2 Mathematical structure of quantum mechanics

2.1 States and observables

A complex Hilbert space is a complete inner product space over the field $\mathbb{C}$. The inner product $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ is

(i) linear in the second variable:

$$\langle \psi, \alpha_1 \phi_1 + \alpha_2 \phi_2 \rangle = \alpha_1 \langle \psi, \phi_1 \rangle + \alpha_2 \langle \psi, \phi_2 \rangle,$$

(ii) hermitian: $\langle \phi, \psi \rangle = \overline{\langle \psi, \phi \rangle}$,

(iii) positive definite: $\langle \psi, \psi \rangle \geq 0$ for all $\psi \in \mathcal{H}$, and $\langle \psi, \psi \rangle = 0 \iff \psi = 0$.

The inner product defines a norm

$$\|\psi\| = \langle \psi, \psi \rangle^{1/2}$$

with the following properties:

- Cauchy–Schwarz inequality:
$$|\langle \psi, \phi \rangle| \leq \|\psi\| \, \|\phi\|$$
- Triangle inequality:
$$\|\psi + \phi\| \leq \|\psi\| + \|\phi\|$$
- Polarization identity:
$$4\langle \psi, \phi \rangle = \|\psi + \phi\|^2 - \|\psi - \phi\|^2 - \mathrm{i}\|\psi + \mathrm{i}\phi\|^2 + \mathrm{i}\|\psi - \mathrm{i}\phi\|^2$$

In this course, $\mathcal{H}$ is separable. In particular, there is a countable orthonormal basis (ONB) and every ONB is countable, see exercises.

The set of pure states of a physical system is the projective space of $\mathcal{H}$:

$$\mathbb{P}(\mathcal{H}) = \big(\mathcal{H} \setminus \{0\}\big)/\!\sim$$

where $\psi \sim \phi$ if and only if $\psi = \lambda \phi$ for $\lambda \in \mathbb{C} \setminus \{0\}$. Equivalently,

$$\mathbb{P}(\mathcal{H}) \simeq \text{set of 1-dimensional subspaces of } \mathcal{H}.$$

One usually identifies $\mathbb{P}(\mathcal{H})$ with the set of unit vectors "up to phase", commonly referring to a (normalized) vector $\psi \in \mathcal{H}$ as a state. A useful identification is with 1-dimensional orthogonal projectors:

$$\mathbb{P}(\mathcal{H}) \ni [\psi] \longmapsto P_\psi = \frac{\langle \psi, \cdot \rangle}{\langle \psi, \psi \rangle} \, \psi.$$

Examples.

(i) $\mathcal{H} = \mathbb{C}^2$, a qubit, with

$$\langle \psi, \phi \rangle = \sum_{i=1}^{2} \overline{\psi_i} \, \phi_i.$$

(ii) $\mathcal{H} = \ell^2(\mathbb{N})$, the particle on a semi-infinite lattice, with

$$\langle \psi, \phi \rangle = \sum_{i=1}^{\infty} \overline{\psi_i}\, \phi_i.$$

(iii) $\mathcal{H} = L^2(\mathbb{R}^3; \mathbb{C})$, the particle in space, with

$$\langle \psi, \phi \rangle = \int_{\mathbb{R}^3;\mathbb{C}} \overline{\psi(x)}\, \phi(x) dx.$$

Physical observables (questions one can ask about the system) are represented by self-adjoint linear operators $A : D(A) \subset \mathcal{H} \to \mathcal{H}$.

Example. On $\mathcal{H} = L^2(\mathbb{R}^3; \mathbb{C})$, the observable "Is the particle in a region $\Omega \subset \mathbb{R}^3$?" is given by

$$(P_\Omega \psi)(x) = \chi_\Omega(x)\, \psi(x),$$

where χ_Ω is the characteristic function of the measurable set Ω.

Given an observable A and a state ψ (with $\|\psi\| = 1$), the expectation value of A is defined as

$$\langle \psi, A\psi \rangle \in \mathbb{C}.$$

For self-adjoint operators, the expectation value is real (see later). It is the average of the measurement outcomes of A in many identical, independent systems, all initially in the state ψ. The possible measurement outcomes are elements of the spectrum of A.

In the example:

- The eigenvalues of P_Ω are 0 and 1 (provided Ω and its complement have positive measure).
- The expectation value of P_Ω in ψ is

$$\langle \psi, P_\Omega \psi \rangle = \int_{\mathbb{R}^3} \overline{\psi(x)}\, \chi_\Omega(x)\, \psi(x) dx = \int_\Omega |\psi(x)|^2 dx. \tag{1}$$

- For any normalized eigenvector, $P_\Omega v = \lambda v$ with $\lambda \in \{0, 1\}$,

$$\langle v, P_\Omega v \rangle = \lambda,$$

  and, since $P_\Omega^2 = P_\Omega$ and $\lambda \in \{0, 1\}$, the variance vanishes,

$$\langle v, P_\Omega^2 v \rangle - \langle v, P_\Omega v \rangle^2 = \lambda - \lambda^2 = 0,$$

  namely: the measurement is deterministic.

Note that (1) is consistent with the interpretation that $|\psi(x)|^2$ is the probability density for the particle to be at $x \in \mathbb{R}^3$.

Another example. The position operator X on $L^2(\mathbb{R}; \mathbb{C})$:

$$(X\psi)(x) = x\, \psi(x), \qquad \mathrm{D}(X) = \left\{ \psi \in L^2(\mathbb{R}; \mathbb{C}) \;:\; x\psi \in L^2(\mathbb{R}; \mathbb{C}) \right\}.$$

Its expectation value is

$$\langle \psi, X\psi \rangle = \int_{\mathbb{R}} x\, |\psi(x)|^2 dx \qquad (\psi \in \mathrm{D}(X),\ \|\psi\| = 1).$$

As we shall see shortly, its spectrum is $\sigma(X) = \mathbb{R}$.

