What is a Hilbert space?

Liam Llamazares

05/30/2023

1 Three line summary

- Hilbert spaces are Banach spaces with an inner product. They correspond to an infinite dimensional version of $\mathbb{R}^n$ or $\mathbb{C}^n$.
- Much of Euclidean geometry carries over to Hilbert spaces including orthogonality, projections, orthonormal basis, and Pythagoras' theorem.
- Compact operators are a simple kind of operator that can be approximated by infinite dimensional ones.

2 Why should I care?

Hilbert spaces come with beautiful geometry and give us tools that are not available in ordinary Banach spaces. These allow us to prove the existence of infinite dimensional equations and even pose them as minimization problems. When you can use a Hilbert space to analyze your problem your life is always a lot easier.

3 Notation

Given $\lambda \in \mathbb{C}$ we will write $\overline{\lambda}$ for the conjugate of λ.
Given a subset A of some topological space, it is also common to write $\overline{A}$ for the closure of A. Though this is a slight abuse of notation we will do the same as the meaning will always be clear from context.

Given two topological vector spaces X, Y we write $\mathcal{L}(X, Y)$ for the space of continuous linear operators from X to Y and X' for the space of continuous linear functions to the base field of X.

4 Introduction

In writing the next post in the series of PDEs I realized that there were many results from functional analysis which had not been discussed and would be necessary. As such, this post is meant as a short, yet rather longer than I intended, introduction to the basic theory that will be required.
I'm going to assume you know nothing (please don't feel insulted) and go over the main ideas and results. Much of the material may be familiar so please feel free to skim briefly through the results or skip entirely. There is a lot of theory to cover so the proofs will only be sketched. For more details see for example [1].

5 A brief recap of Hilbert spaces

Firstly what is a Hilbert space? To answer that question let's start with what kind of spaces we want to model.

Example 1. The Euclidean space $\mathbb{R}^n$ together with the inner product

$$x \cdot y := x_1 y_1 + \dots x_n y_n.$$

Example 2. The complex Euclidean space $\mathbb{C}^n$ together with the inner product

$$\langle x, y \rangle := x_1 \overline{y_1} + \dots x_n \overline{y_n}.$$

You might be wondering why we took the conjugate in the definition of the inner product above. That is simply because we want the inner product to define a norm by

$$\|x\|^2 := \langle x, x \rangle .$$

In the cases above note that this is the standard Euclidean norm

$$\|x\|^2 = \langle x, x \rangle = |x_1|^2 + \dots + |x_n|^2 .$$

A Hilbert space is just this construction generalized to infinite dimensions. Firstly, the inner product is generalized as follows.

Definition 1. Given a vector space V over $\mathbb{K}$ where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$, an inner product is a mapping

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbb{K}; \quad (x, y) \to \langle x, y \rangle .$$

Such that the following hold for all $x, y, z \in V$ and all $\lambda \in \mathbb{K}$

1. Linearity: $\langle \lambda x + y, z \rangle = \lambda \langle x, z \rangle + \langle y, z \rangle$.
2. Conjugate symmetry $\langle x, y \rangle = \overline{\langle y, x \rangle}$.
3. Positivity $\langle x, x \rangle = 0$ if and only if $x = 0$.

Definition 2.

In what follows we will take always take the base field to be $\mathbb{K}$, that is, we will consider real and complex vector spaces as the theory is the same. In the real case, the conjugate function just being the identity. Note that the first two properties also imply "antilinearity" of the second component as

$$\langle z, \lambda x + y \rangle = \overline{\langle \lambda x + y, z \rangle} = \overline{\lambda} \langle z, x \rangle + \lambda_2 \langle z, y \rangle .$$

A vector space with an inner product is called an inner product space. This inner product gives us a concept of geometry.

Definition 3. Given an inner product space V we define

$$\|x\| := \sqrt{\langle x, x \rangle}.$$

Furthermore, we say that x, y are orthogonal if $\langle x, y \rangle = 0$.

With this definition, Pythagoras's theorem follows directly, as if x, y are orthogonal then by the linearity of the inner product.

