Course Notes

Ecole Polytechnique Fédérale de Lausanne

Mathematics Department

Spectral theory

Author
Vincent Dumoncel

Professor
François Genoud

January 31, 2023

Contents

1. Operators on Banach and Hilbert spaces

In this first section, we recall important notions and tools from functional analysis, and we introduce different class of operators.

1.1 Linear operators

In these notes we will only consider vector spaces over complex numbers $\mathbb{C}$.

Definition 1.1. Let E be a $\mathbb{C}-$vector space.
A norm on E is a map $\|\cdot\| \colon E \longrightarrow [0,\infty)$ satisfying the following properties :

(i) $\|x\| = 0$ if and only if $x = 0$.

(ii) $\|\alpha x\| = |\alpha|\|x\|$ for all $x \in E$ and $\alpha \in \mathbb{C}$.

(iii) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in E$.

Such a map automatically induces a metric on E, via the formula $\mathrm{d}(x,y) = \|x - y\|$ (check!).

Definition 1.2. A normed space $(E, \|\cdot\|)$ is a Banach space if the corresponding metric space (E, d) is complete.

We can then properly study linear maps between such vector spaces.

Definition 1.3. Let X_1, X_2 be two Banach spaces.
A linear operator $A \colon X_1 \longrightarrow X_2$ is a map such that $A(u + \lambda v) = A(u) + \lambda A(v)$, for all $u, v \in X_1$ and $\lambda \in \mathbb{C}$.

We shall merely write Au for the image $A(u)$ of $u \in X_1$. As usual, for such a map we define its image (or range) and its kernel by

$$\mathrm{Ker}(A) := \{u \in X_1 : Au = 0\} \quad \text{and} \quad \mathrm{Im}(A) := \{Au : u \in X_1\}.$$

One checks directly that both are vector subspaces of X_1 and X_2 respectively, and in addition that $\mathrm{Ker}(A)$ is closed.

Since X_1 and X_2 both carry a topology, we are moreover interested in operators which preserve these topologies, *i.e.* which are continuous. Such operators are called *bounded* and the following lemma explains why.

Lemma 1.4. Let X_1 and X_2 be two normed spaces, and $A \colon X_1 \longrightarrow X_2$ be linear.
The following are equivalent :

(i) A is continuous on X_1.

(ii) A is continuous at $0 \in X_1$.

(iii) $\sup\{\|Au\| : \|u\| \le 1\} < \infty$.

Proof. (i) $\Longrightarrow$ (ii) : Obvious.
(ii) $\Longrightarrow$ (iii) : Towards a contradiction, suppose $\sup\{\|Au\| : \|u\| \le 1\}$ is not finite. It implies there is a sequence $(u_n)_{n\in\mathbb{N}}$ in X_1 so that $\|u_n\| \le 1$ and $\lim\limits_{n\to\infty} \|Au_n\| = \infty$. We then consider the sequence $(y_n)_{n\in\mathbb{N}}$ defined by $y_n := \frac{x_n}{\|Ax_n\|}$, for $n \in \mathbb{N}$. It converges to 0 in X_1 but $\|Ay_n\| = 1$ for all $n \in \mathbb{N}$, contradicting continuity of A at 0. Thus (iii) holds as well.
(iii) $\Longrightarrow$ (i) : Let $M := \sup\{\|Au\| : \|u\| \le 1\} < \infty$ and $\varepsilon > 0$, $u_0 \in X_1$. Let $\delta := \frac{\varepsilon}{M}$. Then if $\|u - u_0\| < \delta$ one has

$$\|Au - Au_0\| = \|A(u - u_0)\| = \|u - u_0\| \left\| A\left(\frac{u - u_0}{\|u - u_0\|}\right)\right\| < \delta \cdot M = \varepsilon$$

proving that A is continuous at $u_0 \in X_1$. □

From this lemma it immediately follows that a linear map $A : X_1 \longrightarrow X_2$ between two normed spaces is continuous if and only if there is a constant $C > 0$ so that $\|Au\| \le C\|u\|$ for all $u \in X_1$.

The space of bounded linear operators between two normed spaces X_1, X_2 will be denoted $\mathcal{B}(X_1, X_2)$, or simply $\mathcal{B}(X_1)$ if $X_1 = X_2$.

Proposition 1.5. The map $\|\cdot\| : \mathcal{B}(X_1, X_2) \longrightarrow [0, \infty)$ defined by

$$\|A\| := \sup_{u \ne 0} \frac{\|Au\|}{\|u\|} = \sup_{\|u\| \le 1} \|Au\| = \sup_{\|u\| = 1} \|Au\|$$

is a norm on $\mathcal{B}(X_1, X_2)$. In particular, $\|Au\| \le \|A\|\|u\|$ for all $u \in X_1$.

Proof. The second claim directly follows from the definition of $\|A\|$ for $u \ne 0$, and for $x = 0$ the inequality is also satisfied since both sides equal 0.

Now, suppose $\|A\| = 0$. Then for all $u \in X_1$, we have $0 \le \|Au\| \le \|A\|\|u\| = 0$, so $\|Au\| = 0$, and $Au = 0$ for all $u \in X_1$. Thus $A = 0$.

If $A, B \in \mathcal{B}(X_1, X_2)$ and $u \in X_1$ is such that $\|u\| = 1$, then

$$\|(A + B)u\| = \|Au + Bu\| \le \|Au\| + \|Bu\| \le \|A\|\|u\| + \|B\|\|u\| = \|A\| + \|B\|$$

proving the triangle inequality $\|A + B\| \le \|A\| + \|B\|$.

Finally, let $A \in \mathcal{B}(X_1, X_2)$, $c \in \mathbb{R}$ and $u \in X_1$ with $\|u\| = 1$. The computation

$$\|(cA)u\| = \|cAu\| = |c|\|Ax\| \leq |c|\|A\|\|u\| = |c|\|A\|$$

shows that $\|cA\| \leq |c|\|A\|$, and conversely since

$$\|Au\| = \frac{\|c \cdot Au\|}{|c|} = \frac{\|(cA)u\|}{|c|} \leq \frac{\|cA\|\|u\|}{|c|} = \frac{\|cA\|}{|c|}$$

we see that $\|A\| \leq \frac{\|cA\|}{|c|}$, *i.e.* $|c|\|A\| \leq \|cA\|$. Putting the two inequalities together gives $\|cA\| = |c|\|A\|$ and finishes the proof. □

Note furthermore that if X_1, X_2, X_3 are three normed spaces and $A \in \mathcal{B}(X_1, X_2)$, $B \in \mathcal{B}(X_2, X_3)$, then $\|AB\| \leq \|A\|\|B\|$ (check!).

The norm introduced in this proposition is usually called the operator norm. It turns out it $\mathcal{B}(X_1, X_2)$ is a Banach space provided the target space is Banach.

Proposition 1.6. Let X_1, X_2 be two normed spaces, and suppose that X_2 is complete. Then $\mathcal{B}(X_1, X_2)$ is complete.

Among all complete spaces we can imagine, the easiest one, namely $X_2 = \mathbb{C}$, plays a special role.

Definition 1.7. Let X be a normed space.
The space $X^* := \mathcal{B}(X, \mathbb{C})$ is called the dual space of X.

We are now ready to recall some fundamental results of functional analysis, some of them without proofs. The first one is usually called the *Banach-Steinhaus* theorem.

Theorem 1.8. Let X_1, X_2 be two Banach spaces, and $(A_n)_{n \in \mathbb{N}} \subset \mathcal{B}(X_1, X_2)$.
If $\sup\limits_{n \in \mathbb{N}} \|A_n u\| < \infty$ for all $u \in X_1$, then $\sup\limits_{n \in \mathbb{N}} \|A_n\| < \infty$.

The next one is the open mapping theorem and its direct consequences, such as the bounded inverse theorem or the closed graph theorem. Recall that a map $f : X \longrightarrow Y$ is called *open* if it maps any open set of X to an open set of Y.

Theorem 1.9. Let X_1, X_2 be two Banach spaces. Suppose that $A \colon X_1 \longrightarrow X_2$ is continuous, linear and surjective. Then A is open.

Corollary 1.10. Let X_1, X_2 be two Banach spaces. Suppose that $A \colon X_1 \longrightarrow X_2$ is continuous, linear and bijective. Then A^{-1} is continuous.

Proof. All hypotheses of Theorem 1.9 are fullfilled, and thus A is open, which is equivalent to say that A^{-1} is continuous, or bounded. □

Definition 1.11. Let X_1, X_2 be Banach spaces, and $A\colon X_1 \longrightarrow X_2$ be linear. Its graph is the subset of $X_1 \times X_2$ defined as

$$G_A := \{(u, Au) : u \in X_1\}.$$

If X_1, X_2 are two normed spaces, we endow the product $X_1 \times X_2$ with the norm (check!) $\|(u_1, u_2)\| = \|u_1\| + \|u_2\|$, for $u_1 \in X_1, u_2 \in X_2$.

The closed graph theorem states then the following.

Theorem 1.12. Let X_1, X_2 be two Banach spaces. An operator $A\colon X_1 \longrightarrow X_2$ is continuous if and only if G_A is closed in $X_1 \times X_2$.

Proof. To start, suppose that A is continuous, and that $((u_n, Au_n))_{n\in\mathbb{N}}$ is a sequence in G_A converging to $(u, v) \in X_1 \times X_2$. Then

$$v = \lim_{n\to\infty} Au_n = Au$$

by continuity of A, so that $(u, v) = (u, Au) \in G_A$. This proves that G_A is closed.

Conversely, if $A = 0$, the claim is obvious. Suppose then $A \neq 0$. Consider the projection $\pi\colon G_A \longrightarrow X_1$, $\pi(u, Au) = u$. It is a continuous linear bijective map. Moreover, the product $X_1 \times X_2$ is a Banach space, and G_A is closed, so G_A is also a Banach space. Therefore Corollary 1.10 applies, and $\pi^{-1}\colon X_1 \longrightarrow G_A$, $\pi^{-1}(u) = (u, Au)$ is continuous. Equivalently, $\|\pi^{-1}\| < \infty$. Moreover, since $A \neq 0$ we can pick $u_0 \in X_1$ with $Au_0 \neq 0$. It follows that $\|Au_0\| > 0$ and

$$\|\pi^{-1}(u_0)\| = \|(u_0, Au_0)\| = \|u_0\| + \|Au_0\| > \|u_0\|$$

proving that $\|\pi^{-1}\| > 1$. We can then conclude that

$$\|Au\| = \|\pi^{-1}(u)\| - \|u\| \leq (\|\pi^{-1}\| - 1)\|u\|$$

for all $u \in X_1$, and A is continuous. This finishes the proof. □

1.2 Hilbert spaces

Definition 1.13. Let $\mathcal{H}$ be a complex vector space.

A hermitian inner product on $\mathcal{H}$ is a sesquilinear map

$$\begin{aligned} \langle \cdot, \cdot \rangle \colon \mathcal{H} \times \mathcal{H} &\longrightarrow \mathbb{C} \\ (u, v) &\longmapsto \langle u, v \rangle \end{aligned}$$

such that

(i) $\langle \lambda u + \mu v, w \rangle = \lambda \langle u, w \rangle + \mu \langle v, w \rangle$, for all $u, v, w \in \mathcal{H}$, $\lambda, \mu \in \mathbb{C}$.

(ii) $\langle u, v \rangle = \overline{\langle u, v \rangle}$, for all $u, v \in \mathcal{H}$.

(iii) $\langle u, u \rangle \geq 0$ for all $u \in \mathcal{H}$, and $\langle u, u \rangle = 0$ implies $u = 0$.

When $\mathcal{H}$ is equipped with a hermitian inner product, the pair $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ is called a *pre-Hilbert space*.

A priori, for $u \in \mathcal{H}$, $\langle u, u \rangle$ is a complex number, and its sign is undefined. However $\overline{\langle u, u \rangle} = \langle u, u \rangle$ by (ii), so $\langle u, u \rangle$ is in fact a real number. Also, the above properties together imply

$$\langle u, \lambda v + \mu w \rangle = \overline{\langle \lambda v + \mu w, u \rangle} = \overline{\lambda \langle v, u \rangle + \mu \langle w, u \rangle} = \overline{\lambda} \langle u, v \rangle + \overline{\mu} \langle u, w \rangle$$

for all $u, v, w \in \mathcal{H}$, $\lambda, \mu \in \mathbb{C}$. Lastly, for the special case $\lambda = \mu = 0$ in (i), we get $\langle 0, u \rangle = \langle u, 0 \rangle = 0$ for all $u \in \mathcal{H}$. In particular, $\langle u, u \rangle = 0$ if and only if $u = 0$.

The most important examples are the following.

Example 1.14. (i) The space of complex numbers $\mathcal{H} = \mathbb{C}$, equipped with the inner product $\langle u, v \rangle := u\overline{v}$, is a pre-Hilbert space. More generally, the space $\mathbb{C}^n$ with the inner product defined as

$$\langle u, v \rangle := \sum_{i=1}^{n} u_i \overline{v_i}$$

for all $u = (u_1, \ldots, u_n), v = (v_1, \ldots, v_n) \in \mathbb{C}^n$, is a pre-Hilbert space.

(ii) Fix $(X, \mathcal{A}, \mu)$ a measure space, and let $\mathcal{H} = L^2(X, \mathcal{A}, \mu)$. For $f, g \in \mathcal{H}$, the formula

$$\langle f, g \rangle := \int_X f(x)\overline{g(x)}\, \mathrm{d}\mu(x)$$

defines a hermitian inner product on $\mathcal{H}$. If X is countable, we denote this space $\ell^2(X)$ rather than $L^2(X, \mathcal{A}, \mu)$.

Any space $\mathcal{H}$ with an inner product $\langle \cdot, \cdot \rangle$ can be turned into a normed space, by setting $\|u\| := \sqrt{\langle u, u \rangle}$. Indeed, the latter is well defined since $\langle u, u \rangle \geq 0$ for all $u \in \mathcal{H}$, and $\|u\| = 0$ if and only if $u = 0$. Moreover, for all $u \in \mathcal{H}$ and $\lambda \in \mathcal{H}$, we have

$$\|\lambda u\| = \sqrt{\langle \lambda u, \lambda u \rangle} = \sqrt{\lambda \overline{\lambda} \langle u, u \rangle} = |\lambda| \|u\|.$$

Then we are left to show the triangle inequality. This relies on the Cauchy-Schwarz inequality, of which it is difficult to underestimate the importance.

Lemma 1.15. Let $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ be a pre-Hilbert space.
Then, for all $u, v \in \mathcal{H}$, one has

$$|\langle u, v \rangle| \leq \|u\| \|v\|$$

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

Proof. The result is clear if $u = 0$ or $v = 0$. Then we may assume that $u, v \neq 0$ and, up to scaling, we can take $\|u\| = \|v\| = 1$. We start by observing that

$$\langle u - \langle u, v \rangle v, v \rangle = \langle u, v \rangle - \langle u, v \rangle \|v\|^2 = 0$$

and it follows that

$$\begin{aligned} 0 &\leq \|u - \langle u, v \rangle v\|^2 \\ &= \langle u, u - \langle u, v \rangle v \rangle \\ &= \langle u, u \rangle - \overline{\langle u, v \rangle} \langle u, v \rangle \\ &= 1 - |\langle u, v \rangle|^2. \end{aligned}$$

Hence $|\langle u, v \rangle| \leq 1 = \|u\| \|v\|$, and this proves the lemma. □

As announced, this gives the triangle inequality for the map $\| \cdot \|$ defined above.