Another example. The "x component" of the spin:

$$S^x = \frac{1}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{on } \mathbb{C}^2.$$

The spectrum is made of the two eigenvalues $\sigma(S^x) = \{\pm\frac{1}{2}\}$ and the expectation value is given by

$$\langle \phi, S^x \phi \rangle = \tfrac{1}{2}\big(\overline{\phi_1}\, \phi_2 + \overline{\phi_2}\, \phi_1\big) \qquad (\|\phi\| = 1).$$

Just like much of probability theory is about random variables, a large part of mathematical quantum mechanics is concerned with operator theory and spectral theory.

2.2 Elements of operator theory

$$A : \mathrm{D}(A) \longrightarrow \mathrm{R}(A) \subset \mathcal{H}$$

where $\mathrm{D}(A)$ is the domain of A, always assumed to be dense in $\mathcal{H}$ and $\mathrm{R}(A)$ is its range. Linearity means that

$$A(\alpha_1\psi_1 + \alpha_2\psi_2) = \alpha_1 A\psi_1 + \alpha_2 A\psi_2$$

for all $\alpha_{1,2} \in \mathbb{C}$, $\psi_{1,2} \in \mathrm{D}(A)$.

An operator is bounded if there is $C > 0$ such that

$$\|A\psi\| \le C\|\psi\|$$

for all $\psi \in \mathrm{D}(A)$. The smallest possible such constant is the norm of A:

$$\|A\| = \sup\left\{ \frac{\|A\psi\|}{\|\psi\|} \ :\ \psi \in \mathrm{D}(A),\ \psi \neq 0 \right\}$$

(recall: the supremum is the least upper bound). A is called unbounded if the supremum is $+\infty$.

Rules.

- $A = B$ means $\mathrm{D}(A) = \mathrm{D}(B)$ and $A\psi = B\psi$ for all $\psi \in \mathrm{D}(A)$.
- $A \subset B$ if $\mathrm{D}(A) \subset \mathrm{D}(B)$ and $A\psi = B\psi$ for all $\psi \in \mathrm{D}(A)$. B is called an extension of A.
- $\mathrm{D}(A + B) = \mathrm{D}(A) \cap \mathrm{D}(B)$,

  $$\mathrm{D}(AB) = \{\, \psi \in \mathrm{D}(B)\ :\ B\psi \in \mathrm{D}(A)\,\}.$$

  In both cases, there may be extensions of $A + B$, AB so defined.
- A is invertible if it is injective and has a dense range, in which case $\mathrm{D}(A^{-1}) = \mathrm{R}(A)$ and $\mathrm{R}(A^{-1}) = \mathrm{D}(A)$.

Criterion for injectivity.

$$A \text{ is injective} \iff \operatorname{Ker}(A) = \{0\},$$

where $\operatorname{Ker}(A) = \{\psi \in \mathrm{D}(A) : A\psi = 0\}$ is the kernel of A. This follows immediately from linearity:

$$A\psi = A\phi \iff A(\psi - \phi) = 0.$$

A is called closed if its graph

$$\Gamma(A) = \left\{ (\psi, A\psi) \ : \ \psi \in \mathrm{D}(A) \right\} \ \subset \ \mathcal{H} \oplus \mathcal{H}$$

is a closed subspace, namely: if $\psi_n \in \mathrm{D}(A)$ with $\psi_n \to \psi$ and $A\psi_n \to \phi$, then $\psi \in \mathrm{D}(A)$ and $A\psi = \phi$.

A is called closable if it admits a closed extension. In that case there is a smallest one, the closure $\overline{A}$, characterised by $\Gamma(\overline{A}) = \overline{\Gamma(A)}$.

We define the resolvent of A at $z \in \mathbb{C}$ by

$$R_z(A) = (z \cdot \mathbf{1} - A)^{-1}$$

(whenever it exists) and the resolvent set is

$$\rho(A) = \{\, z \in \mathbb{C} \ : \ R_z(A) \text{ exists and is bounded} \,\}.$$

Definition. The spectrum of A is

$$\sigma(A) = \mathbb{C} \setminus \rho(A).$$

For any $z, w \in \rho(A)$, the identity

$$w - z = (w - A) - (z - A)$$

implies that

$$\begin{aligned}(w-z)R_z(A)R_w(A) &= R_z(A)\,\underbrace{(w-A)R_w(A)}_{=\mathbf{1}} - \underbrace{R_z(A)(z-A)}_{=\mathbf{1}}\,R_w(A) \\ &= R_z(A) - R_w(A).\end{aligned}$$

This is the resolvent identity.

Lemma 1. *Let $z, w \in \rho(A)$. Then $R_z(A)$ commutes with $R_w(A)$ and*

$$R_z(A) - R_w(A) = (w - z)\, R_z(A)R_w(A).$$

Proposition 2. *$\sigma(A)$ is a closed subset of $\mathbb{C}$.*

Proof. We prove that $\rho(A)$ is open. Let $z \in \rho(A)$ and let $\|R_z(A)\| = a^{-1}$. Let $w \in \mathbb{C}$ be such that $|w - z| < a$. Then

$$\|(w - A) - (z - A)\| = |w - z| < a.$$

Now, for any $\psi \in \mathrm{D}(A)$, $\psi \neq 0$:

$$a\|\psi\| = a\|R_z(A)(z - A)\psi\| \leq \underbrace{a\|R_z(A)\|}_{=1}\,\|(z - A)\psi\|.$$

Now by the triangle inequality,

$$a\|\psi\| \le \|\big((z-A)-(w-A)\big)\psi\| + \|(w-A)\psi\| \le |w-z|\|\psi\| + \|(w-A)\psi\|.$$

Since $a - |w-z| > 0$, we conclude that $\|(w-A)\psi\| \neq 0$ for all $\psi \in \mathrm{D}(A)$, $\psi \neq 0$. Hence $\mathrm{Ker}(w-A) = \{0\}$ and $w-A$ is injective.