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \|x\|^2 + \|y\|^2.$$

An important property of the inner product is the following

Proposition 1 (Cauchy Schwartz). Let V be an inner product space, then it holds that

$$\langle x, y \rangle \leq \|x\| \|y\|.$$

In particular, this implies that the inner product is continuous jointly, and in each component. An inner product space keeps the key properties of Euclidean space, namely the vector space structure and inner product. The third key property is that of completeness.

Definition 4. A Hilbert space is a complete inner product space $(H, \langle\cdot,\cdot\rangle)$. Where completeness is respect to the inner product norm

$$\|x\|^2 := \langle x, x\rangle .$$

Cool! Now we know what is a Hilbert space. Some examples that have come up in previous posts are the space of square-integrable functions valued in another Hilbert space H with the inner product

$$\langle f, g\rangle_{L^2(\Omega\to H)} := \int_\Omega \langle f, g\rangle \, d\mu.$$

Another are the Sobolev spaces $H^s(\mathbb{R}^d)$ with the inner product

$$\langle f, g\rangle_{H^s(\mathbb{R}^d)} := \int_{\mathbb{R}^d} \langle\omega\rangle^{2s} \widehat{f}(\xi)\overline{\widehat{g}(\xi)} d\xi.$$

(I apologize for using the same notation for the Japanese bracket and the inner product). One result that carries over from Euclidean spaces to Hilbert spaces is the existence of projections.

Theorem 1 (Projection theorem). Let F be a closed convex subset of a Hilbert space H. Then, given $h \in H$ there exists a unique $f \in F$ such that

$$\min_{g\in F}\|h-g\| = \|h-f\|.$$

Furthermore, if F is a (vector) subspace, then it holds that f is the unique element in F such that $h-f$ is orthogonal to F. That is $f \in F$ is the only one verifying

$$h - f \in F^\perp := \{g \in H : \langle g, F\rangle = \{0\}\} .$$

Proof. The idea of the proof is to consider a sequence f_n that approaches the minimum distance

$$d(f_n, h) \to d(h, F).$$

The inner product comes in to show that f_n is a Cauchy sequence. Since H is complete it converges to some f. Then our assumptions on F come in. Firstly, since F is closed shows $f \in F$. Secondly, since F is convex one can get uniqueness. The uniqueness is once more a consequence of the properties of the inner product. □

A consequence of the above theorem is that orthogonal complements in Hilbert spaces exist.

Theorem 2 (Orthogonal complement). Let V be a closed subspace of a vector space H. Then we have that

$$\begin{aligned} V \oplus V^{\perp} &\longrightarrow H \\ (v_1, v_2) &\longmapsto v_1 + v_2. \end{aligned}$$

Is a bijective isomorphism where we consider on $V \oplus V^{\perp}$ the inner product

$$\langle (v_1, v_2), (v_1', v_2') \rangle = \langle v_1, v_1' \rangle + \langle v_2, v_2' \rangle .$$

Proof. Given $h \in H$ we know that there exists a unique closest point $v_1 \in V$ to V (the projection of h onto v). Furthermore, $v_2 := h - v \in V^{\perp}$ so we get the decomposition

$$h = v_1 + v_2.$$

This proves bijectivity, whereas the isomorphism property follows quickly from the orthogonality of v_1, v_2. □

An important corollary of this is the following

Corollary 1. Let V be a vector space, then

$$\overline{V} = H \iff \overline{V}^{\perp} = \{0\} .$$

Note that the imposition that V is closed in Theorem 2 is necessary. Finite-dimensional subspaces are always closed, but infinite-dimensional subspaces may not be. For example, consider $H = L^2(I)$ for some bounded I and take V to polynomials on I . By the Stone-Weierstrass theorem V is dense in H. That is $H = \overline{V}$, whereas

$$H \neq V \oplus V^{\perp} = V \oplus \{0\} .$$

Another property of Euclidean space that Hilbert spaces reproduce is that of the existence of an orthonormal basis

Definition 5. Let H be a Hilbert space and I some index set. We say that $\mathcal{B} = \{\phi_\alpha\}_{\alpha \in I}$ is an orthonormal basis of H if

$$\langle \phi_\alpha, \phi_\beta \rangle = \delta_{\alpha,\beta}, \quad \forall \alpha, \beta \in I.$$

And for every element $f \in H$ there exist $\lambda_\alpha \in \mathbb{C}$ such that

$$f = \sum_{\alpha \in I} \lambda_\alpha \phi_\alpha. \tag{1}$$

Note that we impose no conditions on the Index set I which may be countable or uncountable.

