is non-injective; in this case, T possesses a non-zero *eigenspace* of *eigenvectors* v such that $Tv = \lambda v$. In finite dimensions, the spectrum is the same thing as the set of eigenvalues, and every matrix has an eigenvalue (over $\mathbb{C}$). In infinite dimensions, the latter statement is replaced by Corollary 6.5 below. The set of eigenvalues is called the *point spectrum*, denoted

$$\sigma_{\mathrm{p}} = \{\lambda \in \mathbb{C} \,:\, \lambda - T \,\text{not injective}\}.$$

Obviously, $\sigma_{\mathrm{p}}(T) \subset \sigma(T)$. But not all elements of $\sigma(T)$ are eigenvalues!

Exercise 6.4. Consider the Banach space $\ell^2(\mathbb{Z})$, with canonical basis vectors

$$e_n = (\ldots, 0, \underbrace{1}_{n\text{-th}}, 0 \ldots), \qquad n \in \mathbb{Z}.$$

Let R be the right shift operator on $\ell^2(\mathbb{Z})$ taking each e_n to e_{n+1}.

- Show that $\sigma_{\mathrm{p}}(R)$ is empty.
- Let λ be a unit complex number, $|\lambda| = 1$. Find a sequence of unit vectors $v_1, v_2, \ldots \in \ell^2(\mathbb{Z})$, such that

  $$\lim_{n\to\infty} ||Tv_n - \lambda v_n|| = 0.$$

  (Thus λ is an approximate eigenvalue, see Definition 19.)
- Deduce that $\{\lambda \in \mathbb{C} \,:\, |\lambda| = 1\} \subset \sigma(T)$.

Proposition 6.1 (Neumann series)**.** *Let V be a Banach space and $T \in \mathcal{B}(V)$. If $|\lambda| > ||T||_{\mathrm{op}}$, then $\lambda \in \rho(T)$.*

Proof. By scaling, we may assume $\lambda = 1$ and $||T||_{\mathrm{op}} < 1$. Then sub-multiplicativity of $|| \cdot ||_{\mathrm{op}}$ implies that the Neumann series

$$\sum_{n=0}^{\infty} T^n$$

converges in $\mathcal{B}(V)$, and it is the inverse of $(1-T)$, since

$$(1-T)(1+T+\ldots+T^n) = 1-T^{n+1}, \qquad n \geq 1.$$

$\square$

Proposition 6.2 (Compactness of spectrum)**.** *Let V be a Banach space and $T \in \mathcal{B}(V)$. Then $\sigma(T)$ is a compact subset of $\mathbb{C}$.*

Proof. Fix any $\mu \in \rho(T)$. Then

$$\begin{aligned} ||1-(\mu-T)^{-1}(\lambda-T)|| &= ||(\mu-T)^{-1}\big((\mu-T)-(\lambda-T)\big)|| \\ &= ||(\mu-T)^{-1}|| \cdot |\mu-\lambda|, \qquad \lambda \in \mathbb{C}. \end{aligned}$$

In particular, for all λ sufficiently close to μ, the inequality

$$||1-(\mu-T)^{-1}(\lambda-T)|| < 1$$

holds, and Prop. 6.1 says that

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

is invertible. Thus $\lambda \in \rho(T)$. Since $\mu \in \rho(T)$ was arbitrary, we have shown that $\rho(T)$ is open, thus $\sigma(T)$ is closed. Due to Prop. 6.1, $\sigma(T)$ is bounded; thus it is compact. $\square$

Exercise 6.5. Let V be a Banach space, and $T \in \mathcal{B}(V)$. Show that

- The spectrum is invariant under a similarity transformation,

$$\sigma(T) = \sigma(ATA^{-1}).$$

- If T is invertible, then

$$\sigma(T^{-1}) = \sigma(T)^{-1} := \{\lambda^{-1} \,:\, \lambda \in \sigma(T)\}.$$

- If p is a polynomial, then

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

We end this section with a useful characterization of invertibility.