It remains to prove that $w-A$ has dense range. Let ϕ be such that $\langle \phi, (w-A)\psi\rangle = 0$ for all $\psi \in \mathrm{D}(A)$. Since the range of $(z-A)$ is dense, there is $(\phi_n)_{n\in\mathbb{N}}$ in the range such that $\phi_n \to \phi$. Letting $(z-A)\xi_n = \phi_n$, namely $\xi_n = R_z(A)\phi_n$, we have that

$$(w-A)\xi_n = (w-z)\xi_n + (z-A)\xi_n = (w-z)R_z(A)\phi_n + \phi_n.$$

Taking the inner product with ϕ:

$$0 = \langle \phi, \phi_n\rangle + (w-z)\langle \phi, R_z(A)\phi_n\rangle.$$

By Cauchy–Schwarz, the second inner product is bounded above by $\|\phi\|\|R_z(A)\|\|\phi_n\|$ since $R_z(A)$ is a bounded operator, and so the limit $n\to\infty$ yields

$$\|\phi\|^2 \le |w-z|\,\|R_z(A)\|\,\|\phi\|^2 < \|\phi\|^2$$

unless $\|\phi\| = 0$, namely $\phi = 0$. We conclude that the orthogonal complement of the range of $(w-A)$ is $\{0\}$, so the range is dense.

Finally, $\|(w-A)\psi\| \ge (a-|w-z|)\|\psi\|$ for all $\psi \in \mathrm{D}(A)$ reads $(a-|w-z|)\|R_w(A)\phi\| \le \|\phi\|$ for any $\phi \in \mathrm{R}(w-A) = \mathrm{D}(R_w(A))$, namely $\|R_w(A)\| \le (a-|w-z|)^{-1}$ and the resolvent is bounded. Hence $w \in \rho(A)$. □

If $(z-A)$ has a non-trivial kernel, namely there is $\psi_z \in \mathrm{D}(A)$, $\psi_z \neq 0$ s.t.

$$A\psi_z = z\,\psi_z,$$

then $z \in \sigma(A)$ is called an eigenvalue and ψ_z an eigenvector.

Important note. Not all elements of $\sigma(A)$ are eigenvalues!

Example. Let $f:\mathbb{R}\to\mathbb{R}$ be measurable and let M_f be defined on $L^2(\mathbb{R};\mathbb{C})$ by

$$(M_f\psi)(x) = f(x)\,\psi(x) \qquad \text{(multiplication operator)},$$

with

$$\mathrm{D}(M_f) = \big\{\, \psi \in L^2(\mathbb{R};\mathbb{C}) \ :\ f\psi \in L^2(\mathbb{R};\mathbb{C}) \,\big\}.$$

For any $z \in \mathbb{C}$, let N_z be defined by

$$(N_z\psi)(x) = \frac{1}{z-f(x)}\,\psi(x), \qquad \mathrm{D}(N_z) = \left\{\psi \in \mathcal{H} \ :\ \frac{\psi(x)}{z-f(x)} \in \mathcal{H}\right\}.$$

Now

$$\left\|\frac{\psi(\cdot)}{z-f(\cdot)}\right\|_2 \le \left\|\frac{1}{z-f(\cdot)}\right\|_\infty \|\psi\|_2$$

and so

$$\|N_z\| = \left\|\frac{1}{z-f(\cdot)}\right\|_\infty,$$

which is finite if there is $\varepsilon > 0$ such that

$$\big|\{x\in\mathbb{R} : |f(x)-z| < \varepsilon\}\big| = 0.$$

We conclude:

Theorem 3. *Let f be a measurable function. Then*

$$\sigma(M_f) = \big\{\, z \in \mathbb{C} \ :\ \forall \varepsilon > 0, \big|\{x \in \mathbb{R} : |f(x) - z| < \varepsilon\}\big| > 0 \,\big\}.$$

If f is continuous, then $\sigma(M_f)$ is the closure of the range of f.

Example. Position operator: $f(x) = x$, so $\sigma(X) = \mathbb{R}$.

Does M_f have any eigenvalues? z is an eigenvalue if there is $\psi \neq 0$ such that

$$f(x)\psi(x) = z\,\psi(x) \qquad \text{almost everywhere.}$$

We conclude that $\psi(x) = 0$ for almost every x such that $f(x) \neq z$. In other words: $\psi = 0$ if $|\{x : f(x) = z\}| = 0$, namely

$$\mathrm{Ker}(z - M_f) = \{0\} \quad \text{if} \quad \big|f^{-1}(\{z\})\big| = 0,$$

and z is an eigenvalue if and only if $\big|f^{-1}(\{z\})\big| > 0$.

In particular:

- X has no eigenvalues; Possible measurement outcomes are all of $\mathbb{R}$.
- $P_\Omega = M_f$ with $f = \chi_\Omega$ has eigenvalues $\{0, 1\}$, provided Ω and $\mathbb{R} \setminus \Omega$ have positive measure.

2.2.1 The adjoint

Let $A : \mathrm{D}(A) \to \mathcal{H}$. The adjoint A^* of A is defined as follows:

$$\phi \in \mathrm{D}(A^*) \quad \Longleftrightarrow \quad \text{there is } C(\phi) \text{ such that } |\langle \phi, A\psi \rangle| \leq C(\phi)\|\psi\| \ \text{ for all } \psi \in \mathrm{D}(A).$$

In other words, the map $\psi \mapsto \langle \phi, A\psi \rangle$ is a bounded linear map $\mathcal{H} \to \mathbb{C}$. By Riesz' lemma (see below) there is a unique $\xi \in \mathcal{H}$ such that

$$\langle \phi, A\psi \rangle = \langle \xi, \psi \rangle,$$

and we define A^* on $\mathrm{D}(A^*)$ by $\xi = A^*\phi$. Equivalently,

$$\langle \phi, A\psi \rangle = \langle A^*\phi, \psi \rangle \qquad \big(\psi \in \mathrm{D}(A),\ \phi \in \mathrm{D}(A^*)\big).$$

The subtlety here is that $\mathrm{D}(A)$ is in general different from $\mathrm{D}(A^*)$.

Properties. If A is bounded:

- $\|A^*\| = \|A\|$
- $(\lambda A)^* = \overline{\lambda} A^*$
- $(A + B)^* = A^* + B^*$
- $(AB)^* = B^* A^*$
- $A^{**} = A$

Definition.

(i) A is symmetric if $\langle \phi, A\psi \rangle = \langle A\phi, \psi \rangle$ for all $\phi, \psi \in \mathrm{D}(A)$

(ii) A is self-adjoint if $A = A^*$.

Notes.