Definition 6. Given an index set I and a normed vector space X we say that x_α is absolutely summable to x and write

$$x = \sum_{\alpha \in I} x_\alpha$$

if given $\epsilon > 0$ there exists some finite subset $J_0 \in I$ such that for every J containing J_0 it holds that

$$\|x - \sum_{\alpha \notin J} x_\alpha\| < \epsilon.$$

If I is countable, the definition says that $\sum_{\alpha \in I} x_\alpha$ converges to x regardless of the order in which we sum. In fact, the following shows that, even if we start with an uncountable sequence we will always end up back in this case.

Proposition 2. Let X be a normed space and $\{x_\alpha\}_{\alpha \in I} \subset X$ be absolutely summable to x, then only a countable number of the terms x_α are non-zero. Let us take the nonzero terms and relabel them $\{x_n\}_{n \in \mathbb{N}}$, then

$$x = \sum_{n=0}^{\infty} x_n.$$

Proof. Take J_n such that

$$\|x - \sum_{\alpha \notin J_n} x_\alpha\| < \frac{1}{n}.$$

Then $J := \cup_n J_n$ is countable (it is a countable union of countable sets) and a small reasoning shows that J are the non-zero terms of I. The fact that the nonzero x_n sum to x is just a consequence of the definition of absolute summability. $\square$

Now that we made sense of the sum over a potentially uncountable number of basis elements in (1), it remains to address the question of the existence of orthonormal basis.

Theorem 3 (Existence of orthonormal basis). Every Hilbert space has an orthonormal basis $\mathcal{B}$.

Proof. The proof is formally identical to the proof that every vector space has a basis space . Let $\mathcal{A}$ be the collection of all orthonormal subsets of H. That is, $\mathcal{A}$ is comprised of sets of the form

$$\mathcal{S} = \left\{ \{\phi_\alpha\}_{\alpha \in J} : \langle \phi_\alpha, \phi_\beta \rangle = \delta_{\alpha,\beta}, \quad \forall \alpha, \beta \in J \right\}.$$

Such that ϕ_α are orthonormal. Given an ordered chain $\mathcal{S}_0 \subset \mathcal{S}_1 \subset \mathcal{S}_2 \subset \cdots$ we have the bound

$$\mathcal{S}_0 \subset \mathcal{S}_1 \subset \mathcal{S}_2 \subset \cdots \subset \bigcup_{n=0}^{\infty} \mathcal{S}_n.$$

As a result by Zorn's lemma there exists a maximal element $\mathcal{B}$. If $\mathcal{B}$ is not complete (that is (1) doesn't hold), then there exists $f \in \overline{\mathcal{B}}^{\perp}$. By taking $f/\|f\|$ and forming $\mathcal{B}' := \mathcal{B} \cup \{f\|f\|\}$ we obtain that $\mathcal{B} \subsetneq \mathcal{B}' \in \mathcal{A}$. This contradicts the maximality of $\mathcal{B}$ and concludes the proof. □

The next result is the natural generalization of Pythagoras's theorem to Hilbert spaces.

Theorem 4 (Parseval). Let H be a Hilbert space with orthonormal basis $\mathcal{B} = \{\phi_\alpha\}_{\alpha \in I}$. Then for every $f \in H$ it holds that

$$f = \sum_{\alpha \in I} \langle f, \phi_\alpha \rangle \phi_\alpha; \quad \|f\|^2 = \sum_{\alpha \in I} \|\phi_\alpha\|^2.$$

Proof. We have that by the orthonormality of ϕ_α and the continuity of the inner product

$$\left\langle f - \sum_{\alpha \in I} \langle f, \phi_\alpha \rangle \phi_\alpha, \phi_\alpha \right\rangle = \langle f, \phi_\alpha \rangle - \langle f, \phi_\alpha \rangle = 0.$$