Corollary 1.16. For any $u, v \in \mathcal{H}$, we have $\|u + v\| \leq \|u\| + \|v\|$.
In particular, the pair $(\mathcal{H}, \| \cdot \|)$ is a $\mathbb{C}$−normed vector space.

Proof. Let $u, v \in \mathcal{H}$. Expanding the square of the norm of $u + v$ and using Cauchy-Schwarz, we get

$$\begin{aligned} \|u + v\|^2 &= \langle u + v, u + v \rangle \\ &= \|u\|^2 + 2\mathrm{Re}\langle u, v \rangle + \|v\|^2 \\ &\leq \|u\|^2 + 2|\langle u, v \rangle| + \|v\|^2 \\ &\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \\ &= (\|u\| + \|v\|)^2 \end{aligned}$$

and so $\|u + v\| \leq \|u\| + \|v\|$. This yields the desired claim. □

Exercise 1.17. Prove that the inner product of a pre-Hilbert space is continuous in each variable, and that the norm is a continuous function.

Hence a pre-Hilbert space is also naturally a metric space.

Definition 1.18. A pre-Hilbert $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ is a Hilbert space if it is a Banach space for the norm $\| \cdot \| := \sqrt{\langle \cdot, \cdot \rangle}$.

Example 1.19. (i) $\mathbb{C}$ is complete, so it is a Hilbert space. More generally, $\mathbb{C}^n$ is a Hilbert space for all $n \geq 1$.

(ii) If $(X, \mathcal{A}, \mu)$ is a measure space, $L^2(X, \mathcal{A}, \mu)$ is a complex Hilbert space.

(iii) If $\mathcal{H}_1, \mathcal{H}_2$ are two Hilbert spaces, $\mathcal{H}_1 \times \mathcal{H}_2$ carries naturally the structure of a Hilbert space with the inner product $\langle \cdot, \cdot \rangle_{\mathcal{H}_1 \times \mathcal{H}_2}$ defined as

$$\langle (u_1, u_2), (v_1, v_2) \rangle_{\mathcal{H}_1 \times \mathcal{H}_2} := \langle u_1, v_1 \rangle_{\mathcal{H}_1} + \langle v_1, v_2 \rangle_{\mathcal{H}_2}$$

for all $u_1, v_1 \in \mathcal{H}_1$, $u_2, v_2 \in \mathcal{H}_2$.

The norm induced by an inner product has many properties, which are often useful. The following identities are known, respectively, as the *parallelogram law* and the *Pythagore's theorem*.

Proposition 1.20. For any $u, v \in \mathcal{H}$, we have $\|u+v\|^2 + \|u-v\|^2 = 2(\|u\|^2 + \|v\|^2)$. Moreover, if $\langle u, v \rangle = 0$, then $\|u + v\|^2 = \|u\|^2 + \|v\|^2$.

Proof. On one hand, we compute that

$$\|u + v\|^2 = \langle u + v, u + v \rangle = \|u\|^2 + \langle u, v \rangle + \langle v, u \rangle + \|v\|^2 = \|u\|^2 + 2\mathrm{Re}\langle u, v \rangle + \|v\|^2$$

while on the other hand,

$$\|u - v\|^2 = \langle u - v, u - v \rangle = \|u\|^2 - \langle u, v \rangle - \langle v, u \rangle + \|v\|^2 = \|u\|^2 - 2\mathrm{Re}\langle u, v \rangle + \|v\|^2.$$

Adding these two lines, the first claim follows. Pythagore's theorem is a consequence of

$$\|u + v\|^2 = \|u\|^2 + 2\mathrm{Re}\langle u, v \rangle + \|v\|^2$$

since the middle term of the right hand side vanishes if $\langle u, v \rangle = 0$. $\square$

Now we turn to the notion of *orthogonality*, which will play a prominent role in the sequel.

Two elements $u, v \in \mathcal{H}$ are *orthogonal* if $\langle u, v \rangle = 0$, and if $S \subset \mathcal{H}$ its orthogonal complement is defined as

$$S^\perp = \{u \in \mathcal{H} \mid \forall v \in S,\ \langle u, v \rangle = 0\}.$$

Note that S need not to be a vector subspace of $\mathcal{H}$. However, its orthogonal complement is always a subspace : if $u, v \in S^\perp$ and $\lambda \in \mathbb{C}$, then

$$\langle u + \lambda v, w \rangle = \langle u, w \rangle + \lambda \langle v, w \rangle = 0$$

for all $w \in S$, so $u + \lambda v \in S^\perp$. It is furthermore closed in $\mathcal{H}$, because if $(u_n)_{n\geq 0}$ in $S^\perp$ converges to $u \in \mathcal{H}$, it implies

$$\langle u, v\rangle = \langle \lim_{n\to\infty} u_n, v\rangle = \lim_{n\to\infty} \langle u_n, v\rangle = 0$$

for all $u \in S$, and thus $u \in S^\perp$.

Additionally, note that $S \cap S^\perp \subset \{0\}$. Indeed, if $u \in S \cap S^\perp$, then $\|u\|^2 = \langle u, u\rangle = 0$, so $u = 0$.

Orthogonal complements are useful because they provide a splitting of $\mathcal{H}$ as a direct sum.

Theorem 1.21. Let $\mathcal{H}$ be a Hilbert space, and $G \subset \mathcal{H}$ a closed subspace. Then every $w \in \mathcal{H}$ can be written uniquely as $w = u + v$ with $u \in G$ and $v \in G^\perp$.

In this case, we usually write $\mathcal{H} = G \oplus G^\perp$.

Proof. For the uniqueness part, suppose $w \in \mathcal{H}$ has two decompositions $w = u_1 + v_1 = u_2 + v_2$, $u_1, u_2 \in G$, with $v_1, v_2 \in G^\perp$. Then one gets

$$u_1 - u_2 = v_1 - v_2 \in G \cap G^\perp = \{0\}$$

so $u_1 = u_2$ and $v_1 = v_2$. Let's focus now on the existence part. For brievety, denote $\delta := \inf_{x\in G} \|w - x\|$. Let $(a_n)_{n\geq 0}$ be a sequence in G such that $\lim_{n\to\infty} \|w - a_n\| = \delta$. By the parallelogram law, for all $n, m \geq 0$, we compute

$$2\|w - a_n\|^2 + 2\|w - a_m\|^2 - \|a_n - a_m\|^2 = \|2w - a_n - a_m\|^2 = 4\left\|w - \frac{a_n + a_m}{2}\right\|^2 \geq 4\delta^2$$

and the last inequality holds by definition of δ, since $\frac{a_n+a_m}{2} \in G$. Hence we get

$$\|a_n - a_m\|^2 \leq 2\|w - a_n\|^2 + 2\|w - a_m\|^2 - 4\delta^2 \xrightarrow[n,m\to\infty]{} 0$$

by the choice of the sequence $(a_n)_{n\geq 0}$. This means $(a_n)_{n\geq 0}$ is Cauchy, and since $\mathcal{H}$ is complete, it then converges to $u \in \mathcal{H}$, which must be in G since it is a closed subspace. Now we show that $w - u \in G^\perp$, and $w = u + (w - u)$ will be the desired decomposition. Note that if $x \in G$ and $\lambda \in \mathbb{C}$, then $u + \lambda x \in G$, so

$$\|w - u\|^2 \leq \|w - u - \lambda x\|^2 = \|w - u\|^2 + |\lambda|^2\|x\|^2 - 2\mathrm{Re}\lambda\langle x, w - u\rangle$$

and this leads to

$$2\mathrm{Re}\lambda\langle x, w - u\rangle \leq |\lambda|^2\|x\|^2. \tag{1}$$

If $\lambda > 0$, then dividing by λ and taking the limit $\lambda \to 0$ provides $\mathrm{Re}\langle x, w - u\rangle \leq 0$. On the other hand, replacing λ by $-i\lambda$ in (3), taking $\lambda > 0$, dividing by it and letting $\lambda \to 0$ provides $\mathrm{Im}\langle x, w - u\rangle \leq 0$. G being a subspace, these two inequalities also holds for $-x$ instead of x. Finally, we conclude that

$$\mathrm{Re}\langle x, w - u\rangle = \mathrm{Im}\langle x, w - u\rangle = 0$$

and thus $\langle x, w - u\rangle = 0$ for all $x \in G$. This proves $w - u \in G^\perp$, as claimed. □

This decomposition has many consequences. The first we derive is the famous *Riesz representation theorem*.

Theorem 1.22. Let $\mathcal{H}$ be a Hilbert space, and $f \in \mathcal{H}^*$.
Then there exists a unique $u_0 \in \mathcal{H}$ so that

$$f(u) = \langle u, u_0 \rangle$$

for all $u \in \mathcal{H}$. Moreover, $\|f\| = \|u_0\|$.

Proof. If $f \equiv 0$, we choose $u_0 = 0$ and we are done. Assume now that f is a nontrivial functional. Since f is continuous, $\mathrm{Ker}(f)$ is a closed proper subspace of $\mathcal{H}$, and $\mathrm{Ker}(f)^\perp$ is not empty. Furthermore, it has dimension at least 1, since otherwise Theorem 1.21 would imply $\mathcal{H} = \mathrm{Ker}(f)$, contradicting the fact that f is nontrivial. Let $v_0 \in \mathrm{Ker}(f)^\perp$, with $\|v_0\| = 1$. Then every $u \in \mathcal{H}$ can be written

$$u = u - \langle u, v_0 \rangle v_0 + \langle u, v_0 \rangle v_0$$

and since $\langle u, v_0 \rangle v_0 \in \mathrm{Ker}(f)^\perp$, this forces $u - \langle u, v_0 \rangle v_0 \in \mathrm{Ker}(f)$. Thus $f(u - \langle u, v_0 \rangle v_0) = 0$ for all $u \in \mathcal{H}$, and it follows that

$$f(u) = f(\langle u, v_0 \rangle v_0) = \langle u, v_0 \rangle f(v_0) = \langle u, \overline{f(v_0)} v_0 \rangle$$

for all $u \in \mathcal{H}$. We set then $u_0 := \overline{f(v_0)} v_0$, and the first claim holds. The second follows, since

$$|f(u)| = |\langle u, u_0 \rangle| \leq \|u\| \|u_0\|$$

for every $u \in \mathcal{H}$ by Cauchy-Schwarz, giving $\|f\| \leq \|u_0\|$. For $u = u_0$, one has

$$|f(u_0)| = |\langle u_0, u_0 \rangle| = \|u_0\|^2 = \|u_0\| \|u_0\|$$

and thus $\|f\| = \|u_0\|$. This achieves the proof. □

Conversely one checks directly that for a fixed $u_0 \in \mathcal{H}$, the map

$$\begin{aligned} f : \mathcal{H} &\longrightarrow \mathbb{C} \\ u &\longmapsto \langle u, u_0 \rangle \end{aligned}$$

is a continuous linear functional. Together with Theorem 1.22, this defines an isometric isomorphism between $\mathcal{H}$ and $\mathcal{H}^*$.

We close this subsection by introducing different types of convergence for sequences of bounded linear operators. Fix then $\mathcal{H}$ a Hilbert space and $(A_n)_{n \in \mathbb{N}} \subset \mathcal{B}(\mathcal{H})$.

Definition 1.23. $(A_n)_{n\in\mathbb{N}}$ converges in norm to $A \in \mathcal{B}(\mathcal{H})$ if

$$\lim_{n\to\infty} \|A_n - A\| = 0.$$

Moreover, $(A_n)_{n\in\mathbb{N}}$ is strongly convergent if $(A_n u) \subset \mathcal{H}$ is convergent for all $u \in \mathcal{H}$.

Remark 1.24. The terminology here is confusing, since convergence in norm is stronger than strong convergent. Indeed a sequence $(A_n)_{n\in\mathbb{N}}$ of bounded operators converging to A is strongly convergent (check!).

1.3 Resolvent set and spectrum

The notions of resolvent set and spectrum for an operator generalize that of eigenvalues and eigenvectors for a finite dimensional matrix. The starting point is the following problem: given a normed space X, for which $\lambda \in \mathbb{C}$ does

$$(A - \lambda I)u = v$$

has a unique solution for any $v \in X$?

First of all, observe that the solution will be unique if $\mathrm{Ker}(A - \lambda I) = \{0\}$, *i.e.* if λ is *not* an eigenvalue of A. Next, note that the solution exists for any $v \in X$ provided $\mathrm{Im}(A - \lambda I) = X$. These two observations motivate the next definition.

Definition 1.25. Let X be a Banach space, and $A \in \mathcal{B}(X)$.
The resolvent set of A, denoted $\rho(A)$, is the set of $\lambda \in \mathbb{C}$ so that

$$\mathrm{Ker}(A - \lambda I) = \{0\},\ \overline{\mathrm{Im}(A - \lambda I)} = X \quad \text{and} \quad (A - \lambda I)^{-1} \text{ is bounded.}$$

The spectrum of A is then $\sigma(A) := \mathbb{C} \setminus \rho(A)$.

Remark 1.26. The same definition can be made more generally for bounded operators on a normed space X. However the next exercise gives a good reason to take X a Banach space.

Exercise 1.27. Let $A \in \mathcal{B}(X)$ and X be a Banach space. Prove that if $\lambda \in \rho(A)$, then $(A-\lambda I)^{-1}$ is defined on the whole space X. Conclude that when X is Banach, $\lambda \in \rho(A)$ if and only if $A - \lambda I$ is a bijective bounded operator.

From Definition 1.25 we see that $\sigma(A)$ can be decomposed into three parts, as

$$\sigma(A) = \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A)$$

where

(i) $\sigma_p(A)$ is the *point* spectrum of A. It consists of eigenvalues of A, *i.e.* of complex numbers λ so that $\mathrm{Ker}(A - \lambda I) \neq \{0\}$.

(ii) $\sigma_c(A)$ is the *continuous* spectrum of A. It consists of values of λ for which $\mathrm{Ker}(A - \lambda I) = \{0\}$, $\overline{\mathrm{Im}(A - \lambda I)} = \mathcal{H}$, but $(A - \lambda I)^{-1}$ is not bounded.

(iii) $\sigma_r(A)$ is the *residual* spectrum of A. It consists of values of λ for which $\mathrm{Ker}(A - \lambda I) = \{0\}$ but $\mathrm{Im}(A - \lambda I)$ is not dense in $\mathcal{H}$.

Here is a nice exercise to manipulate definitions of $\rho(A)$ and $\sigma(A)$.

Exercise 1.28. Let X be a Banach space, $\lambda \in \mathbb{C}$, and assume there is a sequence $(u_n)_{n\in\mathbb{N}} \subset X$ so that $\|u_n\| = 1$ and $Au_n - \lambda u_n \longrightarrow 0$ as $n \longrightarrow \infty$. Prove that $\lambda \in \sigma(A)$.

Definition 1.29. Let $A \in \mathcal{B}(\mathcal{H})$.
Its spectral radius, denoted $r(A)$, is defined as

$$r(A) := \sup_{\lambda \in \sigma(A)} |\lambda|.$$

Before looking at important examples, we establish basic properties of the spectrum of an operator.

Proposition 1.30. Let $A \in \mathcal{B}(\mathcal{H})$.
If $|\lambda| > \|A\|$, then $\lambda \in \rho(A)$. In particular, one has

$$\sigma(A) \subset \{\lambda \in \mathbb{C} : |\lambda| \leq \|A\|\}$$

and $r(A) \leq \|A\|$.

Proof. We start by proving the following claim: if $S \in \mathcal{B}(\mathcal{H})$ has $\|S\| < 1$, then $\mathrm{Id}_{\mathcal{H}} - S$ is invertible.