- A is symmetric $\Longleftrightarrow$ $A \subset A^*$, and self-adjoint is symmetric and $D(A) = D(A^*)$.
- If A, B are bounded and defined on all of $\mathcal{H}$ (which can always be done), then symmetric $\Longleftrightarrow$ self-adjoint; this is in particular the case if $\dim \mathcal{H} < \infty$.

2.2.2 The spectral theorem

Spectral theorem. Let $A = A^*$. There is a unique map $\mathcal{B}(\mathbb{R}) \ni f \mapsto f(A)$ such that

$$(\alpha f + g)(A) = \alpha f(A) + g(A) \tag{i}$$

$$(fg)(A) = f(A)\,g(A) \tag{ii}$$

$$\overline{f}(A) = f(A)^* \tag{iii}$$

$$f(A) = \begin{cases} \mathbf{1} & \text{for } f(x) = 1, \\ A & \text{for } f(x) = x, \end{cases} \tag{iv}$$

namely, $f \mapsto f(A)$ is a homomorphism from the unital $*$-algebra of Borel functions to the unital $*$-algebra of self-adjoint operators.

The set $\mathcal{B}(\mathbb{R})$ of Borel functions is the smallest set of functions that contains all continuous functions vanishing at $\pm\infty$ and that is closed under pointwise convergence:

$$(f_n \in \mathcal{B}(\mathbb{R}) \;\text{ and }\; f_n(x) \to f(x)\;\forall x) \quad \text{implies} \quad f \in \mathcal{B}(\mathbb{R}).$$

For finite dimensional self-adjoint matrices, $A = \sum_j a_j\, |\phi_j\rangle\langle\phi_j|$ and

$$f(A) = \sum_j f(a_j)\, |\phi_j\rangle\langle\phi_j|;$$

for example, the exponential of a matrix.

Now, for any interval $I \subset \mathbb{R}$,

$$P_I(A) = \chi_I(A)$$

is an orthogonal projection,

$$P_I(A) = P_I(A)^2 = P_I(A)^* \qquad \text{(by (ii), (iii) above)},$$

and if $I_1 \cap I_2 = \emptyset$,

$$P_{I_1}(A) + P_{I_2}(A) = P_{I_1 \cup I_2}(A) \qquad \text{(by (i))}.$$

Hence, for any $\psi \in \mathcal{H}$ with $\|\psi\| = 1$,

$$\mu_\psi^A(I) = \langle \psi, P_I(A)\psi \rangle$$

is a probability measure on $\mathbb{R}$:

- $\mu_\psi^A(I) = \|P_I(A)\psi\|^2 \geq 0$,
- additivity: $\mu_\psi^A(I_1) + \mu_\psi^A(I_2) = \mu_\psi^A(I_1 \cup I_2)$ $(I_1 \cap I_2 = \emptyset)$,
- $\mu_\psi^A(\mathbb{R}) = \|P_{\mathbb{R}}(A)\psi\| = \|\psi\| = 1$.

Note that, strictly speaking, μ_ψ^A so defined is a premeasure defined on the algebra of intervals. Standard methods of measure theory allow for an extension to a proper Borel measure.

Physical interpretation. $\mu_\psi^A(I)$ is the probability that a measurement of A returns a value $x \in I$, if the system was in the state ψ just before the measurement.

We recover informally the expectation value introduced above. The probability of a measurement in $(\lambda, \lambda + d\lambda)$ is $d\mu_\psi^A((-\infty, \lambda])$, and so the expectation value (now in the sense of probability theory) is

$$\langle A \rangle_\psi \;=\; \int \lambda\, d\mu_\psi^A\big((-\infty, \lambda]\big) = \int \lambda\, d\langle \psi, P_{(-\infty,\lambda]}(A)\psi \rangle = \langle \psi, A\psi \rangle,$$

since $\int \lambda\, d\chi_{(-\infty,\lambda]}(x) = x$ and by (iv) above.

Remark. *One often writes*

$$A = \int \lambda \, dP_\lambda,$$

which is analogous to $A = \sum_j a_j |\phi_j\rangle\langle\phi_j|$ *in the case* $\dim \mathcal{H} < \infty$.

For a self-adjoint operator, the variance is given by

$$\big\langle (\Delta A)^2 \big\rangle_\psi = \big\langle \big(A - \langle A\rangle_\psi\big)^2 \big\rangle_\psi = \big\| \big(A - \langle A\rangle_\psi\big)\psi \big\|^2 = \langle A^2\rangle_\psi - \langle A\rangle_\psi^2.$$

From here,

$$\big\langle (\Delta A)^2 \big\rangle_\psi = 0 \quad \Longleftrightarrow \quad A\psi = \langle A\rangle_\psi \, \psi,$$

namely: the variance in the measurement outcome vanishes, namely the measurement outcome is deterministic, if and only if ψ is an eigenstate of A.

Remarks.

- If $\sigma(A)$ is made up of non-degenerate eigenvalues only, then

  $$P_I(A) = \sum_{a \in \sigma(A) \cap I} |\phi_a\rangle\langle\phi_a|,$$

  so the probability of the measurement outcome $a \in \mathbb{R}$ is $\langle \psi, P_a \psi\rangle$, namely $|\langle \phi_a, \psi\rangle|^2$. In particular, if $\psi = \phi_b$,

  $$\mathrm{Prob}(\text{measurement} = a) = \begin{cases} 1 & \text{if } a = b, \\ 0 & \text{otherwise.} \end{cases}$$

- Repeating a measurement just after the measurement should yield the same result. Hence the state of the system immediately after a measurement with result a is ϕ_a: this is the infamous "collapse" of the wave function induced by the measurement.

Proposition 4 (Heisenberg)**.** *For* A, B *bounded and self-adjoint,*

$$\big\langle (\Delta A)^2 \big\rangle_\psi \big\langle (\Delta B)^2 \big\rangle_\psi \geq \frac{1}{4} \big| \langle [A, B] \rangle_\psi \big|^2.$$

Proof.

$$\big| \langle \psi, [A, B]\psi\rangle \big| = \big| \langle A\psi, B\psi\rangle - \langle B\psi, A\psi\rangle \big| \leq 2\|A\psi\| \, \|B\psi\|.$$

The claim follows by taking the square and replacing $A \to A - \langle A\rangle_\psi$, $B \to B - \langle B\rangle_\psi$ (and noting commutator is unchanged). □

In general, two non-commuting operators do not have common eigenvectors, hence they cannot both have sharply defined values. Note that the famous Heisenberg uncertainty principle is the above inequality but for the unbounded operators of position and momentum, see the exercises.