As a result, $f - \sum_{\alpha \in I} \langle f, \phi_\alpha \rangle \phi_\alpha$ is orthogonal to the closure of the span of $\mathcal{B}$, which by assumption is H. By Corollary 1 we conclude that

$$f - \sum_{\alpha \in I} \langle f, \phi_\alpha \rangle \phi_\alpha = 0.$$

This proves the first part of the theorem. The second follows by the first together with the orthonormality of ϕ_α. □

The above shows that, on fixing a basis, every Hilbert space can be identified with a space of square-integrable sequences by the bijective isometry

$$H \to \ell^2(I); \quad f \to \{\langle f, \phi_\alpha \rangle\}_{\alpha \in I}.$$

In particular, every Hilbert space with a countable basis is isometric to $\ell^2(\mathbb{N})$. However, the identification is not "canonical" as it depends on the bases chosen. The next example is all pervasive (in fact it has even invaded our blog)

Example 3 (Plancherel). The Hilbert space of square integrable complex valued periodic functions $L^2(\mathbb{R}^d/\mathbb{Z}^d \to \mathbb{C})$ has orthonormal basis

$$\{\phi_k(x)\}_{k \in \mathbb{Z}^d} := \left\{e^{2\pi i k \cdot x}\right\}_{k \in \mathbb{Z}^d}.$$

Thus, every function $f \in L^2(\mathbb{R}^d/\mathbb{Z}^d)$ can be written as

$$f(x) = \sum_{k \in \mathbb{Z}^d} \widehat{f}(k) e^{2\pi i k \cdot x}.$$

Where $\widehat{f} \in \ell^2(\mathbb{Z}^d)$ is known as the Fourier transform of f and defined by

$$\widehat{f}(k) := \langle f, \phi_k \rangle = \int_{\mathbb{R}^d/\mathbb{Z}^d} f(x) e^{-2\pi i k \cdot x} \, \mathrm{d}x.$$

Another interesting property of Euclidean space is that every element of the dual $\ell : \mathbb{C}^n \to \mathbb{C}$ is represented by a vector in the space, that is

$$\ell(x) = \langle x, y_\ell \rangle, \quad \forall x \in \mathbb{C}^n.$$

Here one can calculate directly that y_ℓ is the conjugate of the "matrix" defined by ℓ as a linear function

$$y_\ell = (\overline{\ell(e_1)}, \ldots, \overline{\ell(e_n)}).$$

Where e_i is the standard orthonormal basis of $\mathbb{C}^n$. In Hilbert spaces, the same result holds for Hilbert spaces

Theorem 5 (Riesz representation). Let H be a Hilbert space, then given $\ell \in H'$ there exists a unique $f_\ell \in H$ such that

$$\langle h, f_\ell \rangle, \quad \forall h \in H. \tag{2}$$

Furthermore, $\|f_\ell\| = \|\ell\|$.

Proof. Consider an orthonormal basis $\{\phi_\alpha\}_{\alpha \in I}$ for H. Then, just as in Euclidean space we have that

$$f_\ell = \sum_{\alpha \in I} \overline{\ell(\phi_\alpha)} \phi_\alpha.$$

The fact that f_ℓ verifies (2) is a direct application of the (anti)-linearity and continuity of the inner product. Uniqueness following from the fact that if f_ℓ, g_ℓ both verify the equality then for all $h \in h$

$$\langle h, f_\ell - g_\ell \rangle = \ell(h) - \ell(h) = 0.$$

This can only occur if $f_\ell - g_\ell = 0$ (hint take $h = f_\ell - g_\ell$). To verify the norm we can use that, for all $h \in H$ with norm 1

$$\ell\left(\frac{f_\ell}{\|f_\ell\|}\right) = \left\langle \frac{f_\ell}{\|f_\ell\|}, f_\ell \right\rangle = \|f_\ell\|; \quad \ell(h) = \langle h, f_\ell \rangle \leq \|h\| \|f_\ell\| = \|f_\ell\|.$$