Proposition 6.3. *Let V be a Banach space, and $T \in \mathcal{B}(V)$. Then T is invertible iff T is bounded below and has dense range.*

Proof. Suppose T is invertible. Obviously it has dense range. Furthermore, for any $v \in V$,

$$||v|| = ||T^{-1}Tv|| \leq ||T^{-1}|| \cdot ||Tv||,$$

so T is bounded below by $||T^{-1}||^{-1}$.

Conversely, suppose that (i) the range of T is dense, and (ii) there exists some $\alpha > 0$ such that $||Tv|| \geq \alpha||v||$ for all $v \in V$. It follows from (ii) that T has closed range: if $Tv_n = y_n \to y$, then

$$||v_n - v_m|| \leq \alpha^{-1}||Tv_n - Tv_m|| = \alpha^{-1}||y_n - y_m||,$$

so $\{v_n\}$ is Cauchy, and has a limit $v \in V$; continuity of T means $Tv = y$. Combining this closed range property with (i), we see that T is surjective.

For injectivity, suppose $Tv = Tv'$, then $0 = ||T(v-v')|| \geq \alpha||v-v'||$, which implies $v = v'$. Thus the set-theoretic inverse T^{-1} exists (and is linear). It is straightforward to check that T^{-1} is bounded (in fact, $||T^{-1}||_{\text{op}} \leq \alpha^{-1}$.) $\square$

6.1.1 Existence of spectrum (optional)

Theorem 6.4. *Let V be a Banach space, and $T \in \mathcal{B}(V)$. The assignment of resolvents,*

$$\begin{aligned} \rho(T) &\to \mathcal{B}(V) \\ \lambda &\mapsto R_\lambda(T) := (\lambda - T)^{-1} \end{aligned}$$

is analytic on each component of $\rho(T)$.

Proof. For each $\mu \in \rho(T)$, let us find a power series expansion for $R_\lambda(T)$ near μ. Formally,

$$\begin{aligned} R_\lambda(T) \text{``}=\text{''} \frac{1}{\lambda - T} = \frac{1}{(\mu - T) - (\mu - \lambda)} &= \frac{1}{\mu - T}\left(1 - \frac{\mu - \lambda}{\mu - T}\right)^{-1} \\ &= \frac{1}{\mu - T}\left(\sum_{n=0}^{\infty}\left(\frac{\mu - \lambda}{\mu - T}\right)^n\right) \\ &\text{``}=\text{''} \sum_{n=0}^{\infty}(\mu - \lambda)^n (R_\mu(T))^{n+1}. \end{aligned}$$

Using the sub-multiplicativity of the operator norm, the above power series converges to $R_\lambda(T)$ for λ in the open ball

$$|\lambda - \mu| < ||R_\mu(T)||^{-1}.$$

$\square$

Corollary 6.5. *Let V be a Banach space. For any $T \in \mathcal{B}(V)$, the spectrum $\sigma(T)$ is non-empty (and compact).*

Proof. For large $|\lambda| > ||T||_{\mathrm{op}}$, the Neumann series for $R_\lambda(T)$ shows that $||R_\lambda(T)||_{\mathrm{op}} \to 0$ as $\lambda \to \infty$. Fix $\mu \in \rho(T)$. Since linear functionals separate elements[13] of $\mathcal{B}(V)$, we can choose $\Phi \in (\mathcal{B}(V))^*$ such that $\Phi(R_\mu(T)) \neq 0$.

Suppose $\sigma(T)$ is empty. Then the map

$$\mathbb{C} \to \mathbb{C}, \qquad \lambda \mapsto \Phi \circ R_\lambda(T)$$

will be a bounded analytic function tending to 0 at infinity, thus identically zero by Liouville's theorem in complex analysis. In particular, $\Phi(R_\mu(T)) = 0$, which is a contradiction. $\square$

Remark. The complex analysis proofs of non-emptiness of spectrum is the standard approach appearing in most textbooks. It is not so well-known that "elementary" proofs also exist, see, e.g. [Rickart, C., Michigan Math. J. 5 75–78 (1958)] or Kaniuth, E., A Course in Commutative Banach Algebras, pp. 12-13. For self-adjoint operators on Hilbert space, see Prop. 6.11 for a simpler proof.