We turn to Riesz' lemma. For any $\phi \in \mathcal{H}$ the map $\ell_\phi : \mathcal{H} \to \mathbb{C}$ given by $\ell_\phi(\psi) = \langle \phi, \psi\rangle$ is a bounded linear functional of norm $\|\phi\|$ since $|\ell_\phi(\psi)| \leq \|\phi\|\|\psi\|$ by Cauchy-Schwarz. Riesz' lemma claims that *all* bounded linear functionals on $\mathcal{H}$ arise in this way.

Lemma 5 (Riesz)**.** *Let* ℓ *be a bounded linear functional on* $\mathcal{H}$. *There is a unique* $\eta_\ell \in \mathcal{H}$ *such that* $\ell(\psi) = \langle \eta_\ell, \psi\rangle$ *for all* $\psi \in \mathcal{H}$.

Proof. Existence. The set $\mathcal{N} := \{\psi \in \mathcal{H} : \ell(\psi) = 0\} = \ell^{-1}(\{0\})$ is closed, since ℓ is continuous (because bounded). If $\mathcal{N} = \mathcal{H}$ we pick $\eta_\ell = 0$. Otherwise there is $\eta \in \mathcal{N}^\perp$, i.e. $\langle \eta, \psi \rangle = 0$ for all $\psi \in \mathcal{N}$, with $\|\eta\| = 1$; note $\ell(\eta) \neq 0$, since $\eta \in \mathcal{N} \cap \mathcal{N}^\perp$ would force $\eta = 0$. For any $\psi \in \mathcal{H}$,

$$\ell\big(\ell(\psi)\eta - \ell(\eta)\psi\big) = 0 \qquad \text{(by linearity)},$$

so that $\ell(\psi)\eta - \ell(\eta)\psi \in \mathcal{N}$ and

$$0 = \big\langle \eta, \ell(\psi)\eta - \ell(\eta)\psi \big\rangle = \ell(\psi) - \ell(\eta)\langle \eta, \psi \rangle,$$

whence $\langle \eta, \psi \rangle = \ell(\psi)/\ell(\eta)$; namely

$$\eta_\ell := \overline{\ell(\eta)}\,\eta$$

has the right properties. Uniqueness: if η_ℓ' also works, then $\langle \eta_\ell' - \eta_\ell, \psi \rangle = \ell(\psi) - \ell(\psi) = 0$ for all $\psi \in \mathcal{H}$, so $\eta_\ell' - \eta_\ell \in \mathcal{H}^\perp$, namely $\eta_\ell' - \eta_\ell = 0$. □

Corollary 6. *Let $b : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ be a sesquilinear form on $\mathcal{H} \times \mathcal{H}$ (linear in ψ, antilinear in ϕ). If b is bounded,*

$$|b(\phi, \psi)| \le C\|\phi\|\|\psi\|,$$

then there is a unique bounded operator B on $\mathcal{H}$ such that $b(\phi, \psi) = \langle \phi, B\psi \rangle$.

2.2.3 The momentum operator on $\mathcal{H} = L^2([0,1];\mathbb{C})$

The derivative of $\psi \in \mathcal{H}$ is understood in the sense of distributions:

$$\psi' : C_c^\infty((0,1)) \to \mathbb{C}, \qquad v \longmapsto \psi'[v] = -\int_0^1 \psi(x)v'(x)dx$$

(formal integration by parts).

Lemma 7 (Morrey)**.** *If $\psi' \in L^2([0,1];\mathbb{C})$, then ψ is continuous, and for any ϕ such that $\phi' \in L^2([0,1];\mathbb{C})$,*

$$\int_0^1 \big(\psi'(x)\phi(x) + \psi(x)\phi'(x)\big)dx = \psi(x)\phi(x)\Big|_0^1.$$

Proof. Define $\widetilde{\psi}(x) = \int_0^x \psi'(y)dy$.

- Well-defined by Cauchy–Schwarz: $\int_0^x |\psi'(y)|dy \le \big(\int_0^x |\psi'(y)|^2 dy\big)^{1/2}\big(\int_0^x 1\big)^{1/2}$.
- Continuous: by Cauchy–Schwarz again, for $x' \le x$,

  $$\big|\widetilde{\psi}(x) - \widetilde{\psi}(x')\big| = \left|\int_{x'}^x \psi'(y)dy\right| \le \|\psi'\|_2\, |x - x'|^{1/2},$$

  so $\widetilde{\psi}$ is (Hölder-$\frac{1}{2}$) continuous.
- $\psi' = \widetilde{\psi}\,'$. Indeed, for $v \in C_c^\infty((0,1))$ (so $v(0) = v(1) = 0$),

  $$\begin{aligned}\widetilde{\psi}\,'[v] &= -\int_0^1 \widetilde{\psi}(x)v'(x)dx = -\int_0^1 v'(x)\int_0^x \psi'(y)dy\,dx = -\int_0^1\int_y^1 v'(x)\psi'(y)dxdy \\ &= -\int_0^1 \psi'(y)\big(v(1) - v(y)\big)\mathrm{d}y = \int_0^1 \psi'(y)v(y)dy = \psi'[v].\end{aligned}$$

(note that the last equality is precisely what is meant by $\psi' \in L^2([0,1];\mathbb{C})$). In other words, $\widetilde{\psi}\,'[v] - \psi'[v] = 0$ for all $v \in C_c^\infty((0,1))$, and so $\psi - \widetilde{\psi}$ is constant. Therefore ψ is continuous. With this, we multiply the identity $\psi(x) = \psi(0) + \int_0^x \psi'(y)dy$ by $\phi'(x)$ and integrate over $[0,1]$ to get

$$\int_0^1 \psi(x)\phi'(x)dx = \psi(0)(\phi(1) - \phi(0)) + \int_0^1 \int_0^x \psi'(y)\phi'(x)dydx.$$

Repeating the argument with $\phi \leftrightarrow \psi$, $\int_0^1 \psi'(x)\phi(x)dx = \phi(0)(\psi(1) - \psi(0)) + \int_0^1 \int_0^x \phi'(y)\psi'(x)dydx$. Summing the two identities and using that the double integral is over the complementary triangle in $[0,1]^2$ yields

$$\int_0^1 (\psi\phi' + \psi'\phi) = \psi(0)(\phi(1) - \phi(0)) + \phi(0)(\psi(1) - \psi(0)) + (\phi(1) - \phi(0))(\psi(1) - \psi(0))$$

which is equal to $(\phi\psi)|_0^1$ after simplifying. $\square$

Definition.