The equality shows that $\|\ell\| \geq \|f_\ell\|$ whereas the inequality shows the converse $\|\ell\| \leq \|f_\ell\|$, proving the theorem. □

The previous theorem says that we have an antilinear isometry

$$\Phi_1 : H \to H'; \quad f_\ell \to \ell.$$

This allows us to identify H with H'. The identification is canonical as it does not depend on the bases chosen. Yes, we fixed a basis to prove it but the vector f_ℓ is unique independently of the basis. This allows us to make H' into a Hilbert space in a canonical way

Proposition 3. Let H be a Hilbert space, then H' is also a Hilbert space, with inner product given by

$$\langle \ell_1, \ell_2 \rangle_{H'} = \langle f_{\ell_2}, f_{\ell_1} \rangle_H .$$

Where we had to “swap the order” of the representatives of ℓ_1, ℓ_2 due to the anti-linearity of the mapping $\ell \to f_\ell$. Since H' is also a Hilbert space we can apply Riesz's theorem to H' to show that H'' is also a Hilbert space and there exists a canonical antilinear isometry

$$\Phi_2 : H' \to H''; \ell_\varphi \to \varphi.$$

By construction, of Φ_1, Φ_2 it holds that

$$\Phi_2 \circ \Phi_1(f)(\ell) = \ell(f).$$

That is, H is identified canonically with H'' and, the identification is such that

$$f(\ell) = \ell(f).$$

In other words.

Theorem 6. Every Hilbert space is reflexive.

Consider now a Hilbert space H and a linear operator $T : H \to H$. Then, for each $g \in H$ we can define the linear form

$$\ell_g := \langle T\cdot, g\rangle . \tag{3}$$

As a result, there exists a unique representative of ℓ_g in H which, to track the dependence on g, we denote by h_g. That is, h_g verifies

$$\langle Tf, g\rangle = \langle f, h_g\rangle , \quad \forall f \in H. \tag{4}$$

A small reasoning shows that h_g is a linear function of g, that is, there exists $T^* : H \to H$ with $T^* g = h_g$, or in other words

$$\langle Tf, g\rangle = \langle f, T^* g\rangle , \quad \forall f, g \in H. \tag{5}$$

Definition 7. Given $T \in \mathcal{L}(H)$ we denote by T^* the unique element verifying (5) . If $T = T^*$ we say that T is self adjoint.

If we think in terms of complex numbers, the adjoint of an element $\lambda \in \mathbb{C}$ is $\overline{\lambda}$ and T is self adjoint if and only if it is real. If now we consider the case where H is $\mathbb{C}^n$ and T is given by some matrix A then $A^* = A^\dagger := \overline{A^T}$. As in these finite dimensional cases, the following proof is not difficult.

Proposition 4. Let $T \in \mathcal{L}(H)$ with H a Hilbert space. Then $T^* \in \mathcal{L}(H)$ with $\|T\| = \|T^*\|$.

Hilbert spaces provide us a way to guarantee existence and uniqueness to a wide class of problems, an important tool is Lax-Milgram's theorem. We first need two definitions

Definition 8. We say that a mapping

$$B : V \times V \to \mathbb{K}$$

on a vector space V, is sesquilinear if B is linear in the first component and antilinear in the second. That is, for all $x, y, z \in V$ and $\lambda \in \mathbb{K}$:

$$B(\lambda x + y, z) = \lambda B(x, z) + B(y, z); \quad B(x, \lambda y + z) = \overline{\lambda} B(x, z) + B(y, z).$$

Definition 9. Let V be a normed vector space then we say that a sesquilinear form B is α coercive if it is continuous and there exists a constant $\alpha > 0$ such that

$$B(f, f) \geq \alpha \|f\|^2 \quad \forall f \in H.$$

The coercivity condition essentially imposes that the bilinear form is not degenerate. As a particular example, a symmetric sesquilinear form is an inner product.