Indeed, suppose $\|S\| < 1$, and let $S_n := \sum_{k=0}^{n} S^k$. Then for $n, m \in \mathbb{N}$, $n \geq m$, one has

$$\|S_n - S_m\| = \left\| \sum_{k=m+1}^{n} S^k \right\| \leq \sum_{k=m+1}^{n} \|S\|^k$$

using the triangle inequality and the submultiplicativity of the norm. The right-hand side is the rest of a convergent series, since $\|S\| < 1$. We thus see that $\|S_n - S_m\| \longrightarrow 0$ as $n, m \to \infty$, *i.e.* $(S_n)_{n\in\mathbb{N}}$ is Cauchy in $\mathcal{B}(\mathcal{H})$. The latter being complete, $(S_n)_{n\in\mathbb{N}}$ converges in $\mathcal{B}(\mathcal{H})$, and we call T its limit. We compute then that

$$(\mathrm{Id}_{\mathcal{H}} - S)T = \lim_{n\to\infty} \left(\sum_{k=0}^{n} S^k - \sum_{k=0}^{n} S^{k+1} \right) = \lim_{n\to\infty} (\mathrm{Id}_{\mathcal{H}} - S^{n+1}) = \mathrm{Id}_{\mathcal{H}}$$

and similarly $T(\mathrm{Id}_{\mathcal{H}} - S) = \mathrm{Id}_{\mathcal{H}}$. Hence $\mathrm{Id}_{\mathcal{H}} - S$ is invertible and $(\mathrm{Id}_{\mathcal{H}} - S)^{-1} = T$.

This claim implies directly the proposition, because if $|\lambda| > \|A\|$, then $\|\frac{A}{\lambda}\| < 1$, so $\mathrm{Id}_{\mathcal{H}} - \frac{A}{\lambda}$ is invertible, and thus so is

$$-\lambda\left(\mathrm{Id}_{\mathcal{H}} - \frac{A}{\lambda}\right) = A - \lambda \mathrm{Id}_{\mathcal{H}}.$$

This implies that $\lambda \in \rho(A)$, and that $\sigma(A)$ is contained in the closed disk of radius $\|A\|$ centered at the origin, finishing the proof. □

This proposition together with the next one proves that the spectrum of A is always a *compact* subset of $\mathbb{C}$.

Proposition 1.31. Let $A \in \mathcal{B}(\mathcal{H})$. Then $\sigma(A) \subset \mathbb{C}$ is closed.

Proof. Let $\lambda \in \rho(A)$. For $\mu \in \mathbb{C}$ so that $|\mu - \lambda| < \frac{1}{\|(A-\lambda \mathrm{Id}_{\mathcal{H}})^{-1}\|}$, the operator $(\mu - \lambda)(A - \lambda \mathrm{Id}_{\mathcal{H}})^{-1} - \mathrm{Id}_{\mathcal{H}}$ is invertible by the claim in the proof of Proposition 1.30, and since

$$A - \mu \mathrm{Id}_{\mathcal{H}} = -(A - \lambda \mathrm{Id}_{\mathcal{H}})((\mu - \lambda)(A - \lambda \mathrm{Id}_{\mathcal{H}})^{-1} - \mathrm{Id}_{\mathcal{H}})$$

we deduce that $A - \mu \mathrm{Id}_{\mathcal{H}}$ is invertible. Thus $\rho(A)$ contains the open ball centered at λ of radius $\frac{1}{\|(A-\lambda \mathrm{Id}_{\mathcal{H}})^{-1}\|}$, and as this holds for any $\lambda \in \rho(A)$, it is an open set. Its complement $\sigma(A)$ is therefore closed in $\mathbb{C}$. □

Example 1.32. Consider $(\theta_n)_{n\geq 1} \subset \mathbb{C}$ a bounded sequence, and the operator A on ℓ^2 defined as

$$\begin{aligned} A\colon \ell^2 &\longrightarrow \ell^2 \\ u = (u_n)_{n\geq 1} &\longmapsto (\theta_n u_n)_{n\geq 1}. \end{aligned}$$

By assumption, there is $C > 0$ so that $|\theta_n| \leq C$ for every $n \geq 1$, and it follows that

$$\|Au\|_2^2 = \sum_{n\geq 1} |\theta_n u_n|^2 \leq C^2 \|u\|_2^2.$$

Hence $\|A\| \leq C$, and A is bounded. Let us find $\sigma(A)$. First observe that if $u = (\delta_{n,m})_{n\geq 1}$ then $Au = \theta_m u$, so that θ_m is an eigenvalue of A for all $m \geq 1$. Hence $\sigma(A) \supset \{\theta_n : n \geq 1\}$. Since the spectrum is closed (Proposition 1.31), we then deduce

$$\sigma(A) \supset \overline{\{\theta_n : n \geq 1\}}.$$

Now if $\lambda \notin \overline{\{\theta_n : n \geq 1\}}$, the sequence $(|\theta_n - \lambda|)_{n\geq 1}$ is bounded away from 0, so the equation $(A - \lambda I)u = v$ has a solution for any $v \in \ell^2$: consider for instance the sequence $(u_n)_{n\in\mathbb{N}}$ given by $u_n := \frac{1}{\theta_n - \lambda} v_n$ for all $n \geq 1$. This proves that $\sigma(A) \subset \overline{\{\theta_n : n \geq 1\}}$, and we conclude that

$$\sigma(A) = \overline{\{\theta_n : n \geq 1\}}.$$

Example 1.33. Let $a < b$ and t be any continuous function on $[0,1]$ whose range is $[a,b]$. Define an operator A on $C[0,1]$ by

$$\begin{aligned} A\colon C[0,1] &\longrightarrow C[0,1] \\ u &\longmapsto tu \end{aligned}$$

where $(tu)(x) := t(x)u(x)$. Since t is continuous on the compact set $[0,1]$, it is bounded on $[0,1]$, and there is $C > 0$ so that $|t(x)| \le C$ for any $x \in [0,1]$. This implies easily that $\|Au\|_\infty \le C\|u\|_\infty$ and thus A is bounded. Observe that $\lambda \notin [a,b]$ if and only if the equation $(A - \lambda I)u = v$ can be solved uniquely by the formula

$$u(x) = \frac{1}{t(x) - \lambda} v(x),\ x \in [0,1].$$

Thus $\sigma(A) = [a,b]$. This proves that any compact subset of $\mathbb{R}$ can be realized as the spectrum of a bounded operator.

1.4 Symmetric operators

Throughout this section, unless stated otherwise, $\mathcal{H}$ is a complex Hilbert space, and $A\colon \mathcal{H} \longrightarrow \mathcal{H}$ is a bounded linear operator on $\mathcal{H}$.

Fix $v \in \mathcal{H}$. Consider the linear functional φ defined as $\varphi(u) := \langle Au, v\rangle$, for all $u \in \mathcal{H}$. Since A and the first variable of the inner product are linear, φ is linear, and Cauchy-Schwarz inequality tells us it is bounded, as

$$|\varphi(u)| = |\langle Au, v\rangle| \le \|Au\|\|v\| \le \|A\|\|u\|\|v\|$$

for all $u \in \mathcal{H}$. Thus $\|\varphi\| \le \|A\|\|v\|$. Therefore, Riesz representation theorem (Theorem 1.22) gives the existence of a unique element A^*v of $\mathcal{H}$ so that $\varphi(u) = \langle u, A^*v\rangle$ for all $u \in \mathcal{H}$, *i.e.* $\langle Au, v\rangle = \langle u, A^*v\rangle$ for all $u \in \mathcal{H}$. Moreover, $\|\varphi\| = \|A^*v\|$. This correspondence defines a map

$$\begin{aligned} A^*\colon \mathcal{H} &\longrightarrow \mathcal{H} \\ v &\longmapsto A^*v \end{aligned}$$

and one easily checks that A^* is in fact linear. This motivates the next definition.

Definition 1.34. The operator $A^*\colon \mathcal{H} \to \mathcal{H}$ defined above, such that

$$\langle Au, v\rangle = \langle u, A^*v\rangle$$

for all $u, v \in \mathcal{H}$, is called the adjoint operator of A.

As a consequence of Riesz representation theorem, the adjoint A^* of A is the unique bounded linear operator satisfying the equality of Definition 1.34.

Here are the first general properties for computations with adjoint operators.

Proposition 1.35. (i) $\mathrm{Id}_{\mathcal{H}}^* = \mathrm{Id}_{\mathcal{H}}$, and $(A^*)^* = A$ for all $A \in \mathcal{B}(\mathcal{H})$.

(ii) $(A + \lambda B)^* = A^* + \overline{\lambda} B^*$ for all $A, B \in \mathcal{B}(\mathcal{H})$ and $\lambda \in \mathbb{C}$.

(iii) $(B \circ A)^* = A^* \circ B^*$ for all $A, B \in \mathcal{B}(\mathcal{H})$.

(iv) $\|A^*\| = \|A\|$, and $\|A^*A\| = \|A\|^2$ for all $A \in \mathcal{B}(\mathcal{H})$.

(v) If $A \in \mathcal{B}(\mathcal{H})$ is invertible and $A^{-1} \in \mathcal{B}(\mathcal{H})$, then A^* is invertible and $(A^*)^{-1} = (A^{-1})^*$.

Proof. (i) For any $u, v \in \mathcal{H}$, we have $\langle u, \mathrm{Id}_{\mathcal{H}}(v)\rangle = \langle u, v\rangle = \langle \mathrm{Id}_{\mathcal{H}}(u), v\rangle$, so necessarily $\mathrm{Id}_{\mathcal{H}}^* = \mathrm{Id}_{\mathcal{H}}$. In the same way, we compute that

$$\langle u, Av\rangle = \overline{\langle Av, u\rangle} = \overline{\langle v, A^*u\rangle} = \langle A^*u, v\rangle$$

which implies $A = (A^*)^*$.

(ii) Fix $u, v \in \mathcal{H}$, and observe that

$$\langle u, (A^* + \overline{\lambda} B^*)v\rangle = \langle u, A^*v\rangle + \lambda\langle u, B^*v\rangle = \langle Au, v\rangle + \lambda\langle Bu, v\rangle = \langle (A + \lambda B)u, v\rangle$$

by using properties of the inner product. Therefore, $A^* + \overline{\lambda} B^* = (A + \lambda B)^*$.

(iii) Here again, we have

$$\langle u, A^*(B^*v)\rangle = \langle Au, B^*v\rangle = \langle B(Au), v\rangle$$

for all $u, v \in \mathcal{H}$, implying $(B \circ A)^* = A^* \circ B^*$.

(iv) The paragraph preceding Definition 1.34 shows that $\|A^*v\| \leq \|A\|\|v\|$ for all $v \in \mathcal{H}$, giving the upper bound $\|A^*\| \leq \|A\|$. On the other hand, the same inequality with A^* instead of A provides

$$\|(A^*)^*\| \leq \|A^*\|$$

so by (i) we get in fact $\|A\| \leq \|A^*\|$. Henceforth, $\|A^*\| = \|A\|$.
For the last claim, let $u \in \mathcal{H}$ with $\|u\| = 1$. The definition of the operator norm provides

$$\|A^*Au\| \leq \|A^*\|\|Au\| \leq \|A^*\|\|A\|\|u\| = \|A\|^2$$

using $\|A^*\| = \|A\|$ in the last step. On the other hand, an application of Cauchy-Schwarz inequality shows that

$$\|Au\|^2 = \langle Au, Au\rangle = \langle u, A^*Au\rangle \leq |\langle u, A^*Au\rangle| \leq \|A^*Au\| \leq \|A^*A\|$$

providing the other bound $\|A\|^2 \leq \|A^*A\|$. This finishes the proof.

(v) Suppose that A is invertible and $A^{-1} \in \mathcal{B}(\mathcal{H})$. Fix $u, v \in \mathcal{H}$, and write $u = Az$, $v = A^*w$ for some $z, w \in \mathcal{H}$. One has

$$\langle u, (A^*)^{-1}v\rangle = \langle Az, w\rangle = \langle z, A^*w\rangle = \langle A^{-1}u, v\rangle$$

and this proves $(A^*)^{-1} = (A^{-1})^*$ as wanted. □

For our purposes, we will be interested in a special class of operators.

Definition 1.36. If $A \in \mathcal{B}(\mathcal{H})$ satisfies $A^* = A$, then A is called symmetric.

The following lemma is often also useful when working with adjoints. In particular, it will allow us to get a simple description of the residual spectrum of a bounded operator.

Lemma 1.37. Let $A \in \mathcal{B}(\mathcal{H})$. Then $\mathrm{Ker}(A^*) = \mathrm{Im}(A)^\perp$.
In particular, we have
$$\mathcal{H} = \mathrm{Ker}(A^*) \oplus \overline{\mathrm{Im}(A)}.$$

The next exercise will be used in the proof.

Exercise 1.38. Show that $(V^\perp)^\perp = \overline{V}$ for any subset $V \subset \mathcal{H}$. Next, prove that $V^\perp = \overline{V}^\perp$ for any subspace $V \subset \mathcal{H}$.

Proof. First, suppose $v \in \mathrm{Ker}(A^*)$. Fix $u \in \mathcal{H}$. Then we have
$$\langle Au, v\rangle = \langle u, A^* v\rangle = 0$$
showing that $v \in \mathrm{Im}(A)^\perp$. Conversely, if $v \in \mathrm{Im}(A)^\perp$, then $\langle Au, v\rangle = 0$ for all $u \in \mathcal{H}$. In other words, $\langle u, A^* v\rangle = 0$ for all $u \in \mathcal{H}$. Applying this inequality for the choice $u = A^* v$ yields to $\|A^*\|^2 = 0$. Hence $A^* v = 0$ and indeed $v \in \mathrm{Ker}(A^*)$.
For the second claim, note that $(\mathrm{Ker}(A^*))^\perp = (\mathrm{Im}(A)^\perp)^\perp = \overline{\mathrm{Im}(A)}$ by Exercise 1.38. Theorem 1.21, which we may apply since $\mathrm{Ker}(A^*)$ is a closed subspace of $\mathcal{H}$, then implies
$$\mathcal{H} = \mathrm{Ker}(A^*) \oplus (\mathrm{Ker}(A^*))^\perp = \mathrm{Ker}(A^*) \oplus \overline{\mathrm{Im}(A)}$$
as announced. This finishes the proof. □

Corollary 1.39. Let $A \in \mathcal{B}(\mathcal{H})$. Then it holds that
$$\sigma_r(A) = \{\lambda \in \mathbb{C} : \lambda \notin \sigma_p(A),\ \overline{\lambda} \in \sigma_p(A^*)\}.$$

Proof. By definition, and using Proposition 1.35 and Lemma 1.37, we see that
$$\begin{aligned}
\lambda \in \sigma_r(A) &\iff \mathrm{Ker}(A - \lambda I) = \{0\},\ \overline{\mathrm{Im}(A - \lambda I)} \subsetneq \mathcal{H} \\
&\iff \mathrm{Ker}(A - \lambda I) = \{0\},\ \mathrm{Ker}((A - \lambda I)^*) \supsetneq \{0\} \\
&\iff \mathrm{Ker}(A - \lambda I) = \{0\},\ \mathrm{Ker}(A^* - \overline{\lambda} I) \supsetneq \{0\} \\
&\iff \lambda \notin \sigma_p(A),\ \overline{\lambda} \in \sigma_p(A^*)
\end{aligned}$$
proving the announced equality. □

Exercise 1.40. Define the left and the right shift $S, T\colon \ell^2 \longrightarrow \ell^2$ by $(Su)_n := u_{n+1}$ and $(Tu)_1 := 0$, $(Tu)_n := u_{n-1}$, $n \geq 1$.
Show that S and T are bounded, and compute $\|S\|, \|T\|$. Determine S^*, T^*, and find $\sigma_p(S), \sigma_c(S), \sigma_r(S), \sigma_p(T), \sigma_c(T), \sigma_r(T)$.