(i) $\mathrm{D}(\widetilde{P}) = \{\psi \in \mathcal{H} : \psi' \in \mathcal{H}\}$, $\widetilde{P}\psi = -\mathrm{i}\psi'$;

(ii) $\mathrm{D}(P_0) = \{\psi \in \mathrm{D}(\widetilde{P}) : \psi(0) = \psi(1) = 0\}$, (well defined by the lemma). $P_0\psi = \widetilde{P}\psi$

Clearly $P_0 \subsetneq \widetilde{P}$.

Claim. $P_0 = \widetilde{P}^*$.

Proof. $\phi \in \mathrm{D}(\widetilde{P}^*)$ if $|\langle \phi, \widetilde{P}\psi\rangle| \le C\|\psi\|$ for all $\psi \in \mathrm{D}(\widetilde{P})$. Testing this on $\psi = \overline{v}$ with $v \in C_c^\infty((0,1))$,

$$\langle \phi, \widetilde{P}\overline{v}\rangle = -\mathrm{i}\int_0^1 \overline{\phi}\,\overline{v}\,'dx = \mathrm{i}\,\overline{\phi'[v]},$$

yields $|\phi'[v]| \le C\|v\|$. Hence $v \mapsto \phi'[v]$ is a bounded linear functional on the dense subspace $C_c^\infty((0,1)) \subset L^2([0,1];\mathbb{C})$; it extends uniquely by continuity to $L^2([0,1];\mathbb{C})$, and Riesz' lemma gives $\phi' \in L^2([0,1];\mathbb{C})$, namely $\phi \in \mathrm{D}(\widetilde{P})$. Now for any $\psi \in \mathrm{D}(\widetilde{P})$,

$$\langle \widetilde{P}\psi, \phi\rangle - \langle \psi, \widetilde{P}\phi\rangle = \mathrm{i}\int_0^1 \big(\overline{\psi}\,'\phi + \overline{\psi}\phi'\big) = \mathrm{i}\big(\overline{\psi(1)}\phi(1) - \overline{\psi(0)}\phi(0)\big) \tag{2}$$

by the lemma. We note

$$\big|\langle \psi, \widetilde{P}\phi\rangle\big|/\|\psi\| \le \|\widetilde{P}\phi\|,$$

but the right-hand side of (2) can be made arbitrarily large by concentrating ψ at 1 or at 0. Hence $\big|\langle \widetilde{P}\psi, \phi\rangle\big|/\|\psi\|$ is bounded for all ψ if and only if $\phi(0) = \phi(1) = 0$, so $\mathrm{D}(\widetilde{P}^*) = \mathrm{D}(P_0)$ and $\widetilde{P}^*\phi = \widetilde{P}\phi = P_0\phi$. $\square$

Remark. *For $\phi, \psi \in \mathrm{D}(P_0)$, (2) gives $\langle P_0\psi, \phi\rangle = \langle \psi, P_0\phi\rangle$, namely P_0 is symmetric — but not self-adjoint:*

$$P_0 \subsetneq \widetilde{P} = \widetilde{P}^{**} = P_0^*$$

where the middle equality is because $\widetilde{P}$ is closed.

Question. Can we extend P_0 — namely find a domain between $\mathrm{D}(P_0)$ and $\mathrm{D}(\widetilde{P})$ — so that $-\mathrm{i}\frac{d}{dx}$ becomes self-adjoint?
Answer. Yes; there is a one-parameter family of self-adjoint extensions. For $\alpha \in \mathbb{C}$ with $|\alpha| = 1$, define

$$\mathrm{D}(P_\alpha) = \{\psi \in \mathrm{D}(\widetilde{P}) : \psi(1) = \alpha\psi(0)\}, \qquad P_\alpha\psi = \widetilde{P}\psi.$$

Clearly $P_0 \subset P_\alpha \subset \widetilde{P}$.

Proposition 8. $P_\alpha = P_\alpha^*$.

Proof. As above, first $\mathrm{D}(P_\alpha^*) \subset \mathrm{D}(\widetilde{P})$. Second, for $\psi \in \mathrm{D}(P_\alpha)$, (2) reads

$$\langle \widetilde{P}\psi, \phi\rangle - \langle \psi, \widetilde{P}\phi\rangle = \mathrm{i}\,\overline{\psi(0)}\big(\overline{\alpha}\phi(1) - \phi(0)\big),$$

and by the same reasoning $\phi \in \mathrm{D}(P_\alpha^*)$ if and only if $\overline{\alpha}\phi(1) - \phi(0) = 0$, namely $\phi(1) = \alpha\phi(0)$ (using $|\alpha| = 1$); hence $\phi \in \mathrm{D}(P_\alpha)$. ☐

Summary.

- All the P's above act as $-\mathrm{i}\frac{d}{dx}$.
- With Dirichlet boundary conditions, P_0 is symmetric, not self-adjoint.
- With periodic boundary conditions (and a magnetic flux), P_α is self-adjoint.
- P_0 and P_α differ non-trivially despite agreeing on the dense set $\mathrm{D}(P_0)$. For example ($\alpha = 1$): P_0 has no eigenvector, but P_1 has an orthonormal basis of eigenvectors $\phi_n(x) = \mathrm{e}^{2\pi\mathrm{i}nx}$.
- Self-adjointness is crucial for a physical dynamics; see later (Stone's theorem).