6.2 Adjoints of Hilbert space operators

We shift focus to Hilbert spaces, and recall the bra-ket duality (Remark 2). There is a corresponding duality at the level of Hilbert space operators.

Definition 17. Let $\mathcal{H}_1, \mathcal{H}_2$ be Hilbert spaces, and $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$. The *adjoint* of T is defined to be the operator $T^* \in \mathcal{B}(\mathcal{H}_2, \mathcal{H}_1)$ satisfying

$$\langle T^*\varphi|\psi\rangle = \langle \varphi|T\psi\rangle, \qquad \forall\, \varphi \in \mathcal{H}_2, \psi \in \mathcal{H}_1. \tag{6.3}$$

[13] One form of the *Hahn–Banach theorem* states:

> Let V be a normed vector space, and $v, v' \in V$ be distinct. Then there exists $\Phi \in V^*$ such that $\Phi(v) \neq \Phi(v')$.

This is another non-trivial core theorem in functional analysis.

Now, the existence and uniqueness of the adjoint operator is not completely obvious. One way to show this is via the following correspondence result:

Lemma 6.6. *If $T \in \mathcal{B}(\mathcal{H})$, then*

$$f_T : \mathcal{H} \oplus \mathcal{H} \to \mathbb{C}, \qquad (\varphi, \psi) \mapsto \langle \varphi | T\psi \rangle$$

is a bounded sesquilinear form, whose norm

$$||f_T|| := \sup_{||\varphi||, ||\psi|| = 1} |f_T(\varphi, \psi)|$$

equals $||T||_{\mathrm{op}}$. Conversely, if f is a bounded sesquilinear form, there exists a unique $T_f \in \mathcal{B}(\mathcal{H})$, such that

$$f(\varphi, \psi) = \langle \varphi | T_f \psi \rangle \qquad \forall \varphi, \psi \in \mathcal{H}.$$

Proof. Exercise. The Riesz representation theorem is needed here. □

Remark. If you know about weak operator topologies, try to figure out how a bounded operator is approximated by linear combinations of ket-bras, $|\psi\rangle\langle\varphi|$, i.e., rank-1 operators. With this in mind, the adjoint duality for Hilbert space operators is a natural generalization of the Riesz duality for Hilbert spaces and their duals.

Exercise 6.6. For operators $T \in \mathcal{B}(\mathcal{H}_1, \mathcal{H}_2)$ and $T' \in \mathcal{B}(\mathcal{H}_2, \mathcal{H}_3)$, show that

- $T \mapsto T^*$ is antilinear.
- $(T^*)^* = T$.
- $||T^*|| = ||T||$.
- $(T'T)^* = T^* T'^*$.
- If T is invertible, then so is T^*, and $(T^*)^{-1} = (T^{-1})^*$.
- $||T^*T|| = ||TT^*|| = ||T||^2 = ||T^*||^2$.

As mentioned in Remark 2, we can think of $|\psi\rangle$ as a linear map $\mathbb{C} \to \mathcal{H}$ with adjoint $\langle\psi| : \mathcal{H} \to \mathbb{C}$. The kernel of $\langle\psi|$ comprises those vectors which are orthogonal to ψ, i.e.,

$$\ker(\langle\psi|) = (\mathrm{Ran}|\psi\rangle)^{\perp}.$$

More generally, the adjoint of $T \in \mathcal{B}(\mathcal{H})$ has the same property,

$$\ker(T^*) = (\operatorname{Ran} T)^\perp. \tag{6.4}$$

This follows from

$$T^*\psi = 0 \Leftrightarrow \langle\varphi|T^*\psi\rangle = 0 \Leftrightarrow \langle T\varphi|\psi\rangle = 0, \qquad \varphi, \psi \in \mathcal{H}.$$

6.2.1 Normal operators

Definition 18. An operator $T \in \mathcal{B}(\mathcal{H})$ is *normal* if $TT^* = T^*T$. Particular classes of normal operators are:

- *Self-adjoint*, if $T = T^*$;
- *Unitary*, if $TT^* = 1 = T^*T$;
- *(Orthogonal) projection*, if $T = T^* = T^2$.