Theorem 7 (Lax Milgram). Let B, L be respectively an α coercive sesquilinear form and a linear form on a Hilbert space H. Then there exists an invertible linear operator $\mathcal{L} : H \to H$ and $f \in H$ such that

$$B(v, u) = \langle v, \mathcal{L}u \rangle ; \quad L(v) = \langle v, f \rangle .$$

As a result, equation

$$B(v, u) = L(v) \quad \forall v \in H \tag{6}$$

has a unique solution $u = \mathcal{L}^{-1} f$. Furthermore, the solution operator $\mathcal{L}^{-1}$ is continuous with

$$\|\mathcal{L}^{-1}\| \leq \alpha^{-1}.$$

Proof. For each fixed $u \in H$, we have that $\ell_u := B(\cdot, u) \in H'$. As a result by Riesz's representation theorem (Theorem 5) there exists a unique $f_{\ell_u} \in H$ such that

$$B(v, u) = \ell_u(v) = \langle v, f_{\ell_u} \rangle . \tag{7}$$

Furthermore, it can be simply verified that the mapping $u \to f_{\ell_u}$ is linear in u. That is, there exists $\mathcal{L} : H \to H$ such that

$$\mathcal{L}u = f_{\ell_u} \quad \forall u \in H. \tag{8}$$

The existence of the representative $f \in H$ of L is once more by Riesz's representation theorem. We now show that $\mathcal{L}$ verifies the desired properties. Firstly $\mathcal{L}$ is continuous as, given $u \in H$

$$\|\mathcal{L}u\|^2 = \langle \mathcal{L}u, \mathcal{L}u \rangle = B(\mathcal{L}u, u) \leq \|B\|\|u\|\|\mathcal{L}u\|.$$

So dividing by $\|\mathcal{L}u\|$ on either side shows that $\|\mathcal{L}\| \leq \|B\|$. Now, $\mathcal{L}$ is injective as if $\mathcal{L}u = 0$ then

$$0 = \langle u, \mathcal{L}u \rangle = B(u, u) \geq \alpha\|u\|^2.$$

We now prove surjectivity of $\mathcal{L}$. Consider $u \in \mathrm{Im}(\mathcal{L})^\perp$, then it holds that

$$\langle u, \mathcal{L}u \rangle = B(u, u) \geq \alpha\|u\|^2. \tag{9}$$

As a result, we deduce from the corollary of the orthogonal complement theorem 1 that $\overline{\mathrm{Im}(\mathcal{L})} = 0$. Thus, if we show that $\mathcal{L}$ is closed invertibility follows. The estimate in (9) together with Cauchy Schwartz show that for all $u \in H$

$$\|\mathcal{L}u\| \geq \|u\|.$$

As if $\mathcal{L}u_n \in \mathrm{Im}(\mathcal{L})$ is a Cauchy sequence then so must be u_n. By completeness of $\mathcal{L}$, the sequence u_n converges to some $u \in H$ and we deduce, by continuity of $\mathcal{L}$, that $\mathcal{L}u_n \to \mathcal{L}u \in \mathrm{Im}(\mathcal{L})$. In consequence, $\mathcal{L}$ is invertible, finally to show the bound on $\mathcal{L}^{-1}$ let us write $u = \mathcal{L}^{-1}f$ then

$$\alpha\|u\|^2 \leq B(u, u) = \langle u, f \rangle \leq \|u\|\|f\|.$$

Dividing on either side by $\alpha\|u\|$ concludes the proof. □

If we had assumed that B were anti-symmetric then the proof would have been simplified as B would define an inner product $\langle \cdot, \cdot \rangle_B$. Applying Riesz's theorem to this inner product (as opposed to the original one) would transform our equation (6) into

$$\langle v, u \rangle_B = \langle v, f \rangle , \quad \forall v \in H.$$

That is, to solve (6) it would suffice to take $u = f$. In the case where B is symmetric and real, we can also find u by solving a minimization problem

Proposition 5 (Minimization formulation). Let $B : H \times H \to \mathbb{R}$ be a symmetric coercive bilinear operator on a real Hilbert space H. Then, problem (6) is equivalent to minimizing

$$J(u) := \frac{1}{2} B(u, u) - L(u).$$

Proof. To prove that a solution to (6) minimizes J we can develop $J(u + v)$ and simplify it using $B(v, u) = L(v)$ to obtain

$$J(u + v) \geq J(u), \quad \forall v \in H.$$

To prove that a minimum of J solves (6) one can show by taking limits as $\lambda \to 0$ in the expression

$$J(u) \leq J(u + \lambda(h - u)).$$

That for all h

$$L(h - u) \leq B(u, h - u).$$

Taking $h = u + v$ and $h = u - v$ where v is any shows $-L(v) \leq -B(u, v)$ and $L(v) \leq B(u, v)$, which concludes the proof. $\square$

6 A little bit of operator theory

Finally, we wrap up with some operator theory. This will be revisited in a more detailed blog post on spectral theory. For now, we give the essentials.