In fact, as a differential equation the eigenvalue equation $-\mathrm{i}\psi' = z\psi$ has the unique solution $\psi(x) = C\,\mathrm{e}^{\mathrm{i}zx}$, for all $z \in \mathbb{C}$. Hence:

(i) $\widetilde{P}$: for any $z \in \mathbb{C}$, $C\mathrm{e}^{\mathrm{i}zx} \in L^2([0,1];\mathbb{C})$ and $\psi' = \mathrm{i}z\,\mathrm{e}^{\mathrm{i}zx} \in L^2([0,1];\mathbb{C})$, so all $z \in \mathbb{C}$ are eigenvalues and $\sigma(\widetilde{P}) = \mathbb{C}$.

(ii) P_α: $\psi(1) = \alpha\psi(0) \iff \mathrm{e}^{\mathrm{i}z} = \alpha$, namely, writing $\alpha = \mathrm{e}^{\mathrm{i}\theta}$ with $\theta \in [0, 2\pi)$, $z = \theta + 2\pi n$, $n \in \mathbb{Z}$.

(iii) P_0: $\psi(1) = 0$ and $\psi(0) = 0$ force $C = 0$, so there is no eigenvector.

2.2.4 Quadratic forms and the Friedrichs extension

We have seen:

(i) Symmetric operators may have many self-adjoint extensions.

(ii) Physics often comes with a formal quadratic form for the energy (Hamiltonian) which is bounded below. Example:

$$E(\psi) = \int \left(|\nabla\psi(x)|^2 + V(x)|\psi(x)|^2\right) dx \geq E_0 > -\infty.$$

Question. Is there a natural self-adjoint operator associated with a bounded quadratic form?
Answer. Yes: the *Friedrichs extension.*

As an example, consider the free particle in a box, $T = -\frac{d^2}{dx^2}$ defined on $C_c^\infty((0,1))$, which is dense in $L^2([0,1])$. Then

$$\langle \psi, T\psi\rangle = -\int_0^1 \overline{\psi(x)}\,\psi''(x)dx \overset{\text{IBP}}{=} \int_0^1 \overline{\psi'(x)}\,\psi'(x)dx = \int_0^1 |\psi'(x)|^2 dx \geq 0.$$

In general, a densely defined operator T is called nonnegative if the associated quadratic form satisfies

$$q_T(\psi) = \langle \psi, T\psi\rangle \geq 0 \qquad \text{for all } \psi \in \mathrm{D}(T).$$

Lemma 9. *$T \geq 0$ implies that T is symmetric.*

Proof. $T \geq 0$ implies in particular that $\langle \psi, T\psi\rangle \in \mathbb{R}$, namely $\langle \psi, T\psi\rangle = \overline{\langle \psi, T\psi\rangle} = \langle T\psi, \psi\rangle$ for all $\psi \in D(T)$. In other words $b(\phi,\psi) = \langle \phi, T\psi\rangle - \langle T\phi, \psi\rangle$ is such that $b(\psi,\psi) = 0$. That $b = 0$ follows by polarization. □

We introduce an inner product on $\mathrm{D}(T)$,

$$\langle \psi, \phi\rangle_T = \langle \psi, T\phi\rangle + \langle \psi, \phi\rangle,$$

and thereby a norm $\|\cdot\|_T$. Let $Q(T)$ be the completion of $\mathrm{D}(T)$ with respect to $\|\cdot\|_T$ — i.e. $Q(T)$ is the set of equivalence classes of Cauchy sequences in $\mathrm{D}(T)$, where $\psi_n \sim \phi_n$ if $\|\psi_n - \phi_n\|_T \to 0$. In particular $\psi_n \to \psi$ in $\mathcal{H}$, and so

$$\mathrm{D}(T) \subset Q(T) \subset \mathcal{H}.$$

$Q(T)$ is called the form domain of q_T (or, sometimes, of T).

Theorem 10 (Friedrichs)**.** *Let T be a symmetric operator such that $T \geq 0$. There is a unique self-adjoint extension $\widetilde{T}$ of T such that (i) $\widetilde{T} \geq 0$ and (ii) $\mathrm{D}(\widetilde{T}) \subset Q(T)$. In fact*

$$\mathrm{D}(\widetilde{T}) = \mathrm{D}(T^*) \cap Q(T), \qquad \widetilde{T}\psi = T^*\psi \quad \textit{for all } \psi \in \mathrm{D}(\widetilde{T}). \tag{3}$$

Proof. $T \subset \widetilde{T}$. For a symmetric operator, $T \subset T^*$: indeed T symmetric gives $\langle T\psi, \phi\rangle = \langle \psi, T\phi\rangle$ for all $\psi, \phi \in \mathrm{D}(T)$, namely $\phi \in \mathrm{D}(T^*)$ for all $\phi \in \mathrm{D}(T)$, namely $\mathrm{D}(T) \subset \mathrm{D}(T^*)$, and $T^*\psi = T\psi$ for $\psi \in \mathrm{D}(T)$. Since $Q(T) \supset \mathrm{D}(T)$ and $\mathrm{D}(T^*) \supset \mathrm{D}(T)$, we get $\mathrm{D}(T) \subset Q(T) \cap \mathrm{D}(T^*) = \mathrm{D}(\widetilde{T})$, and with (3), $T \subset \widetilde{T}$.

$\widetilde{T} \geq 0$. Write $a(u,v) = \langle u, Tv\rangle$ for the quadratic form of T. It is Hermitian (T being symmetric) and nonnegative, so Cauchy–Schwarz applies to a itself:

$$|a(u,v)| \leq a(u,u)^{1/2}\, a(v,v)^{1/2} \leq \|u\|_T \|v\|_T \qquad (u, v \in \mathrm{D}(T)),$$

and by continuity a extends to $Q(T)$ with the same bound.