Here is another subset of $\mathbb{C}$ associated to a bounded operator.

Definition 1.41. Let $A \in \mathcal{B}(\mathcal{H})$. Its numerical range is defined as

$$\mathrm{nr}(A) := \{\langle Au, u\rangle : \|u\| = 1\} \subset \mathbb{C}.$$

For instance, $\mathrm{nr}(0) = \{0\}$ while $\mathrm{nr}(\mathrm{Id}_{\mathcal{H}})=\{1\}$.

It turns out this subset provides a tool for checking a given operator is symmetric.

Exercise 1.42. Let $\mathcal{H}$ be a complex Hilbert space, and $A \in \mathcal{B}(\mathcal{H})$. Prove that if $\langle Au, u\rangle = 0$ for all $u \in \mathcal{H}$ then $A = 0$. What happens if $\mathcal{H}$ is rather a real Hilbert space?

Proposition 1.43. Let $A \in \mathcal{B}(\mathcal{H})$. Then A is symmetric if and only if $\mathrm{nr}(A) \subset \mathbb{R}$.

Proof. Suppose A is symmetric, and let $u \in \mathcal{H}$, $\|u\| = 1$. Then

$$\langle Au, u\rangle = \langle u, Au\rangle = \overline{\langle Au, u\rangle}$$

and so $\langle Au, u\rangle \in \mathbb{R}$. Hence $\mathrm{nr}(A) \subset \mathbb{R}$.

The other way around, let $v \in \mathcal{H} \setminus \{0\}$, and observe

$$\langle Av, v\rangle = \|v\|^2 \left\langle A\frac{v}{\|v\|}, \frac{v}{\|v\|}\right\rangle \in \mathbb{R}$$

since $\mathrm{nr}(A) \subset \mathbb{R}$ and $\frac{v}{\|v\|}$ has norm 1. It follows that

$$\langle Au, u\rangle = \overline{\langle Au, u\rangle} = \langle u, Au\rangle = \langle A^*u, u\rangle$$

for any $u \in \mathcal{H}$. This means that $\langle (A^* - A)u, u\rangle = 0$ for all $u \in \mathcal{H}$, and by Exercise 1.42 one gets $A^* - A = 0$, so A is symmetric. □

Symmetric operators also have an important property, an other way of computing their norms.

Theorem 1.44. Let $A \in \mathcal{B}(\mathcal{H})$ be a symmetric operator on $\mathcal{H}$.

Then, one has
$$\|A\| = \sup_{\|u\|=1} |\langle Au, u\rangle|.$$

Proof. Denote $s := \sup_{\|u\|=1} |\langle Au, u\rangle|$. Note that s also equals $\sup_{u\neq 0} \dfrac{|\langle Au, u\rangle|}{\langle u, u\rangle}$.

If $\|u\| = 1$, note that
$$|\langle Au, u\rangle| \leq \|A\|\|u\|^2 = \|A\|$$
by Cauchy-Schwarz. Hence $s \leq \|A\|$ already.

Conversely, let $u, v \in \mathcal{H}$. As A is symmetric, one has
$$\langle A(u+v), u+v\rangle - \langle A(u-v), u-v\rangle = 2(\langle Au, v\rangle + \overline{\langle Au, v\rangle}) = 4\mathrm{Re}\langle Au, v\rangle$$
and we can estimate
$$\begin{aligned} |\mathrm{Re}\langle Au, v\rangle| &\leq \frac{1}{4}|\langle A(u+v), u+v\rangle - \langle A(u-v), u-v\rangle| \\ &\leq \frac{1}{4}(s\|u+v\|^2 + s\|u-v\|^2) \\ &= \frac{s}{2}(\|u\|^2 + \|v\|^2) \end{aligned}$$
for any $u, v \in \mathcal{H}$. Now let $u \in \mathcal{H}$ with $\|u\| = 1$ and $v := \frac{Au}{\|Au\|}$. Then $\mathrm{Re}\langle Au, v\rangle$ reduces to $\|Au\|$, so that
$$\|Au\| = \mathrm{Re}\langle Au, v\rangle \leq \frac{s}{2}\left(\|u\|^2 + \left\|\frac{Au}{\|Au\|}\right\|^2\right) = s.$$
Hence $\|A\| \leq s$, and from the first part of the proof it follows that $\|A\| = s$. $\square$

Exercise 1.45. Prove that if $\mathcal{H}$ is a pre-Hilbert space, then
$$\langle u, v\rangle = \frac{1}{4}(\|u+v\|^2 - \|u-v\|^2 + i\|u+iv\|^2 - i\|u-iv\|^2)$$
for any $u, v \in \mathcal{H}$. Next prove that if A is an operator on a Hilbert space $\mathcal{H}$, then
$$\langle Au, v\rangle = \frac{1}{4}(\langle A(u+v), u+v\rangle - \langle A(u-v), u-v\rangle + i\langle A(u+iv), u+iv\rangle - i\langle A(u-iv), u-iv\rangle)$$
then for any $u, v \in \mathcal{H}$.

Exercise 1.46. Show that the subspace $\mathcal{S}(\mathcal{H})$ of symmetric operators is closed in $\mathcal{B}(\mathcal{H})$. Show also that if $S, B \in \mathcal{S}(\mathcal{H})$, then $SB \in \mathcal{S}(\mathcal{H})$ if and only if $SB = BS$.

Another fundamental property of symmetric operators is the localization of their spectrum.

Theorem 1.47. Let $A \in \mathcal{B}(\mathcal{H})$ be symmetric. Then $\sigma(A) \subset \mathbb{R}$.

The following result is needed in the proof.

Exercise 1.48. Let $A \in \mathcal{B}(\mathcal{H})$ be normal, *i.e.* $AA^* = A^*A$. Prove that A is invertible and has bounded inverse if and only if there exists a constant $C > 0$ so that $\|Au\| \geq C\|u\|$ for all $u \in \mathcal{H}$.
Hint: Prove first that A is normal if and only if $\|Au\| = \|A^*u\|$ for any $u \in \mathcal{H}$. Deduce that a normal operator is injective if and only if it has dense range.

Proof. Suppose $A \in \mathcal{B}(\mathcal{H})$ is symmetric, and let $\lambda = a + ib$ be a complex number with $b \neq 0$. We compute that

$$\begin{aligned}
\|(A-\lambda I)u\|^2 &= \langle (A-\lambda I)u, (A-\lambda I)u\rangle \\
&= \|Au\|^2 - (\lambda + \overline{\lambda})\langle Au, u\rangle + |\lambda|^2\|u\|^2 \\
&= \|Au\|^2 - 2a\langle Au, u\rangle + a^2\|u\|^2 + b^2\|u\|^2 \\
&= \|Au\|^2 - 2\mathrm{Re}\langle Au, au\rangle + \|au\|^2 + b^2\|u\|^2 \\
&= (\|Au + au\|)^2 + b^2\|u\|^2 \\
&\geq b^2\|u\|^2
\end{aligned}$$

with $b^2 > 0$ since $b \neq 0$. Above we used that $\langle Au, u\rangle$ is a real number, due to the fact that A is symmetric. By Exercise 1.48, since $A - \lambda I$ is normal, we deduce it is invertible and its inverse is bounded. Thus $\lambda \in \rho(A)$. In other words, we proved that $\mathbb{C} \setminus \mathbb{R} \subset \rho(A) = \mathbb{C} \setminus \sigma(A)$, which implies $\sigma(A) \subset \mathbb{R}$. □

Exercise 1.49. Prove directly that the eigenvalues (if any) of a symmetric operator $A \in \mathcal{B}(\mathcal{H})$ are real.

Here is an immediate consequence of our previous results.

Corollary 1.50. Let $A \in \mathcal{B}(\mathcal{H})$ be symmetric. Then $\sigma_r(A) = \emptyset$.

Proof. Combine Corollary 1.39 and Theorem 1.47. □

Note that the converse does not hold, as for instance the left shift on $\ell^2(\mathbb{N})$ has empty residual spectrum but is not symmetric.

Definition 1.51. Let $A \in \mathcal{B}(\mathcal{H})$ be symmetric.
The numbers

$$m := \inf_{\|u\|=1} \langle Au, u\rangle, \ M := \sup_{\|u\|=1} \langle Au, u\rangle$$

are called the lower bound and the upper bound of A.

An immediate consequence of Theorem 1.44 is that $\|A\| = \max\{|m|, |M|\}$.

These numbers provide an additional information on the localization of the spectrum.

Theorem 1.52. Let $A \in \mathcal{B}(\mathcal{H})$ be a symmetric operator. Then $\sigma(A) \subset [m, M]$.

Proof. Suppose that $\lambda \in \sigma(A)$. As A is symmetric, the converse statement of Exercise 1.28 holds (check!), and there exists a sequence $(u_n)_{n\in\mathbb{N}} \subset \mathcal{H}$ so that $\|u_n\| = 1$ for all $n \in \mathbb{N}$ and $\|(A - \lambda \mathrm{Id}_{\mathcal{H}})u_n\| \longrightarrow 0$ as $n \to \infty$. By Cauchy-Schwarz, it follows that

$$|\langle Au_n, u_n\rangle - \lambda| = |\langle (A - \lambda \mathrm{Id}_{\mathcal{H}})u_n, u_n\rangle| \leq \|(A - \lambda \mathrm{Id}_{\mathcal{H}})u_n\| \longrightarrow 0$$

whence $\lambda = \lim_{n\to\infty} \langle Au_n, u_n\rangle$. As $m \leq \langle Au_n, u_n\rangle \leq M$ for all $n \in \mathbb{N}$, we deduce that λ is real and that $\lambda \in [m, M]$. This concludes the proof. □

1.5 Projection operators

The main theme of the spectral theorem is to decompose any bounded symmetric operators as a linear combination of orthogonal projections. This part is devoted to this class of operators and their main properties.

Definition 1.53. Let $\mathcal{M} \subset \mathcal{H}$ be a closed subspace.
The orthogonal projection onto $\mathcal{M}$ is the map

$$\begin{aligned} P\colon \mathcal{H} &\longrightarrow \mathcal{H} \\ u &\longmapsto Pu := v \end{aligned}$$

where $u = v + w$, $v \in \mathcal{M}$, $w \in \mathcal{M}^\perp$ is the decomposition provided by Theorem 1.21.

One checks easily that it indeed defines a linear operator on $\mathcal{H}$.

Theorem 1.54. Let $\mathcal{M}, \mathcal{N}$ be closed subspaces of $\mathcal{H}$. Denote by P, Q the corresponding projections.

(i) $\mathcal{M} = \mathrm{Im}(P)$ and $\mathcal{M}^\perp = \mathrm{Ker}(P)$.

(ii) P is symmetric, $P^2 = P$ and $\|P\| = 1$.

(iii) $\mathrm{Id}_{\mathcal{H}} - P$ is the projection onto $\mathcal{M}^\perp$.

(iv) $\mathcal{M} \subset \mathcal{N}$ if and only if $PQ = QP = P$.

(v) $\langle Pu, u\rangle \leq \langle Qu, u\rangle$ for all $u \in \mathcal{H}$ if and only if $\mathcal{M} \subset \mathcal{N}$.

Proof. (i) These are straightforward consequences of the definition of an orthogonal projection.

(ii) Take $u, v \in \mathcal{H}$, and write $u = u_1 + u_2$, $v = v_1 + v_2$, $u_1, v_1 \in \mathcal{M}$, $v_1, v_2 \in \mathcal{M}^\perp$. Then

$$\langle Pu, v\rangle = \langle u_1, v\rangle = \langle u_1, v_1\rangle = \langle u, Pv\rangle$$

using that $\langle u_1, v_2\rangle = \langle u_2, v_1\rangle = 0$. Thus $P = P^*$. Also $P = P^2$ by definition of P. This implies

$$\|P\| = \|P^2\| \leq \|P\|^2$$

and then $1 \leq \|P\|$. Conversely, for $u \in \mathcal{H}$ the vectors Pu and $(I - P)u$ are orthogonal, and Pythagore's theorem (Proposition 1.20) gives

$$\|u\|^2 = \|Pu\|^2 + \|(I - P)u\|^2 \geq \|Pu\|^2$$

for any $u \in \mathcal{H}$. Hence $\|P\| \leq 1$ as announced.

(iii) For any $u \in \mathcal{H}$, we have a unique decomposition $u = Pu + u'$ with $u' \in \mathcal{M}^\perp$, so that $(I - P)u = u'$ is the projection on the second component of u, *i.e.* on $\mathcal{M}^\perp$.

(iv) Suppose $\mathcal{M} \subset \mathcal{N}$. It suffices to write

$$PQ = P \iff P(Q - \mathrm{Id}_{\mathcal{H}}) = 0 \iff P(\mathrm{Id}_{\mathcal{H}} - Q) = 0 \iff \mathrm{Im}(\mathrm{Id}_{\mathcal{H}} - Q) \subset \mathrm{Ker}(P).$$

Combining (i) and (iii), the latter condition means $\mathcal{N}^\perp \subset \mathcal{M}^\perp$, which holds since $\mathcal{M} \subset \mathcal{N}$. The same reasoning proves $QP = P$.

Conversely, suppose $PQ = QP = P$. Let $u \in \mathcal{M} = \mathrm{Im}(P)$, so there is $v \in \mathcal{H}$ so that $u = Pv$. This reads as $u = Pv = QPv$, so $u \in \mathrm{Im}(Q) = \mathcal{N}$.

(v) Suppose first that $\mathcal{M} \subset \mathcal{N}$. Let $u \in \mathcal{H}$, and write it as $u = v + v'$ with $v \in \mathcal{N}$ and $v' \in \mathcal{N}^\perp$. As $\mathcal{M} \subset \mathcal{N}$, $\mathcal{N}^\perp \subset \mathcal{M}^\perp$ and the latter is the kernel of P, whence $Pv' = 0$. Hence

$$\langle Pu, u\rangle = \langle Pv, v + v'\rangle = \langle Pv, v\rangle + \langle Pv, v'\rangle = \langle Pv, v\rangle.$$

Also $\langle Qu, u\rangle = \langle v, v + v'\rangle = \|v\|^2$, and thus

$$\langle Pu, u\rangle = \langle Pv, v\rangle \leq \|P\|\|v\|^2 = \|v\|^2 = \langle Qu, u\rangle$$

by the Cauchy-Schwarz inequality.

Conversely, suppose that $\langle Pu, u\rangle \leq \langle Qu, u\rangle$ for any $u \in \mathcal{H}$. As $\mathcal{M} = \mathrm{Im}(P)$ and $\mathcal{N} = \mathrm{Im}(Q)$, the inclusion to show is equivalent to

$$\mathrm{Im}(P) \subset \mathrm{Im}(Q) \iff \mathrm{Im}(Q)^\perp \subset \mathrm{Im}(P)^\perp \iff \mathrm{Ker}(Q) \subset \mathrm{Ker}(P).$$

Let thus $u \in \mathrm{Ker}(Q)$, so $Qu = 0$. This implies $\langle Pu, u\rangle \leq \langle Qu, u\rangle = 0$, and $\langle Pu, u\rangle \geq 0$, whence in fact $\langle Pu, u\rangle = 0$. We can conclude that

$$\|Pu\|^2 = \langle Pu, Pu\rangle = \langle Pu, u\rangle = 0$$

using the symmetry and the idempotence of P for the second equality. Hence $Pu = 0$, and $u \in \mathrm{Ker}(P)$, as wished. □

For orthogonal projections, the numerical range introduced above can be computed easily.