Definition 10. We say that a linear operator $K : X \to Y$ where X, Y are two metric spaces is compact if $T(B)$ is relatively compact for all bounded $B \subset X$.

The above will be abbreviated $K \in \mathcal{K}(X,Y)$. Note that, since every compact set is bounded K must be bounded and thus continuous. That is

$$\mathcal{K}(X,Y) \subset \mathcal{L}(X,Y).$$

In practice, the following equivalent characterizations are useful.

Proposition 6. Let X, Y be two metric spaces, then the following are equivalent

a) $K \in \mathcal{K}(X,Y)$.

b) $K(B_X)$ is relatively compact where $B_X = \{x \in X : \|x\| < 1\}$ is the unit ball in X.

c) From every sequence Kx_n where $x_n \in B_X$ one can extract a subsequence Kx_{n_j} converging in Y.

Proof. The prove the first two points are equivalent we observe that every bounded set $B \subset X$ is contained in λB_X for λ big enough. A general fact from topology is that closed subsets of compact sets are compact. This is enough to conclude the equivalence.
To prove that the last two points are equivalent we recall that in metric spaces compact is equvalent to sequentially compact. $\square$

An important property of compact operators is that they are preserved by continuous ones.

Proposition 7. Let $K \in \mathcal{K}(X,Y), T_1 \in \mathcal{L}(W,X), T_2 \in L(Y,Z)$. Where W, X, Y, Z are metric spaces , then $K \circ T_1$ and $T_2 \circ K$ are compact

Proof. This follows by the last equivalent characterization in Proposition 6. $\square$

We already saw that the space of compact operators $\mathcal{K}(X,Y)$ is a subset of the space of linear operators. The structure of $\mathcal{K}(X,Y)$ as a subspace is as follows

Proposition 8. Let X, Y be metric spaces then $\mathcal{K}(X,Y)$ is a vector space. Furthermore, if Y is completee then $\mathcal{K}(X,Y)$ is closed in $\mathcal{L}(X,Y)$. That is, if K_n are compact and $K_n \to K \in \mathcal{L}(X,Y)$ then K is compact.

Proof. The first part again follows fromb the last equivalent characterization in Proposition 6. The second part relies on the fact that in complete spaces compact and totally bounded are equivalent. □

The reason we are interested in compact operators is that they are particularly simple. Note that every finite-dimensional operator (that is operators whose image is finite-dimensional) is compact by the Heine-Borel theorem. In fact, a good way of thinking of compact operators is to see them as finite-dimensional operators . Or more precisely, as the limit of them

Proposition 9. Let $K \in \mathcal{K}(X, H)$ where H is a separable Hilbert space . Then there exists a sequence of finite-dimensional operators K_n such that

$$\lim_{n\to\infty} K_n = K.$$

Proof. Since H is separable there exists a countable orthonormal basis $\{\phi_n\}_{n\in\mathbb{N}}$. If we denote T_n for the projection of K onto the space generated by $\{\phi_1, \ldots, \phi_n\}$ then, by Parseval's Theorem pointwise convergence holds

$$K_n(x) \to K(x), \quad \forall x \in H.$$

Consider the unit ball B_X in X. Then $K(B_X)$ is relatively compact. So $K(B_X)$ is totally bounded and given $\epsilon > 0$ we can form a finite ϵ net $Kx_1, \ldots, Kx_m$ of $K(B_X)$. By pointwise convergence, we can now take n_0 large enough so that for all $N \geq n_0$

$$\|K(x_j) - K_N(x_j)\|, \quad \forall j = 1, \ldots, m.$$

Now for any $x \in B_X$ we can find x_j such that $\|Tx_j - Tx\| < \epsilon$. Using the triangle inequality

$$\|Kx - K_N x\| \leq \|Kx - Kx_j\| + \|Kx_j - K_N x_j\| + \|K_N x_j - K_N x\| < 3\epsilon.$$