There is a useful equivalent criterion for normality:

Proposition 6.7. *An operator $T \in \mathcal{B}(\mathcal{H})$ is normal iff $||T\psi|| = ||T^*\psi||$ for all $\psi \in \mathcal{H}$.*

Proof. A preparatory result is that $\langle\psi|T\psi\rangle = 0$ for all $\psi \in \mathcal{H}$ implies $T = 0$. To see this, observe that $0 = \langle\psi' + \psi|T(\psi' + \psi)\rangle = \langle\psi'|T\psi\rangle + \langle\psi|T\psi'\rangle$ as well. Replace ψ by $i\psi$ to get $0 = \langle\psi'|T\psi\rangle - \langle\psi|T\psi'\rangle$. Thus $0 = \langle\psi|T\psi'\rangle$ for all $\psi, \psi' \in \mathcal{H}$, and $T = 0$ follows readily.

Now, we have the equations

$$\begin{aligned} ||T\psi||^2 &= \langle T\psi|T\psi\rangle = \langle\psi|T^*T\psi\rangle, \\ ||T^*\psi||^2 &= \langle T^*\psi|T^*\psi\rangle = \langle\psi|TT^*\psi\rangle, \qquad \forall\psi \in \mathcal{H}. \end{aligned}$$

So $||T\psi|| = ||T^*\psi||$ for all $\psi \in \mathcal{H}$ iff $\langle\psi|(T^*T - TT^*)\psi\rangle = 0$ for all $\psi \in \mathcal{H}$ iff $T^*T - TT^* = 0$ iff T is normal. □

Note that any $T \in \mathcal{B}(\mathcal{H})$, can be decomposed as

$$T = \underbrace{\frac{1}{2}(T + T^*)}_{\text{real part}} + i\,\underbrace{\frac{1}{2i}(T - T^*)}_{\text{imaginary part}},$$

with the real/imaginary part being self-adjoint. This mimics the decomposition of a complex number. The key difference is that the real/imaginary parts of T are operators, so they are not guaranteed to commute with each other. In fact, they commute iff T is normal (Exercise).

Exercise 6.7. Show that an invertible $T \in \mathcal{B}(\mathcal{H})$ preserves inner products iff it is unitary. (Thus unitaries are the automorphisms of Hilbert spaces.)

Exercise 6.8. Let $T \in \mathcal{B}(\mathcal{H})$. Check that

$$\sigma(T^*) = \sigma(T)^* := \{\bar{\lambda} \,:\, \lambda \in \sigma(T)\}. \tag{6.5}$$

Deduce that if T is unitary, $\sigma(T)$ is a subset of the unit circle.

Normal operators are well-behaved from the viewpoint of spectral theory. At the level of eigenvalues, there is the following easy result:

Exercise 6.9. For a bounded normal operator, show that the eigenspaces for distinct eigenvalues are orthogonal to each other.

Next, consider the following generalization of eigenvalue:

Definition 19. Let $T \in \mathcal{B}(\mathcal{H})$. We say that $\lambda \in \mathbb{C}$ is an *approximate eigenvalue* of T if for each $\epsilon > 0$, there exists a unit vector $\psi \in \mathcal{H}$ such that $||(\lambda - T)\psi|| \leq \epsilon$. The set of approximate eigenvalues of T is its *approximate point spectrum*, denoted $\sigma_{\mathrm{ap}}(T)$.

Exercise 6.10. Check that $\sigma_{\mathrm{p}}(T) \subset \sigma_{\mathrm{ap}}(T) \subset \sigma(T)$.

Theorem 6.8. *Let $T \in \mathcal{B}(\mathcal{H})$ be a normal operator. Then $\sigma(T) = \sigma_{\mathrm{ap}}(T)$.*

Proof. The inclusion $\sigma_{\mathrm{ap}}(T) \subset \sigma(T)$ is Exercise 6.10.