Proposition 1.55. Suppose that $P \notin \{0, \mathrm{Id}_{\mathcal{H}}\}$ is an orthogonal projection. Then $\mathrm{nr}(P) = [0,1]$.

Proof. First of all, let $u \in \mathcal{H}$ with $\|u\| = 1$. We have

$$\langle Pu, u\rangle = \langle Pu, Pu + (\mathrm{Id}_{\mathcal{H}} - P)u\rangle = \langle Pu, Pu\rangle = \|Pu\|^2 \geq 0$$

and also $\langle Pu, u\rangle \leq \|Pu\|\|u\| \leq \|P\| = 1$. Hence $\mathrm{nr}(P) \subset [0,1]$.

Conversely, fix $t \in [0,1]$. Let $v \in \mathrm{Im}(P)$ and $w \in \mathrm{Ker}(P) = (\mathrm{Im}(P))^{\perp}$ with $\|v\| = \|w\| = 1$. Consider $u_t = tv + \sqrt{1-t^2}w$. Then

$$\|u_t\| = \|tv\|^2 + \left\|\sqrt{1-t^2}w\right\|^2 = t^2 + (1-t^2) = 1$$

and also $\langle Pu_t, u_t\rangle = \langle tv, tv + \sqrt{1-t^2}w\rangle = t^2$. This proves that $t^2 \in \mathrm{nr}(P)$ for any $t \in [0,1]$, and since the square is a bijection from $[0,1]$ onto $[0,1]$, this shows that $[0,1] \subset \mathrm{nr}(P)$. Hence $\mathrm{nr}(P) = [0,1]$ as claimed. □

1.6 Positive operators

Definition 1.56. An operator $P \in \mathcal{B}(\mathcal{H})$ is positive if $\langle Pu, u\rangle \geq 0$ for all $u \in \mathcal{H}$.

Note that a positive operator is automatically symmetric.

Also, if P, Q are symmetric, we can compare them, saying that P is smaller than Q, denoted $P \leq Q$, if $Q - P$ is positive.

Exercise 1.57. Check that $\leq$ is a partial order on the class of symmetric operators on $\mathcal{H}$.

Example 1.58. (i) Let A be the operator on $L^2([0,1])$ defined by $(Au)(t) = tu(t)$. We have that

$$\langle Au, u\rangle = \int_0^1 tu(t)\overline{u(t)}\,\mathrm{d}t = \int_0^1 t|u(t)|^2\,\mathrm{d}t \geq 0$$

for any $u \in L^2([0,1])$, so A is positive.

(ii) If $A \in \mathcal{B}(\mathcal{H})$ then A^*A is positive, because

$$\langle A^*Au, u\rangle = \langle Au, Au\rangle = \|Au\|^2 \geq 0$$

for any $u \in \mathcal{H}$. The same applies for AA^*. We will prove below that in fact all positive operators arise in this form.

(iii) Likewise, if A is positive, then A^* is positive since

$$\langle A^*u, u\rangle = \langle u, Au\rangle = \overline{\langle Au, u\rangle} = \langle Au, u\rangle \geq 0$$

for any $u \in \mathcal{H}$. Also, if A is invertible, then A^{-1} is positive.

(iv) An orthogonal projection is always positive, since equals to its square.

Exercise 1.59. Deduce from the proof of Theorem 1.47 that if A is a positive operator, then $\sigma(A) \subset [0, \infty)$.

Definition 1.60. Let $P \in \mathcal{B}(\mathcal{H})$ be positive.
A square root of P is a bounded operator R so that $R^2 = P$.

When it exists, the square root of an operator P is denoted $\sqrt{P}$, or $P^{1/2}$.

The main goal of this subsection is to prove that any positive operator has a unique positive square root, as for real numbers. To prove this claim, we need to go through some technical lemmas. The first one is the following.

Lemma 1.61. Let $A, S \in \mathcal{B}(\mathcal{H})$, and suppose $AS = SA$.
Then any polynomial in A commutes with S. Moreover if A is symmetric, then any polynomial in A is symmetric.

Proof. Since a polynomial with real coefficients is a linear combination of monomials, it is enough to prove both claims for monomials of the type A^j, $j \geq 0$. This is straightforward, as

$$A^j S = A^{j-1}AS = A^{j-1}SA = A^{j-2}SA^2 = \cdots = SA^j.$$

If A is symmetric, then for $u, v \in \mathcal{H}$ one has

$$\langle A^j u, v\rangle = \langle A^{j-1}u, Av\rangle = \cdots = \langle u, A^j v\rangle$$

which proves that A^j is symmetric. Thus we are done. □

Lemma 1.62. Let $P \in \mathcal{B}(\mathcal{H})$ be positive.
There exists a sequence $(P_n)_{n\geq 1}$ of polynomials in P so that the sequence of partial sums

$$\left(\sum_{k=1}^{n} P_n^2\right)_{n\geq 1}$$

converges strongly to P, *i.e.* $\lim_{n\to\infty} \left\| Pu - \sum_{k=1}^{n} P_n^2 u \right\| = 0$ for all $u \in \mathcal{H}$.

Proof. Suppose $P \neq 0$, otherwise the result is clear. Define $B_1 := \frac{1}{\|P\|}P$ and by induction

$$B_{n+1} = B_n - B_n^2, \; n \geq 1.$$

By Lemma 1.61, each B_n is symmetric since it is a polynomial in P. We now prove that $0 \leq B_n \leq I$ for all $n \geq 1$, by induction. For the case $n = 1$, we compute

$$\langle B_1 u, u\rangle = \frac{1}{\|P\|}\langle Pu, u\rangle \geq 0$$

for any $u \in \mathcal{H}$ since $P \geq 0$, and also

$$\langle (I - B_1)u, u\rangle = \langle u, u\rangle - \langle B_1 u, u\rangle = \langle u, u\rangle - \frac{1}{\|P\|}\langle Pu, u\rangle \geq 0$$

for any $u \in \mathcal{H}$, by Cauchy-Schwarz inequality, so $B_1 \leq I$. Now suppose $0 \leq B_m \leq I$ for some $m \geq 1$. Observe that

$$B_{m+1} = B_m - B_m^2 = B_m(I - B_m)^2 + B_m^2(I - B_m).$$

Now fix $u \in \mathcal{H}$. Then we have

$$\begin{aligned}\langle B_m(I - B_m)^2 u, u\rangle &= \langle (I - B_m)B_m(I - B_m)u, u\rangle \\ &= \langle B_m(I - B_m)u, (I - B_m)u\rangle \geq 0\end{aligned}$$

using the fact that B_m commutes with $I - B_m$, that $I - B_m$ is symmetric and that B_m is positive by assumption. Thus $B_m(I - B_m)^2$ is positive. Likewise

$$\begin{aligned}\langle B_m^2(I - B_m)u, u\rangle &= \langle B_m(I - B_m)B_m u, u\rangle \\ &= \langle (I - B_m)B_m u, B_m u\rangle \geq 0\end{aligned}$$

since B_m is symmetric and $I - B_m$ is positive. Hence $B_m^2(I - B_m) \geq 0$ as well, and B_{m+1} is positive. On the other hand, $I - B_{m+1} = (I - B_m) + B_m^2$ is the sum of two positive operators, so is positive. This concludes the inductive step, and thus

$$\forall n \geq 1, \; 0 \leq B_n \leq I.$$

Now we observe that $\sum_{k=1}^{n} B_k^2 = \sum_{k=1}^{n}(B_k - B_{k+1}) = B_1 - B_{n+1}$, so that

$$\begin{aligned}\sum_{k=1}^{n}\langle B_k u, B_k u\rangle &= \sum_{k=1}^{n}\langle B_k^2 u, u\rangle \\ &= \Big\langle \sum_{k=1}^{n} B_k^2 u, u\Big\rangle \\ &= \langle B_1 u, u\rangle - \langle B_{n+1}u, u\rangle \\ &\leq \langle B_1 u, u\rangle\end{aligned}$$

for all $u \in \mathcal{H}$. In particular, we see that $\sum_{k=1}^{\infty} \|B_k^2 u\|$ converges, so $\lim_{k\to\infty} \|B_k^2 u\| = 0$ for all $u \in \mathcal{H}$. It thus follows that

$$\left\| \sum_{k=1}^{n} B_k^2 u - B_1 u \right\| = \|B_{n+1} u\| \xrightarrow{n\to\infty} 0.$$

for all $u \in \mathcal{H}$. Letting $P_n = \sqrt{\|P\|} B_n$, we have

$$\lim_{k\to\infty} \left\| Pu - \sum_{k=1}^{n} P_k^2 u \right\| = \lim_{k\to\infty} \left\| \|P\| B_1 u - \sum_{k=1}^{n} \|P\| B_k^2 u \right\| = 0$$

for all $u \in \mathcal{H}$, as claimed. This finishes the proof. □

The first corollary we derive and that we will use for the proof of the existence of square roots is the stability under multiplication of the class of positive operators, provided they commute.

Corollary 1.63. Let $P, Q \in \mathcal{B}(\mathcal{H})$ be positive, and suppose that $PQ = QP$. Then PQ is positive.

Proof. Let $(P_n)_{n\geq 1} \subset \mathcal{B}(\mathcal{H})$ be the sequence given by the previous lemma. Since each P_n is a polynomial in P, and that P commutes with Q, $P_n Q = Q P_n$ for all $n \geq 1$. This implies that $\langle P_n^2 Q u, u\rangle = \langle Q P_n u, P_n u\rangle$ for any $u \in \mathcal{H}$ and any $n \geq 1$, using that P_n is symmetric. Hence, from the continuity of the inner product in the first variable (Exercise 1.17) we get

$$\langle PQu, u\rangle = \left\langle \sum_{n=1}^{\infty} P_n^2 Q u, u \right\rangle = \sum_{n=1}^{\infty} \langle P_n^2 Q u, u\rangle$$

for all $u \in \mathcal{H}$, meaning that PQ is positive. □

Exercise 1.64. Show that the above result is false if P and Q do not commute.

The next theorem we derive is an analog of the dominated convergence theorem in analysis.

Theorem 1.65. Let $(S_n)_{n\in\mathbb{N}} \subset \mathcal{B}(\mathcal{H})$ be symmetric operators, and $T \in \mathcal{B}(\mathcal{H})$ be a symmetric operator so that

(i) $S_m \leq S_n$ for any $0 \leq m \leq n$.

(ii) $S_n S_m = S_m S_n$ for any $n, m \in \mathbb{N}$.

(iii) $S_n T = T S_n$ for any $n \in \mathbb{N}$.

(iv) $S_n \leq T$ for any $n \in \mathbb{N}$.

Then $(S_n)_{n\in\mathbb{N}}$ converges strongly to a symmetric operator $S \in \mathcal{B}(\mathcal{H})$.

Proof. Let $P_n := T - S_n$, $n \in \mathbb{N}$. By condition (iv), $P_n \geq 0$ for all $n \in \mathbb{N}$, and by (ii) and (iii), $P_n P_m = P_m P_n$ for all $n, m \in \mathbb{N}$. By (i), one has

$$P_n - P_m = (T - S_n) - (T - S_m) = S_m - S_n \leq 0$$

if $0 \leq m \leq n$. Thus $(P_n)_{n\in\mathbb{N}}$ is decreasing. By Corollary 1.63, we then have that $P_m(P_m - P_n) \geq 0$ and $P_n(P_m - P_n) \geq 0$ if $0 \leq m \leq n$, and it follows that

$$P_m^2 \geq P_n P_m \geq P_n^2$$

for $0 \leq m \leq n$, meaning that $(P_n^2)_{n\in\mathbb{N}}$ is decreasing. Hence if $u \in \mathcal{H}$, the sequence $(\langle P_n^2 u, u\rangle)_{n\in\mathbb{N}}$ is a decreasing sequence of positive numbers, and therefore it has a limit, that we call α_u. The above inequalities then provides that

$$\langle P_m P_n u, u\rangle \longrightarrow \alpha_u$$

as $n \to \infty$. In turn, this implies that

$$\begin{aligned}
\|S_n u - S_m u\|^2 &= \|P_n u - P_m u\|^2 \\
&= \langle (P_n - P_m)u, (P_n - P_m)u\rangle \\
&= \langle (P_n - P_m)^2 u, u\rangle \\
&= \langle P_n^2 u, u\rangle + \langle P_m^2 u, u\rangle - 2\langle P_n P_m u, u\rangle \\
&\longrightarrow \alpha_u + \alpha_u - 2\alpha_u = 0
\end{aligned}$$

as $n \to \infty$. Hence $(S_n u)_{n\in\mathbb{N}}$ is Cauchy in $\mathcal{H}$, and by completeness it must converges. Call the limit Su. This defines a linear operator on $\mathcal{H}$, which is bounded by Banach-Steinhaus (Theorem 1.8). S is also symmetric as the inner product is continuous in each variable (Exercise 1.17). □

We can now proceed to show the existence of square roots for positive operators.

Theorem 1.66. Let $P \in \mathcal{B}(\mathcal{H})$ be positive.
Then P has a unique positive square root, which commutes with any $A \in \mathcal{B}(\mathcal{H})$ commuting with P.

Proof. Without loss of generality, we may suppose $0 \leq P \leq I$. Set $R_0 = 0$ and

$$R_{n+1} := R_n + \frac{1}{2}(P - R_n^2), \ n \in \mathbb{N}.$$

Note that R_n is a polynomial in P for all $n \in \mathbb{N}$, so R_n is symmetric for any $n \in \mathbb{N}$ by Lemma 1.61. Now note the relation

$$I - R_{n+1} = \frac{1}{2}(I - R_n)^2 + \frac{1}{2}(I - P). \tag{2}$$

By induction using (2), one sees directly that $R_n \leq I$ for any $n \in \mathbb{N}$. Furthermore, since

$$\begin{aligned} R_{n+1} - R_n &= (I - R_n) - (I - R_{n+1}) \\ &= \frac{1}{2}(I - R_{n-1})^2 + \frac{1}{2}(I - P) - \left(\frac{1}{2}(I - R_n)^2 + \frac{1}{2}(I - P)\right) \\ &= \frac{1}{2}(I - R_{n-1})^2 - \frac{1}{2}(I - R_n)^2 \\ &= \frac{1}{2}(R_n - R_{n-1})((I - R_{n-1}) + (I - R_n)) \end{aligned}$$

an immediate induction proves that $R_{n+1} \geq R_n$ for all $n \in \mathbb{N}$. We have $R_0 = 0$, so in particular each R_n is positive. To sum up, $(R_n)_{n\in\mathbb{N}}$ is an increasing sequence of positive symmetric operators, bounded from above by I, and pairwise commuting. Theorem 1.65 therefore ensures the existence a symmetric operator $R \in \mathcal{B}(\mathcal{H})$ so that

$$R_n u \longrightarrow Ru$$

for any $u \in \mathcal{H}$. Additionally, by Banach-Steinhaus there is $C > 0$ so that $\|R_n\| \leq C$, whence

$$\begin{aligned} \|R_n^2 u - R^2 u\| &= \|R_n^2 u - R_n(Ru) + R_n(Ru) - R^2 u\| \\ &\leq \|R_n(R_n u) - R_n(Ru)\| + \|R_n(Ru) - R(Ru)\| \\ &\leq C\|R_n u - Ru\| + \|R_n(Ru) - R(Ru)\| \end{aligned}$$

for any $u \in \mathcal{H}$. Both terms go to 0 as $(R_n)_{n\in\mathbb{N}}$ converges strongly to R. This proves that $(R_n^2)_{n\in\mathbb{N}}$ converges strongly to R^2. Passing to the limit in the recursive relation defining R_n, we get $R^2 = P$ as wanted. Since R is the strong limit of $(R_n)_{n\in\mathbb{N}}$ and these are positive operators, R is positive. Additionally, if $A \in \mathcal{B}(\mathcal{H})$ commutes with P, then A commutes with R_n for any $n \in \mathbb{N}$, because these are polynomials in P. It follows that

$$ARu = A(\lim_{n\to\infty} R_n u) = \lim_{n\to\infty} AR_n u = \lim_{n\to\infty} R_n Au = RAu$$

for any $u \in \mathcal{H}$, showing the second claim.