This concludes the proof. □

For us, an important example of compact operators will be the solution operator $\mathcal{L}^{-1}$ of a PDE. This is because of the following theorem

Theorem 8 (Rellich-Kondrachov). Let $U \subset \mathbb{R}^n$ be a bounded open domain in $\mathbb{R}^n$ with smooth boundary. Then, given $s > 0$ it holds that the natural inclusion

$$i : H^{s+\sigma}(U) \hookrightarrow H^s(U)$$

is compact for all $\sigma > 0$.

For the proof of a more general version see [2] page 334. It is important to observe the restriction that U is bounded is necessary and the theorem no longer holds if U is replaced by $\mathbb{R}^n$.

Corollary 2 (Adding differentiability is compact). Let $U \subset \mathbb{R}^n$ be a bounded open domain in $\mathbb{R}^n$ with smooth boundary and $s, \sigma > 0$, then every continuous operator

$$T : H^{s+\sigma}(U) \hookrightarrow H^s(U)$$

is compact.

Proof. We have that $T = i \circ T$ so we conclude by Rellich-Kondrachov's theorem 8 and the preservation of compact operators by continuous ones (Proposition 7). □

Compact operators also have a nice spectral theory. Where we recall that the spectrum of an operator T is defined as

$$\sigma(T) := \{\lambda \in \mathbb{K} : \lambda \mathbf{Id} - T \text{ is not invertible }\}.$$

Firstly, the spectrum of a compact operator is equal to it's eigenvalues. Secondly the following holds

Theorem 9 (Spectral theorem). Let $K \in \mathcal{K}(H)$ be compact and self adjoint on a Hilbert space H. Then T diagonalizes in an orthonormal basis. That is, there exists an orthonormal basis $\{\phi_\alpha\}_{\alpha \in I}$ and $\lambda_\alpha \in \mathbb{K}$ such that

$$Kx = \left(\sum_{\alpha \in I}^{\infty} \lambda_\alpha \phi_\alpha \otimes \phi_\alpha \right) x = \sum_{\alpha \in I} \lambda_\alpha \langle x, \phi_\alpha \rangle \phi_\alpha.$$

A more general result is possible but requires more theory so we reserve it for another day. That said, this type of discrete representation of T is very useful and links up with the theory of trace class and Hilbert-Schmidt operators which we will discuss more in the future.
To end it all off we state without proof a theorem that will be useful in proving properties about the solution space to the solution of PDEs

Theorem 10 (Fredholm alternative). Let H be a Hilbert space and $K \in \mathcal{K}(H)$. Consider $T := Id - K$ then it holds that

$$T \text{ is injective} \iff T \text{ is surjective.}$$

Furthermore, it holds that

a) $\mathbf{ker}(T)$ is finite dimensional.

b) T is closed.

c) $\text{Im}(T) = \mathbf{ker}(T^*)^{\perp}$.

d) $\dim(\mathbf{ker}(T)) = \dim(\mathbf{ker}(T^*))$

We delay the proof till another day, in the meantime, see [3] page 725.

References

[1] G. Bachman, L. Narici, Functional analysis, Courier Corporation, 2000.
URL https://books.google.co.uk/books?hl=en&lr=&id=wCHtLumoGY4C&oi=fnd&pg=PR3&dq=narici+bachman&ots=vxU8cRul6s&sig=tGyszvXv-1psk-qyJ8ruM6S7l5I&redir_esc=y#v=onepage&q=narici%20bachman&f=false

[2] M. Taylor, Partial differential equations II: Qualitative studies of linear equations, Vol. 116, Springer Science & Business Media, 2013.

[3] L. C. Evans, Partial differential equations, Vol. 19, American Mathematical Society, 2022.
URL https://math24.files.wordpress.com/2013/02/partial-differential-equations-by-evans.pdf