For the reverse inclusion, suppose $\lambda \notin \sigma_{\mathrm{ap}}(T)$. So there exists $\alpha > 0$ such that $||(\lambda - T)\psi|| \geq \alpha||\psi||$ for all $\psi \in \mathcal{H}$, i.e., $\lambda - T$ is bounded below. By Prop. 6.3, it suffices to show that $\lambda - T$ has dense range, to deduce that it is invertible, thus $\lambda \in \sigma(T)$.

To that end, suppose $\varphi \in (\mathrm{Ran}(\lambda - T))^{\perp}$. Thus

$$0 = \langle \varphi | (\lambda - T)\psi \rangle = \langle (\bar{\lambda} - T^*)\varphi | \psi \rangle, \qquad \forall \psi \in \mathcal{H},$$

so $(\bar{\lambda} - T^*)\varphi = 0$. Now $\lambda - T$ is normal, so Prop. 6.7 says that

$$||(\bar{\lambda} - T^*)\varphi|| = ||(\lambda - T)\varphi|| \geq \alpha||\varphi||$$

also holds. We learn that φ has to be zero. Thus $(\mathrm{Ran}(\lambda - T))^{\perp} = \{0\}$, which is the dense range condition. $\square$

For non-normal operators, Theorem 6.8 is not true; the spectrum will generally include some rather mysterious elements.

Exercise 6.11. On $\ell^2 = \ell^2(\mathbb{N})$, consider the *unilateral* shift,

$$S : (x_1, x_2, x_3 \dots) \mapsto (0, x_1, x_2, \dots).$$

- What is the adjoint operator S^*?
- Show that $S^*S = 1$. What is SS^*? Is S normal?
- Determine the operator norm of S.
- Does S have any eigenvalues?

Let $\lambda \in \mathbb{C}$ with $|\lambda| < 1$.

- Show that λ is an eigenvalue of S^*.
- Show that $\lambda - S$ does not have dense range. Is λ an approximate eigenvalue of S?

6.2.2 Bounded self-adjoint operators and quantum observables

It follows from (6.5) that the spectrum of a self-adjoint operator is symmetric under reflection in the real axis. Actually, much more is true!

Proposition 6.9. *$T \in \mathcal{B}(\mathcal{H})$ is self-adjoint iff $\langle \psi | T\psi \rangle \in \mathbb{R}$ for all $\psi \in \mathcal{H}$.*

Proof. Let f_T be the sesquilinear form for T (Lemma 6.6). It determines a quadratic form

$$\hat{f}_T : \psi \mapsto f_T(\psi, \psi) = \langle \psi | T\psi \rangle.$$

Similarly, for T^*, we have the sesquilinear form f_{T^*} with associated quadratic form

$$\hat{f}_{T^*} : \psi \mapsto f_{T^*}(\psi, \psi) = \langle \psi | T^*\psi \rangle = \langle T\psi | \psi \rangle.$$

So $\hat{f}_{T^*} = \hat{f}_T$ iff $\langle \psi | T\psi \rangle$ is real-valued for all $\psi \in \mathcal{H}$.

Now, the polarization identity (Eq. (5.1)) is actually a general procedure for recovering a sesquilinear form from its associated quadratic form. Thus $T = T^*$ iff $f_T = f_{T^*}$ iff $\langle \psi | T\psi \rangle$ is real-valued for all $\psi \in \mathcal{H}$. □

In quantum mechanics,

$$\langle\psi|T\psi\rangle, \qquad ||\psi|| = 1$$

is the expectation value of the observable modelled by T, in the normalized state $\psi \in \mathcal{H}$. These expectation values must always be real-valued, so Prop. 6.9 gives an initial reason for modelling observables by self-adjoint operators. Why this actually works well is due to an *orthogonal spectral decomposition* property of self-adjoint operators (Spectral theorem). Let us illustrate this in the finite-dimensional case.