We can now prove uniqueness. Suppose that there is another bounded positive operator S so that $S^2 = P$. Let $u \in \mathcal{H}$, and $v := (R - S)u$. Then

$$\begin{aligned} \langle Rv, v\rangle + \langle Sv, v\rangle &= \langle R(R - S)u, v\rangle + \langle S(R - S)u, v\rangle \\ &= \langle (R + S)(R - S)u, v\rangle \\ &= \langle (R^2 - S^2)u, v\rangle \end{aligned}$$

$$= 0$$

and since $\langle Rv, v\rangle, \langle Sv, v\rangle \geq 0$, this forces $\langle Rv, v\rangle = \langle Sv, v\rangle = 0$. Now if T is a positive square root of R, then $\|Tv\|^2 = \langle Tv, Tv\rangle = \langle T^2v, v\rangle = \langle Rv, v\rangle = 0$, whence $Tv = 0$ and $Rv = T(Tv) = 0$. Likewise, by considering a square root of S, we get $Sv = 0$. Finally, this implies

$$\begin{aligned}\|Ru - Su\|^2 &= \langle (R-S)u, (R-S)u\rangle \\ &= \langle (R-S)v, u\rangle \\ &= \langle Rv, u\rangle + \langle Sv, u\rangle \\ &= 0\end{aligned}$$

for any $u \in \mathcal{H}$ and we get $R = S$. This concludes the proof. □

As promised, the following exercise completes Example 1.58(ii).

Exercise 1.67. Prove that an operator $P \in \mathcal{B}(\mathcal{H})$ is positive if and only if there exists $A \in \mathcal{B}(\mathcal{H})$ so that $P = A^*A$.

Exercise 1.68. Let $A\colon \ell^2 \longrightarrow \ell^2$ be the *double right shift*, defined by $(Au)_1 = (Au)_2 := 0$ and $(Au)_n := u_{n-2}$, $n \geq 3$.
Show that A is bounded and compute $\|A\|$. Determine A^*, and find $\sigma_p(A)$, $\sigma_c(A)$ and $\sigma_r(A)$. Is A positive? Find $B\colon \ell^2 \longrightarrow \ell^2$ so that $A = B^2$. What can we conclude?

Definition 1.69. Let $S \in \mathcal{B}(\mathcal{H})$ be symmetric.
The absolute value of S is defined as $|S| := \sqrt{S^2}$.

To prove properties of the absolute value, we will make use of the following more general lemma.

Lemma 1.70. Let $S, T \in \mathcal{B}(\mathcal{H})$ be symmetric, $ST = TS$, and $S^2 = T^2$. Let P be the projection onto $\operatorname{Ker}(S-T)$. Then

(i) If $A \in \mathcal{B}(\mathcal{H})$ commutes with $S - T$, then $AP = PA$.

(ii) $Su = 0$ implies $Pu = u$.

(iii) $P(S+T) = S+T$ and $P(S-T) = 0$.

Proof. (i) Suppose that $A \in \mathcal{B}(\mathcal{H})$ commutes with $S - T$. According to Theorem 1.21, we have the splitting

$$\mathcal{H} = \operatorname{Ker}(S-T) \oplus \operatorname{Ker}(S-T)^\perp.$$

Fix then $v \in \mathcal{H}$, and write $v = v_1 + v_2$, $v_1 \in \operatorname{Ker}(S-T)$, $v_2 \in \operatorname{Ker}(S-T)^\perp$. Then $Av = Av_1 + Av_2$ by linearity, and

$$(S-T)Av_1 = A(S-T)v_1 = A(0) = 0$$

since A commutes with $S - T$ and $v_1 \in \operatorname{Ker}(S - T)$. Also, if $w \in \operatorname{Ker}(S - T)$, one has that

$$(S - T)A^*w = (S^* - T^*)A^*w = (AS - AT)^*w = (SA - TA)^*w = A^*(S - T)w = 0$$

showing that $A^*w \in \operatorname{Ker}(S - T)$, and thus

$$\langle Av_2, w\rangle = \langle v_2, A^*w\rangle = 0$$

since $v_2 \in \operatorname{Ker}(S - T)^\perp$ and $A^*w \in \operatorname{Ker}(S - T)$. This implies that $Av_2 \in \operatorname{Ker}(S - T)^\perp$, and $Av = Av_1 + Av_2$ is the unique decomposition of Av as a sum of an element of $\operatorname{Ker}(S - T)$ and an element of its orthogonal. Hence $APv = Av_1 = PAv$, whence $AP = PA$ as claimed.

(ii) Suppose $Su = 0$. We first observe that

$$\|Tu\|^2 = \langle Tu, Tu\rangle = \langle T^2u, u\rangle = \langle S^2u, u\rangle = \langle Su, Su\rangle = \|Su\|^2$$

using symmetry of S, T and $S^2 = T^2$. As $Su = 0$, it implies also that $Tu = 0$, whence $(S - T)u = Su - Tu = 0$. This means $u \in \operatorname{Ker}(S - T)$, and thus is invariant by P, so $Pu = u$.

(iii) Likewise, for the first equality it is enough to check that vectors of the form $(S+T)u$ are in $\operatorname{Ker}(S - T)$. As $ST = TS$ and $S^2 = T^2$, $(S - T)(S + T) = (S + T)(S - T) = S^2 - T^2 = 0$, whence $(S - T)(S + T)u = 0$ for any $u \in \mathcal{H}$, providing $P(S + T) = S + T$. Also $(S - T)u \in \operatorname{Ker}(S - T)^\perp$ for any $u \in \mathcal{H}$, as

$$\langle v, (S - T)u\rangle = \langle (S - T)^*v, u\rangle = \langle (S - T)v, u\rangle = 0$$

for all $v \in \operatorname{Ker}(S - T)$, because S, T are symmetric. Since $\operatorname{Ker}(S - T)^\perp = \operatorname{Ker}(P)$, we conclude $P(S - T) = 0$. □

We have then the next result, about the absolute value of an operator.

Lemma 1.71. Let $S \in \mathcal{B}(\mathcal{H})$ be symmetric. Let E_+ be the projection onto $\operatorname{Ker}(S - |S|)$. Then

(i) If $A \in \mathcal{B}(\mathcal{H})$ commutes with S, then $AE_+ = E_+A$.

(ii) $Su = 0$ implies $E_+u = u$.

(iii) $SE_+ \geq 0$ and $S(I - E_+) \leq 0$.

Proof. (i) By Lemma 1.70(i), it is enough to check that A commutes with $S - |S|$. But A commutes with S, so also with S^2, and then also with $|S| = \sqrt{S^2}$ by Theorem 1.66. Thus A commutes with $S - |S|$, as wanted.

(ii) Directly follows from Lemma 1.70(ii).

(iii) By (iii) of Lemma 1.70, we have $E_+(S - |S|) = 0$, and by (i) of the present proof, $S - |S|$ commutes with E_+, so $(S - |S|)E_+ = 0$, whence $SE_+ = |S|E_+$. Now E_+ and $|S|$ are both positive and commute, so Corollary 1.63 tells us that $|S|E_+ \geq 0$. Now combining $E_+(S - |S|) = 0$ and $E_+(S + |S|) = S + |S|$, we have $2E_+|S| = S + |S|$, *i.e.* $S = (2E_+ - I)|S|$. This yields to

$$S(I - E_+) = S - SE_+ = (2E_+ - I)|S| - SE_+ = -(I - E_+)|S|.$$

Now as above $(I - E_+)|S| \geq 0$, so $S(I - E_+) \leq 0$. This finishes the proof. □

2. The spectral theorem for bounded symmetric operators

In this chapter we establish the first goal of the course, namely the spectral decomposition of bounded symmetric operators on Hilbert spaces.

2.1 Integration with respect to a spectral family

Definition 2.1. A spectral family is a mapping $E\colon \mathbb{R} \longrightarrow \mathcal{B}(\mathcal{H})$, denoted $(E_\lambda)_{\lambda\in\mathbb{R}}$, so that

(i) E_λ is a projection for all $\lambda \in \mathbb{R}$.

(ii) $\lambda < \mu \Longrightarrow E_\lambda \le E_\mu$.

(iii) E is strongly left-continuous, *i.e.*

$$\forall u \in \mathcal{H},\ \forall \mu \in \mathbb{R},\ \lim_{\lambda\to\mu^-} E_\lambda u = E_\mu u.$$

(iv) There exists $m, M \in \mathbb{R}$, $m < M$, so that $E_\lambda = 0$ if $\lambda < m$ and $E_\lambda = \mathrm{Id}_{\mathcal{H}}$ if $\lambda > M$.

We now consider $f \in C([m, M+\varepsilon], \mathbb{C})$ and we extend it continuously to $[m, M+\varepsilon]$, for some $0 < \varepsilon < 1$.

Let $\Pi = (\lambda_k)_{k=0}^n$ be a partition of $[m, M+\varepsilon]$, *i.e.* a subdivision $m = \lambda_0 < \lambda_1 < \cdots < \lambda_k = M + \varepsilon$ and define its size as

$$|\Pi| := \max_{k=1,\dots,n} \lambda_k - \lambda_{k-1}.$$

We then select points $\mu_i \in [\lambda_{i-1}, \lambda_i]$, $1 \le i \le n$, and we set

$$S_\Pi := \sum_{i=1}^n f(\mu_i)(E_{\lambda_i} - E_{\lambda_{i-1}}) \in \mathcal{B}(\mathcal{H}).$$

Lemma 2.2. There exists a unique $S \in \mathcal{B}(\mathcal{H})$ so that

$$\forall \eta > 0,\ \exists \delta_\eta > 0,\ |\Pi| < \delta_\eta \Longrightarrow \|S_\Pi - S\| < \eta.$$

Furthermore, S is independent of the choice of ε, the extension of f and the choice of points $(\mu_i)_{i=1}^n$.

This idea here really comes from Riemann sums. The idea of the proof is that if $(\Pi_n)_{n\in\mathbb{N}}$ is a sequence of partitions which size tends to 0, then $(S_{\Pi_n})_{n\in\mathbb{N}}$ is Cauchy in $\mathcal{B}(\mathcal{H})$, and therefore converges.

Proof. Let $\eta > 0$. Since f is continuous on the compact set $[m, M+\varepsilon]$, f is uniformly continuous, and there is $\delta_\eta > 0$ so that

$$\mu, \mu' \in [m, M+\varepsilon],\ |\mu - \mu'| < \delta_\eta \Longrightarrow |f(\mu) - f(\mu')| < \frac{\eta}{2}.$$

The first part consists at proving that if two partitions Π and Π' have $|\Pi|, |\Pi'| < \delta_\eta$, then $\|S_\Pi - S_{\Pi'}\| < \eta$. Let us write then $\Pi = (\lambda_i)_{i=0}^n$, and consider

$$\overline{\Pi} := \Pi \cup \Pi' = (\overline{\lambda_j})_{j=0}^{\overline{n}}$$

where $\overline{n} > n$. Choose $\mu_i \in [\lambda_{i-1}, \lambda_i]$, $1 \le i \le n$, and $\overline{\mu_j} \in [\overline{\lambda_{j-1}}, \overline{\lambda_j}]$, $1 \le j \le \overline{n}$. Now, each λ_i is some $\overline{\lambda_{k_i}}$. This gives a subset of $\{0, \dots, \overline{n}\}$, namely $0 = k_0 < k_1 < \cdots < k_n = \overline{n}$. We will now estimate $\|S_\Pi - S_{\overline{\Pi}}\|$. The trick to do this is to use those indices $k_0 < k_1 < \cdots < k_n$ and write

$$f(\mu_i)(E_{\lambda_i} - E_{\lambda_{i-1}}) = f(\mu_i) \sum_{j=k_{i-1}+1}^{k_i} (E_{\overline{\lambda_j}} - E_{\overline{\lambda_{j-1}}})$$

in order to get

$$S_\Pi = \sum_{i=1}^{n} \sum_{j=k_{i-1}+1}^{k_i} f(\mu_i)(E_{\overline{\lambda_j}} - E_{\overline{\lambda_{j-1}}}),\ S_{\overline{\Pi}} = \sum_{i=1}^{n} \sum_{j=k_{i-1}+1}^{k_i} f(\overline{\mu_j})(E_{\overline{\lambda_j}} - E_{\overline{\lambda_{j-1}}}).$$

If $u \in \mathcal{H}$ has $\|u\| = 1$, we compute that

$$\begin{aligned}
|\langle (S_\Pi - S_{\overline{\Pi}})u, u\rangle| &= \left|\left\langle \sum_i \sum_j (f(\mu_i) - f(\overline{\mu_j}))(E_{\overline{\lambda_j}} - E_{\overline{\lambda_{j-1}}})u, u \right\rangle\right| \\
&= \left| \sum_{i,j} (f(\mu_i) - f(\overline{\mu_j}))\langle (E_{\overline{\lambda_j}} - E_{\overline{\lambda_{j-1}}})u, u\rangle \right| \\
&< \frac{\eta}{2} \sum_{i,j} \langle (E_{\overline{\lambda_j}} - E_{\overline{\lambda_{j-1}}})u, u\rangle \\
&= \frac{\eta}{2} \left\langle \sum_{i,j} (E_{\overline{\lambda_j}} - E_{\overline{\lambda_{j-1}}})u, u \right\rangle \\
&= \frac{\eta}{2} \langle (E_{M+\varepsilon} - E_m)u, u\rangle \\
&= \frac{\eta}{2} \|u\|^2
\end{aligned}$$

$$= \frac{\eta}{2}.$$

and the upper bound follows from the uniform continuity of f, which we may use since

$$|\mu_i - \overline{\mu_j}| \leq \lambda_i - \lambda_{i-1} \leq |\Pi| < \delta_\eta.$$

We then obtain, using Theorem 1.44, that $\|S_\Pi - S_{\overline{\Pi}}\| < \frac{\eta}{2}$. Likewise, since $|\Pi'| < \delta_\eta$, we get $\|S_{\Pi'} - S_{\overline{\Pi}}\| < \frac{\eta}{2}$. The triangle inequality now yields to

$$\|S_\Pi - S_{\Pi'}\| \leq \|S_\Pi - S_{\overline{\Pi}}\| + \|S_{\overline{\Pi}} - S_{\Pi'}\| < \frac{\eta}{2} + \frac{\eta}{2} = \eta$$

as announced above. This concludes the first step of the proof.