Let $\psi \in \mathrm{D}(\widetilde{T}) \subset Q(T)$ and let $\psi_n \in \mathrm{D}(T)$ with $\|\psi_n - \psi\|_T \to 0$; being convergent, (ψ_n) is bounded, say $M := \sup_n \|\psi_n\|_T < \infty$. For all n, m,

$$\begin{aligned}\langle \psi_n, \widetilde{T}\psi\rangle = \langle \psi_n, T^*\psi\rangle = \langle T\psi_n, \psi\rangle &= \langle T\psi_n, \psi_m\rangle + \langle T\psi_n, \psi - \psi_m\rangle \\ &= \langle T(\psi_n - \psi_m), \psi_m\rangle + \langle T\psi_m, \psi_m\rangle + \langle T\psi_n, \psi - \psi_m\rangle.\end{aligned}$$

The first and third terms are $\overline{a(\psi_m, \psi_n - \psi_m)}$ and $\overline{a(\psi - \psi_m, \psi_n)}$, hence bounded in modulus by $M\|\psi_n - \psi_m\|_T$ and $M\|\psi - \psi_m\|_T$ respectively, while the second is $a(\psi_m, \psi_m) \geq 0$.

Let $\varepsilon > 0$ and choose N with $\|\psi_n - \psi_m\|_T < \varepsilon$ and $\|\psi - \psi_m\|_T < \varepsilon$ for all $n, m \geq N$. Then

$$\left|\langle \psi_n, \widetilde{T}\psi\rangle - a(\psi_m, \psi_m)\right| \leq 2\varepsilon M \qquad (n, m \geq N),$$

and letting $n \to \infty$, which is legitimate since $\|\psi_n - \psi\| \leq \|\psi_n - \psi\|_T \to 0$,

$$\left|\langle \psi, \widetilde{T}\psi\rangle - a(\psi_m, \psi_m)\right| \leq 2\varepsilon M \qquad (m \geq N).$$

So $\langle \psi, \widetilde{T}\psi\rangle$ lies within $2\varepsilon M$ of $[0, \infty)$ for every $\varepsilon > 0$, whence $\langle \psi, \widetilde{T}\psi\rangle \geq 0$ for all $\psi \in \mathrm{D}(\widetilde{T})$.

Self-adjointness. It suffices to show that $\mathrm{R}(1 + \widetilde{T}) = \mathcal{H}$. We postpone this for later. $\square$

Back to the example $T = -\frac{d^2}{dx^2}$ on $C_c^\infty((0,1))$:

$$\|\psi\|_T^2 = \langle \psi, T\psi\rangle + \langle \psi, \psi\rangle = \|\psi'\|_2^2 + \|\psi\|_2^2,$$

which is the H^1-norm. Let now $\psi \in Q(T)$ and let $\psi_n \in C_c^\infty((0,1))$ with $\psi_n \to \psi$ in $\|\cdot\|_T$. In particular $\psi' \in L^2([0,1];\mathbb{C})$. As in the proof of Morrey's lemma, we have for every $x \in [0,1]$,

$$|\psi_n(x) - \psi(x)| = \left|\int_0^x \left(\psi_n' - \psi'\right)\right| \leq \|\psi_n' - \psi'\|_2 \longrightarrow 0.$$

The bound being independent of x, $\psi_n \to \psi$ uniformly on $[0,1]$. Hence $\psi(0) = \lim_{n\to\infty} \psi_n(0) = 0$ and similarly $\psi(1) = 0$ because each ψ_n has compact support in $(0,1)$.

Hence for all $\psi \in Q(T)$ we must have (i) $\psi' \in L^2([0,1];\mathbb{C})$ and (ii) $\psi(0) = \psi(1) = 0$, and so

$$Q(T) \subset \left\{ \psi \in L^2([0,1];\mathbb{C}) \ : \ \psi' \in L^2([0,1];\mathbb{C}) \ \text{ and } \ \psi(0) = \psi(1) = 0 \right\}.$$

The reverse inclusion holds as well (sketch): if $\psi' \in L^2([0,1];\mathbb{C})$ and $\psi(0) = \psi(1) = 0$, then truncating ψ near the endpoints and mollifying produces $\psi_n \in C_c^\infty((0,1))$ with $\|\psi_n - \psi\|_T \to 0$, so $\psi \in Q(T)$ and the two sets are equal.

Now, $\phi \in \mathrm{D}(T^*)$ if and only if $|\langle \phi, Tv\rangle| \leq C\|v\|$ for all $v \in C_c^\infty((0,1))$, namely if and only if $\phi'' \in L^2([0,1];\mathbb{C})$. Hence, $\mathrm{D}(T^*) = \{\phi \in L^2([0,1];\mathbb{C}) : \phi'' \in L^2([0,1];\mathbb{C})\}$.

Conclusion. By (3), the Friedrichs extension of $-\frac{d^2}{dx^2}$ therefore has domain

$$\mathrm{D}(\widetilde{T}) = \left\{ \psi \in L^2([0,1];\mathbb{C}) \ : \ \psi', \psi'' \in L^2([0,1];\mathbb{C}) \ \text{ and } \ \psi(0) = \psi(1) = 0 \right\}.$$

This is the Dirichlet Laplacian.

Finally, we compute its spectrum. For $\lambda \neq 0$ the equation $-\psi''(x) = \lambda\psi(x)$ has the two independent solutions $\mathrm{e}^{\pm \mathrm{i}\sqrt{\lambda}x}$, which lie in $L^2([0,1];\mathbb{C})$ for every $\lambda \in \mathbb{C}$. Imposing $\psi(0) = 0$ leaves $\psi(x) = C\sin(\sqrt{\lambda}\, x)$, and $\psi(1) = 0$ then forces $\sin\sqrt{\lambda} = 0$, namely

$$\lambda = (n\pi)^2, \quad \psi_n(x) = \sin(n\pi x) \qquad (n \in \mathbb{N}).$$

For $\lambda = 0$ the two exponentials coincide and the solution space is spanned by 1 and x instead; $\psi(x) = a + bx$ with $\psi(0) = a = 0$ and $\psi(1) = a + b = 0$ gives $\psi = 0$, so 0 is not an eigenvalue.

Note that $\widetilde{T} - \pi^2 \geq 0$, and hence the same must be true for the original T, namely

$$\int_0^1 |\psi'(x)|^2 dx \geq \pi^2 \int_0^1 |\psi(x)|^2 dx.$$

This is also known as *Poincaré's inequality* (for functions vanishing at the boundary).