In linear algebra, we learn that an $n \times n$ self-adjoint matrix $T = T^*$ can be unitarily diagonalized, with real eigenvalues. The eigenspaces for distinct eigenvalues are mutually orthogonal. For simplicity, assume non-degenerate eigenvalues. Thus T has an orthonormal basis of eigenvectors e_i, with respective eigenvalues λ_i. In ket-bra notation, the orthogonal projection onto the i-th eigenspace is $|e_i\rangle\langle e_i|$, and we may write

$$T = \sum_{i=1}^{n} \lambda_i |e_i\rangle\langle e_i|.$$

Given a normalized $\psi \in \mathbb{C}^n$ ("state"), the expectation value of T is thus

$$\langle\psi|T\psi\rangle = \sum_{i=1}^{n} \lambda_i \langle\psi|e_i\rangle\langle e_i|\psi\rangle = \sum_{i=1}^{n} \lambda_i |\langle e_i|\psi\rangle|^2.$$

The real number $|\langle e_i|\psi\rangle|^2$ is the probability that the outcome λ_i occurs when measuring the observable T for the state ψ. Note that the sum of these probabilities over $i = 1, \ldots, n$ is 1, as required.

Each single measurement produces only one outcome, and expectation values are actually averaged over many repeated measurements of the identically prepared state ψ. So the observable T must encode a characteristic set of possible outcomes (via its real eigenvalues λ_i), with each outcome associated with a projection $|e_i\rangle\langle e_i|$, mutually orthogonal for distinct outcomes.

The last sentence is a demand that T is unitarily diagonalizable with real eigenvalues. Unitary diagonalizability characterizes the normal operators on $\mathbb{C}^n$. (This was probably covered in linear algebra.) The requirement that the eigenvalues are real, then restricts T to be self-adjoint.

Returning to the general infinite-dimensional Hilbert space setting, we see that T should be unitarily diagonalizable (this will be made precise), and have real spectrum. Let us check that self-adjoint $T = T^*$ indeed possess the second property.

Theorem 6.10. *If $T \in \mathcal{B}(\mathcal{H})$ is self-adjoint, then $\sigma(T) \subset \mathbb{R}$.*

Proof. Let $\lambda \in \mathbb{C} \setminus \mathbb{R}$. For any unit vector $\psi \in \mathcal{H}$,

$$\begin{aligned} 0 < |\lambda - \bar{\lambda}| = |\lambda - \bar{\lambda}| \cdot ||\psi||^2 &= \big|\langle \psi | (\lambda - \bar{\lambda}\psi\rangle\big| \\ &= \big|\langle \psi | (\bar{\lambda} - T)\psi\rangle - \langle \psi | (\lambda - T)\psi\rangle\big| \\ &= \big|\langle (\lambda - T)\psi | \psi\rangle - \langle \psi | (\lambda - T)\psi\rangle\big| \\ &\leq 2||(\lambda - T)\psi|| \cdot ||\psi|| \\ &= 2||(\lambda - T)\psi||. \end{aligned}$$

Thus λ is not an approximate eigenvalue of T. But T is normal, so Theorem 6.8 says that λ is not in the spectrum of T. $\square$

The following is called a *spectral radius formula.*

Proposition 6.11. *If $T \in \mathcal{B}(\mathcal{H})$ is self-adjoint, then*

$$||T|| = \sup_{\lambda \in \sigma(T)} |\lambda|.$$

Proof. By definition of operator norm, there is a sequence of unit vectors ψ_n such that $||T\psi_n|| \to ||T||$. For these unit vectors,

$$\begin{aligned} ||(||T||^2 - T^2)\psi_n||^2 &= ||T||^4 - 2||T||^2||T\psi_n||^2 + ||T^2\psi_n||^2 \\ &\leq ||T||^4 - 2||T||^2||T\psi_n||^2 + ||T||^2||T\psi_n||^2 \\ &= ||T||^2(||T||^2 - ||T\psi_n||^2) \to 0. \end{aligned}$$

holds. This shows that $||T||^2 \in \sigma_{\mathrm{ap}}(T^2) = \sigma(T^2)$. (Note that T^2 is self-adjoint so Theorem 6.8 applies). Now, we also saw from an exercise that

$$\sigma(T^2) = \sigma(T)^2.$$

So $||T||$ and/or $-||T||$ belong to $\sigma(T)$. Thus $\sup_{\lambda \in \sigma(T)} |\lambda| \geq ||T||$. The reverse inequality is a general result proved in Prop. 6.1. $\square$