Now take a sequence $(\Pi_n)_{n\in\mathbb{N}}$ with $|\Pi_n| \longrightarrow 0$ as $n \to \infty$. Hence there is $N \in \mathbb{N}$ so that $n \geq N \Longrightarrow |\Pi_n| < \delta_\eta$, and thus

$$n, m \geq N \Longrightarrow |\Pi_n|, |\Pi_m| < \delta_\eta \Longrightarrow \|S_{\Pi_n} - S_{\Pi_m}\| < \eta.$$

This means that $(S_{\Pi_n})_{n\in\mathbb{N}}$ is Cauchy in $\mathcal{B}(\mathcal{H})$, and therefore there is $S \in \mathcal{B}(\mathcal{H})$ so that $\|S_{\Pi_n} - S\| \longrightarrow 0$ as $n \to \infty$. The latter convergence implies that there is $N_\eta \in \mathbb{N}$ so that $|\Pi_{N_\eta}| < \delta_{\eta/2}$ and $\|S_{\Pi_{N_\eta}} - S\| < \frac{\eta}{2}$. Thus if Π is a partition with $|\Pi| < \delta_{\eta/2}$, we get

$$\|S_\Pi - S\| \leq \|S_\Pi - S_{\Pi_{N_\eta}}\| + \|S_{\Pi_{N_\eta}} - S\| < \frac{\eta}{2} + \frac{\eta}{2} = \eta.$$

Finally it is clear that S does not depend on any choice we made above. We are done. □

Definition 2.3. The operator S obtained in Lemma 2.2 is called the integral of S with respect to $(E_\lambda)_{\lambda\in\mathbb{R}}$, and is denoted

$$S = \int_m^{M+\varepsilon} f(\lambda)\, \mathrm{d}E_\lambda.$$

Observe for instance that $\displaystyle\int_m^{M+\varepsilon} \mathrm{d}E_\lambda = E_{M+\varepsilon} - E_m = \mathrm{Id}_{\mathcal{H}}.$

2.2 Spectral theorem I

Here is the statement and the proof of the spectral theorem.

Theorem 2.4. Let S be a bounded symmetric operator.
There exists a unique spectral family $(E_\lambda)_{\lambda\in\mathbb{R}}$ such that the following hold.

(i) If $A \in \mathcal{B}(\mathcal{H})$ and $AS = SA$, then $AE_\lambda = E_\lambda A$ for all $\lambda \in \mathbb{R}$.

(ii) For all $u \in \mathcal{H}$, for all $\mu \in \mathbb{R}$, the limit $\lim_{\lambda \to \mu^+} E_\lambda u$ exists in $\mathcal{H}$.

(iii) The upper and lower bounds of S are the upper and lower bounds for $(E_\lambda)_{\lambda \in \mathbb{R}}$.

(iv) S has the representation

$$S = \int_m^{M+\varepsilon} \lambda \, \mathrm{d}E_\lambda.$$

$(E_\lambda)_{\lambda \in \mathbb{R}}$ is then called the spectral family of $S \in \mathcal{B}(\mathcal{H})$.

Proof. Let $E_+(\lambda)$ be the projection onto $\operatorname{Ker}((S - \lambda \mathrm{Id}_{\mathcal{H}}) - |S - \lambda \mathrm{Id}_{\mathcal{H}}|)$, and let

$$E_\lambda := \mathrm{Id}_{\mathcal{H}} - E_+(\lambda).$$

Clearly, E_λ is a projection for all $\lambda \in \mathbb{R}$. From the definition of the family $(E_\lambda)_{\lambda \in \mathbb{R}}$, it appears that Lemma 1.71 will play a key role in the proof. For instance Lemma 1.71(i) already implies point (i) of the theorem, and also the relation $E_\lambda E_\mu = E_\mu E_\lambda$ for every $\lambda, \mu \in \mathbb{R}$.

Let $\lambda < \mu$, and define $P = E_\lambda(\mathrm{Id}_{\mathcal{H}} - E_\mu)$. First, note the relations $E_\lambda P = P$ and $(\mathrm{Id}_{\mathcal{H}} - E_\mu)P = P$ (as projections are idempotent). Fix $u \in \mathcal{H}$ and $v := Pu$. Then

$$E_\lambda v = E_\lambda P u = Pu = v, \ (\mathrm{Id}_{\mathcal{H}} - E_\mu)v = (\mathrm{Id}_{\mathcal{H}} - E_\mu)Pu = Pu = v$$

and thus

$$\begin{aligned}
\langle (S - \lambda \mathrm{Id}_{\mathcal{H}})v, v\rangle &= \langle (S - \lambda \mathrm{Id}_{\mathcal{H}})E_\lambda v, v\rangle = \langle (S - \lambda \mathrm{Id}_{\mathcal{H}})(\mathrm{Id}_{\mathcal{H}} - E_+(\lambda))v, v\rangle \le 0 \\
\langle (S - \mu \mathrm{Id}_{\mathcal{H}})v, v\rangle &= \langle (S - \mu \mathrm{Id}_{\mathcal{H}})(\mathrm{Id}_{\mathcal{H}} - E_\mu)v, v\rangle = \langle (S - \mu \mathrm{Id}_{\mathcal{H}})E_+(\mu)v, v\rangle \ge 0
\end{aligned}$$

where we used that $(S - \mu \mathrm{Id}_{\mathcal{H}})E_+(\mu) \ge 0$ and $(S - \lambda \mathrm{Id}_{\mathcal{H}})(\mathrm{Id}_{\mathcal{H}} - E_+(\lambda)) \le 0$ by Lemma 1.70(iii). Substracting these two lines provide $(\mu - \lambda)\|v\|^2 \le 0$, which implies $\|v\|^2 = 0$. Hence $\|Pu\|^2 = 0$, so $Pu = 0$, and $P = 0$. It follows that $E_\lambda = E_\lambda E_\mu$ (developing the expression for P) and thus $E_\lambda \le E_\mu$ by Theorem 1.54(iv). This shows that $(E_\lambda)_{\lambda \in \mathbb{R}}$ is increasing.

Now we turn to proving (iii) of the theorem. We treat the case of m the lower bound of S, and the case of M is similar. By contradiction, suppose that $\lambda < m$ and $E_\lambda \neq 0$. Take $u \in \mathcal{H}$ so that $v := E_\lambda u \neq 0$. Without restriction, we may assume that $\|v\| = 1$. Again using Lemma 1.71(iii), we have

$$\begin{aligned}
\langle Sv, v\rangle - \lambda &= \langle (S - \lambda \mathrm{Id}_{\mathcal{H}})v, v\rangle \\
&= \langle (S - \lambda \mathrm{Id}_{\mathcal{H}})E_\lambda u, E_\lambda u\rangle \\
&= \langle (S - \lambda \mathrm{Id}_{\mathcal{H}})E_\lambda u, u\rangle \\
&\le 0
\end{aligned}$$

as E_λ is symmetric, idempotent and commutes with $(S - \lambda \mathrm{Id}_{\mathcal{H}})$, and the last step is a consequence of $(S - \lambda \mathrm{Id}_{\mathcal{H}})E_\lambda \leq 0$. We deduce then that $m \leq \langle Sv, v\rangle \leq \lambda$, which contradicts the way we chose λ. Thus $E_\lambda = 0$ for every $\lambda < m$, as claimed. Point (iii) of the theorem is proven.

Let now $u \in \mathcal{H}$, and $\lambda, \mu \in \mathbb{R}$. We show that the two limits $\lim\limits_{\lambda\to\mu^\pm} E_\lambda u$ exists. We handle the case $\lambda \to \mu^+$, and the other one is dual. As the spectral family $(E_\lambda)_{\lambda\in\mathbb{R}}$ is increasing, the map $\lambda \longmapsto \langle E_\lambda u, u\rangle$ is non-decreasing, so the limit

$$\ell_{\mu^+} := \lim_{\lambda\to\mu^+} \langle E_\lambda u, u\rangle$$

exists. Let $\eta > 0$. Then there is $\delta > 0$ so that $\mu < \lambda < \mu + \delta \Longrightarrow |\langle E_\lambda u, u\rangle - \ell_{\mu^+}| < \frac{\eta}{2}$. Hence, if $\mu < \nu < \lambda < \mu + \delta$, one has

$$\begin{aligned}
\|E_\lambda u - E_\nu u\|^2 &= \langle (E_\lambda - E_\nu)^2 u, u\rangle \\
&= \langle (E_\lambda - E_\nu) u, u\rangle \\
&= \langle E_\lambda u, u\rangle - \ell_{\mu^+} + \ell_{\mu^+} - \langle E_\nu u, u\rangle \\
&\leq |\langle E_\lambda u, u\rangle - \ell_{\mu^+}| + |\ell_{\mu^+} - \langle E_\nu u, u\rangle| \\
&< \frac{\eta}{2} + \frac{\eta}{2} = \eta.
\end{aligned}$$

Using completeness of $\mathcal{H}$ (in the same way as in the proof of Lemma 2.2), we see that $\lim\limits_{\lambda\to\mu^+} E_\lambda u$ exists, and we proved point (ii) of the theorem. Likewise, we obtain the existence of $\lim\limits_{\lambda\to\mu^-} E_\lambda u$ for any $u \in \mathcal{H}$, and

We must still argue that this last limit coincides with $E_\mu u$. For $\lambda < \mu$, set $E_\Delta := E_\mu - E_\lambda \geq 0$. First, we have

$$\begin{aligned}
E_\mu E_\Delta &= E_\mu^2 - E_\mu E_\lambda = E_\mu - E_\lambda = E_\Delta \\
(\mathrm{Id}_{\mathcal{H}} - E_\lambda)E_\Delta &= E_\Delta - E_\lambda E_\mu + E_\lambda^2 = E_\Delta
\end{aligned}$$

so we may write $(S-\mu\mathrm{Id}_{\mathcal{H}})E_\Delta = (S-\mu\mathrm{Id}_{\mathcal{H}})E_\mu E_\Delta$ and $(S-\lambda\mathrm{Id}_{\mathcal{H}})E_\Delta = (S-\lambda\mathrm{Id}_{\mathcal{H}})(\mathrm{Id}_{\mathcal{H}}-E_\lambda)E_\Delta$. As $E_\Delta \geq 0$ and as

$$(S - \mu\mathrm{Id}_{\mathcal{H}})E_\mu \leq 0,\ (S - \lambda\mathrm{Id}_{\mathcal{H}})(\mathrm{Id}_{\mathcal{H}} - E_\lambda) \geq 0$$

we conclude from Corollary 1.63 that $(S - \mu\mathrm{Id}_{\mathcal{H}})E_\Delta \leq 0$ and $(S - \lambda\mathrm{Id}_{\mathcal{H}})E_\Delta \geq 0$. Developing these inequalities, we arrive at

$$\lambda E_\Delta \leq SE_\Delta \leq \mu E_\Delta. \tag{3}$$

Now consider $E_{\Delta_0} := E_\mu - E_{\mu^-}$, and note that $\lim\limits_{\lambda\to\mu^-} E_\lambda u = E_{\Delta_0} u$ for all $u \in \mathcal{H}$. Taking strong limits in (3), we have

$$\mu E_{\Delta_0} \leq SE_{\Delta_0} \leq \mu E_{\Delta_0}$$

whence $\mu E_{\Delta_0} = SE_{\Delta_0}$. To finish, let $u \in \mathcal{H}$ and $v := E_{\Delta_0}u$. Then $(S - \mu \mathrm{Id}_{\mathcal{H}})v = (S - \mu \mathrm{Id}_{\mathcal{H}})E_{\Delta_0}u = 0$, so $Sv = \mu v$. Lemma 1.71(ii) therefore gives $E_+(\mu)v = v$, and thus

$$v = E_{\Delta_0}u = \lim_{\lambda \to \mu^-} E_\Delta u = \lim_{\lambda \to \mu^-} E_\mu E_\Delta u = E_\mu E_{\Delta_0} u = (\mathrm{Id}_{\mathcal{H}} - E_+(\mu))v = 0.$$

We have then $E_{\Delta_0} = 0$, thus $E_\mu u = \lim_{\lambda \to \mu^-} E_\lambda u$ for all $u \in \mathcal{H}$, and $(E_\lambda)_{\lambda \in \mathbb{R}}$ is strongly left-continuous.

We are left to prove point (iv) in the statement. Relations (3) are gonna be useful here. Consider a sequence of partitions $(\Pi_j)_{j=1}^{\infty}$, with $|\Pi_j| \longrightarrow 0$ as $j \to \infty$. We denote

$$m = \lambda_0^j < \lambda_1^j < \cdots < \lambda_{n_j}^j = M + \varepsilon.$$

Let $E_{\Delta_k}^j := E_{\lambda_k^j} - E_{\lambda_{k-1}^j}$. By (3), one has $\lambda_{k-1}^j E_{\Delta_k^j} \leq SE_{\Delta_k^j} \leq \lambda_k^j E_{\Delta_k^j}$ and summing gives then

$$\sum_{k=1}^{n_j} \lambda_{k-1}^j E_{\Delta_k^j} \leq S \sum_{k=1}^{n_j} E_{\Delta_k^j} \leq \sum_{k=1}^{n_j} \lambda_k^j E_{\Delta_k^j}.$$

The first sum is exactly S_{Π_j} with chosen points $\mu_k^j = \lambda_{k-1}^j$, while the last one is the same but with chosen points $\mu_k^j = \lambda_k^j$. The middle sum is telescopic and we are left with $S\mathrm{Id}_{\mathcal{H}} = S$. Letting then $j \to \infty$, we get

$$\int_m^{M+\varepsilon} \lambda \, \mathrm{d}E_\lambda \leq S \leq \int_m^{M+\varepsilon} \lambda \, \mathrm{d}E_\lambda$$

which finishes our proof. □

To prove the uniqueness part of the spectral theorem, we will appeal the next lemma and exercise.

Lemma 2.5. Let $S \in \mathcal{B}(\mathcal{H})$ be bounded and symmetric, and let $(E_\lambda)_{\lambda \in \mathbb{R}}$ be a spectral family, as in Spectral Theorem I. For any $P \in \mathbb{R}[X]$, it holds that

$$P(S) = \int_m^{M+\varepsilon} P(\lambda) \, \mathrm{d}E_\lambda.$$

Proof. It is enough to prove the theorem for monomials $P(\lambda) = \lambda^l$, $l \in \mathbb{N}$. For $l = 0$ there is nothing to prove, and the case $l = 1$ is exactly Spectral theorem I. We treat the general case by induction, assuming it holds for some $l \geq 0$. Fix $0 < \eta < 1$. By the theorem and induction hypothesis, there exists $\delta > 0$ so that for any partition $\Pi = (\lambda_k)_{k=0}^n$ with $|\Pi| < \delta$, we have

$$\left\| S - \sum_{k=1}^n \lambda_k E_{\Delta_k} \right\| \leq \eta \quad \text{and} \quad \left\| S^l - \sum_{k=1}^n \lambda_k^l E_{\Delta_k} \right\| \leq \eta$$

where $E_{\Delta_k} = E_{\lambda_k} - E_{\lambda_{k-1}}$. Let us abbreviate $T = \sum_{k=1}^n \lambda_k E_{\Delta_k}$ and $T^{(l)} = \sum_{k=1}^n \lambda_k^l E_{\Delta_k}$, and observe that with this notation, we have

$$T^{(l)}T = \left(\sum_{k=1}^n \lambda_k^l E_{\Delta_k}\right)\left(\sum_{k=1}^n \lambda_k E_{\Delta_k}\right) = \sum_{k=1}^n \lambda^{l+1} E_{\Delta_k}$$

since

$$E_{\Delta_k}^2 = (E_{\lambda_k} - E_{\lambda_{k-1}})^2 = E_{\lambda_k}^2 - 2E_{\lambda_k}E_{\lambda_{k-1}} + E_{\lambda_{k-1}}^2 = E_{\Delta_k}$$

using Theorem 1.54, and also $E_{\Delta_i}E_{\Delta_j} = 0$ for $i \neq j$. Now the triangle inequality provides

$$\begin{aligned}
\|S^{l+1} - T^{(l)}T\| &= \|(S^{l+1} + T^{(l)}T - S^lT - ST^{(l)}) + (S^lT - S^{l+1}) + (ST^{(l)} - S^{l+1})\| \\
&\le \|S^{l+1} + T^{(l)}T - S^lT - ST^{(l)}\| + \|S^lT - S^{l+1}\| + \|ST^{(l)} - S^{l+1}\| \\
&= \|(S^l - T^{(l)})(S-T)\| + \|S^lT - S^{l+1}\| + \|ST^{(l)} - S^{l+1}\| \\
&\le \|S^l - T^{(l)}\| + \|S^l\|\|T - S\| + \|S\|\|T^{(l)} - S^l\| \\
&\le \eta^2 + \|S\|^l\eta + \|S\|\eta.
\end{aligned}$$

It thus follows from the uniqueness part of Lemma 2.2 that

$$S^{l+1} = \int_m^{M+\varepsilon} \lambda^{l+1} \,\mathrm{d}E_\lambda$$

as announced. This concludes the inductive step and the proof. □

Exercise 2.6. Let $S \in \mathcal{B}(\mathcal{H})$ be symmetric, and $(E_\lambda)_{\lambda\in\mathbb{R}}$ be a corresponding spectral family. Prove that for any $P \in \mathbb{R}[X]$ and $u, v \in \mathcal{H}$, we have

$$\langle P(S)u, v\rangle = \int_m^{M+\varepsilon} P(\lambda) \,\mathrm{d}\langle E_\lambda u, v\rangle$$

where the right-hand side is the Riemann-Stieltjes integral of P with respect to $\phi(\lambda) := \langle E_\lambda u, v\rangle$.

The above exercise allows us to exploit properties of the Riemann-Stieltjes integral to establish uniqueness of the spectral family.

Corollary 2.7. Let $S \in \mathcal{B}(\mathcal{H})$ be symmetric.
The spectral family provided by Spectral Theorem I is unique.

Proof. Suppose $(E_\lambda)_{\lambda\in\mathbb{R}}$, $(F_\lambda)_{\lambda\in\mathbb{R}}$ are two spectral families for S, satisfying (i)-(iv) of Spectral Theorem I. Fix $u \in \mathcal{H}$, and consider

$$\phi(\lambda) := \langle (E_\lambda - F_\lambda)u, u\rangle, \ \lambda \in \mathbb{R}.$$

By Exercise 1.42, it is enough to show that ϕ is identically 0. Since $\phi(m) = 0$, it is enough to prove that ϕ is constant. By a property of the Riemann-Stieltjes integral, it is enough to prove that

$$\int_m^{M+\varepsilon} f(\lambda)\,\mathrm{d}\phi(\lambda) = 0$$

for any continuous function $f\colon [m, M+\varepsilon] \longrightarrow \mathbb{R}$. Let then f be such a function. By the Stone-Weierstrass theorem, there exists a sequence $(P_n)_{n\in\mathbb{N}}$ of polynomials so that $\|P_n - f\|_\infty \longrightarrow 0$ as $n \to \infty$. By another property of the Riemann-Stieltjes integral, it follows that

$$\begin{aligned}
\int_m^{M+\varepsilon} f(\lambda)\,\mathrm{d}\phi(\lambda) &= \lim_{n\to\infty}\int_m^{M+\varepsilon} P_n(\lambda)\,\mathrm{d}\phi(\lambda)\\
&= \lim_{n\to\infty}\int_m^{M+\varepsilon} P_n(\lambda)\,\mathrm{d}\langle E_\lambda u, v\rangle - \lim_{n\to\infty}\int_m^{M+\varepsilon} P_n(\lambda)\,\mathrm{d}\langle F_\lambda u, v\rangle\\
&= \lim_{n\to\infty} P_n(S) - \lim_{n\to\infty} P_n(S)\\
&= 0
\end{aligned}$$

and the third equality comes from Exercise 2.6, that we may apply since $(E_\lambda)_{\lambda\in\mathbb{R}}$ and $(F_\lambda)_{\lambda\in\mathbb{R}}$ are both spectral families for S. As explained above, we deduce that $E_\lambda = F_\lambda$ for all $\lambda \in \mathbb{R}$, concluding the proof. □

2.3 Caracterisations of the different parts of the spectrum

We conclude this chapter by discussing the relation between the spectral family of a symmetric operator S and the two components of its spectrum $\sigma_p(S), \sigma_c(S)$.

Theorem 2.8. Let $S \in \mathcal{B}(\mathcal{H})$ be a symmetric operator and $(E_\lambda)_{\lambda\in\mathbb{R}}$ be the corresponding spectral family.
Then $\lambda_0 \in \sigma_p(S)$ if and only if the map $E\colon \mathbb{R} \longrightarrow \mathcal{B}(\mathcal{H})$ is discontinuous at λ_0.
In this case, one has

$$\operatorname{Ker}(S - \lambda_0 \mathrm{Id}_{\mathcal{H}}) = \operatorname{Im}(E_{\lambda_0^+} - E_{\lambda_0}).$$

Proof. Clearly the first statement is a consequence of the equality

$$\operatorname{Ker}(S - \lambda_0 \mathrm{Id}_{\mathcal{H}}) = \operatorname{Im}(E_{\lambda_0^+} - E_{\lambda_0}).$$

Let $\lambda_0 \in [m, M]$, and first suppose that $u \in \operatorname{Ker}(S - \lambda_0\mathrm{Id}_{\mathcal{H}})$. Then $(S - \lambda_0\mathrm{Id}_{\mathcal{H}})^2 u = 0$ and, by Exercise 2.6, we get

$$\int_a^b (\lambda - \lambda_0)^2\,\mathrm{d}\langle E_\lambda u, u\rangle = \langle (S - \lambda_0\mathrm{Id}_{\mathcal{H}})^2 u, u\rangle = 0$$

for any $a < m$ and $b > M$. Hence, the integral of $\lambda \longmapsto (\lambda - \lambda_0)^2$ must be 0 on any subinterval of positive length. Let then $\varepsilon > 0$, so that

$$0 = \int_a^{\lambda_0-\varepsilon} (\lambda - \lambda_0)^2 \, \mathrm{d}\langle E_\lambda u, u\rangle \geq \varepsilon^2 \int_a^{\lambda_0-\varepsilon} \mathrm{d}\langle E_\lambda u, u\rangle = \varepsilon^2 \langle E_{\lambda_0-\varepsilon} u, u\rangle$$

and likewise

$$0 = \int_{\lambda_0+\varepsilon}^{b} (\lambda - \lambda_0)^2 \, \mathrm{d}\langle E_\lambda u, u\rangle \geq \varepsilon^2 \langle (\mathrm{Id}_{\mathcal{H}} - E_{\lambda_0+\varepsilon})u, u\rangle.$$

It follows that

$$\langle E_{\lambda_0-\varepsilon} u, u\rangle = 0 = \langle u, u\rangle - \langle E_{\lambda_0+\varepsilon} u, u\rangle$$

and thus $E_{\lambda_0-\varepsilon} u = 0 = u - E_{\lambda_0+\varepsilon} u$, using that $E_{\lambda_0-\varepsilon}, E_{\lambda_0+\varepsilon}$ are projections. We conclude that

$$u = (E_{\lambda_0+\varepsilon} - E_{\lambda_0-\varepsilon})u$$

for all $\varepsilon > 0$, and letting $\varepsilon \longrightarrow 0$ now provides $u = (E_{\lambda_0^+} - E_{\lambda_0})u$ since the map $\lambda \longmapsto E_\lambda$ is strongly left-continuous. Thus $u \in \mathrm{Im}(E_{\lambda_0^+} - E_{\lambda_0})$ as claimed.

Conversely, let $n \geq 1$, and use inequalities (3) from the proof of the spectral theorem to get

$$\lambda_0(E_{\lambda_0+\frac{1}{n}} - E_{\lambda_0}) \leq S(E_{\lambda_0+\frac{1}{n}} - E_{\lambda_0}) \leq (\lambda_0 + \frac{1}{n})(E_{\lambda_0+\frac{1}{n}} - E_{\lambda_0}).$$

Letting $n \longrightarrow \infty$ provides $\lambda_0(E_{\lambda_0^+} - E_{\lambda_0}) \leq S(E_{\lambda_0^+} - E_{\lambda_0}) \leq \lambda_0(E_{\lambda_0^+} - E_{\lambda_0})$, whence

$$S(E_{\lambda_0^+} - E_{\lambda_0}) = \lambda_0(E_{\lambda_0^+} - E_{\lambda_0}).$$

It follows that $(S - \lambda_0 \mathrm{Id}_{\mathcal{H}})(E_{\lambda_0^+} - E_{\lambda_0}) = 0$, which means that

$$\mathrm{Im}(E_{\lambda_0^+} - E_{\lambda_0}) \subset \mathrm{Ker}(S - \lambda_0 \mathrm{Id}_{\mathcal{H}}).$$

The proof is now complete. □

The next result provides a remarkable description of the resolvent set in terms of the spectral family.

Theorem 2.9. Let $S \in \mathcal{B}(\mathcal{H})$ be a symmetric operator and $(E_\lambda)_{\lambda\in\mathbb{R}}$ be the corresponding spectral family.
Then $\lambda_0 \in \rho(S)$ if and only if there exists $\varepsilon > 0$ so that the map $E\colon \mathbb{R} \longrightarrow \mathcal{B}(\mathcal{H})$ is constant on $[\lambda_0 - \varepsilon, \lambda_0 + \varepsilon]$.

The next corollary is then a direct consequence of the previous results.

Corollary 2.10. Let $S \in \mathcal{B}(\mathcal{H})$ be a symmetric operator and $(E_\lambda)_{\lambda\in\mathbb{R}}$ be the corresponding spectral family.
Then $\lambda_0 \in \sigma_c(S)$ if and only if $E\colon \mathbb{R} \longrightarrow \mathcal{B}(\mathcal{H})$ is continuous but not locally constant at λ_0.

3. The spectral theorem for self-adjoint operators

The goal of this third chapter is to extend the spectral theorem seen in chapter 2 to not necessarily bounded operators.

3.1 Unbounded linear operators

As from the beginning, $\mathcal{H}$ denotes a complex Hilbert space.

Definition 3.1. A linear operator is a linear map $T\colon \mathcal{D}_T \longrightarrow \mathcal{H}$, where $\mathcal{D}_T$ is a subspace of $\mathcal{H}$.

As usual, if T is a linear operator, we define its range and its kernel by

$$\mathrm{Im}(T) := \{Tu : u \in \mathcal{D}_T\}, \ \mathrm{Ker}(T) := \{u \in \mathcal{D}_T : Tu = 0\}.$$

For two operators $T\colon \mathcal{D}_T \longrightarrow \mathcal{H}$, $T'\colon \mathcal{D}_{T'} \longrightarrow \mathcal{H}$, we say that T' is an *extension of* T, and we denote $T \subset T'$, if $\mathcal{D}_T \subset \mathcal{D}_{T'}$ and $T'u = Tu$ for any $u \in \mathcal{D}_T$.

Exercise 3.2. Check that $\subset$ is a partial order on the set of operators on $\mathcal{H}$.

Our first proposition says that that we can always define a bounded operator on a closed subspace of $\mathcal{H}$.

Proposition 3.3. If $T\colon \mathcal{D}_T \longrightarrow \mathcal{H}$ is bounded on $\mathcal{D}_T$, then there exists a unique bounded extension of T to $\overline{\mathcal{D}_T}$.

Proof. Let $u \in \overline{\mathcal{D}_T}$ and pick a sequence $(u_n)_{n\in\mathbb{N}}$ converging to u. As T is bounded on $\mathcal{D}_T$, we have

$$\|Tu_n - Tu_m\| \leq \|T\|\|u_n - u_m\|$$

for all $n, m \in \mathbb{N}$, and as $(u_n)_{n\in\mathbb{N}}$ is Cauchy (because it is a convergent sequence), $(Tu_n)_{n\in\mathbb{N}}$ is also Cauchy. By completeness of $\mathcal{H}$, it converges, and we denote Tu its limit. This defines a linear operator on $\overline{\mathcal{D}_T}$, and letting $n \to \infty$ in the inequality $\|Tu_n\| \leq C\|u_n\|$ proves that this linear operator is still bounded on $\overline{\mathcal{D}_T}$. This finishes the proof. □

Definition 3.4. An operator $T\colon \mathcal{D}_T \longrightarrow \mathcal{H}$ is called densely defined if $\overline{\mathcal{D}_T} = \mathcal{H}$.

It follows from Proposition 3.3 that a densely defined bounded operator can be uniquely extended to a bounded operator on the whole space $\mathcal{H}$.

As in the bounded case, the *graph* of an operator $T\colon \mathcal{D}_T \longrightarrow \mathcal{H}$ is

$$G_T := \{(u, Tu) : u \in \mathcal{D}_T\} \subset \mathcal{H} \times \mathcal{H}$$

where $\mathcal{H} \times \mathcal{H}$ has the structure of a Hilbert space with the inner product

$$\langle (u,v), (u',v')\rangle_{\mathcal{H}\times\mathcal{H}} := \langle u, u'\rangle + \langle v, v'\rangle,\ u, u', v, v' \in \mathcal{H}.$$

Moreover, T is *closed* if its graph is closed in $\mathcal{H} \times \mathcal{H}$. Lastly, it is easy to check that

$$T \subset T' \Longleftrightarrow G_T \subset G_{T'}.$$

Definition 3.5. Let $T\colon \mathcal{D}_T \longrightarrow \mathcal{H}$.
The resolvent set of T, denoted $\rho(T)$, is the set of $\lambda \in \mathbb{C}$ so that $T - \lambda \mathrm{Id}_{\mathcal{H}}$ is a bijection from $\mathcal{D}_T$ to $\mathcal{H}$ with bounded inverse. Its spectrum is $\sigma(T) := \mathbb{C} \setminus \rho(T)$.

Again, as for the bounded case the spectrum of an operator splits into three parts, the point spectrum, the continuous spectrum and the residual spectrum.

One aspect which is different from the bounded case is the operations we may realize on operators. Let $T_1\colon \mathcal{D}_{T_1} \longrightarrow \mathcal{H}$, $T_2\colon \mathcal{D}_{T_2} \longrightarrow \mathcal{H}$ be two operators, and $\lambda \in \mathbb{C}$. Then

(i) λT_1 is defined as $(\lambda T_1)(u) := \lambda T_1 u$, for all $u \in \mathcal{D}_{T_1}$.

(ii) $T_1 + T_2$ is defined on $\mathcal{D}_{T_1}$ as $(T_1 + T_2)(u) := T_1 u + T_2 u$.

(iii) $T_1 T_2$ is defined on $D_{T_1 T_2} := \{u \in \mathcal{D}_{T_2} : T_2 u \in \mathcal{D}_{T_1}\}$ as $(T_1 T_2)(u) := T_1(T_2 u)$.

If $T\colon \mathcal{D}_T \longrightarrow \mathcal{H}$ is injective, it has an inverse $T^{-1}\colon \mathcal{D}_{T^{-1}} \longrightarrow \mathcal{H}$, and clearly $\mathcal{D}_{T^{-1}} = \mathrm{Im}(T)$, $\mathrm{Im}(T) := \mathcal{D}_T$. One has then $T^{-1}T \subset \mathrm{Id}_{\mathcal{H}}$ and $TT^{-1} \subset \mathrm{Id}_{\mathcal{H}}$ (check!).

3.2 The adjoint operator

If $T\colon \mathcal{D}_T \longrightarrow \mathcal{H}$ is not bounded, we cannot apply Riesz representation theorem as we did in Chapter 1 to define the adjoint of T everywhere. We must restrict ourselves to a subspace of $\mathcal{H}$.

More precisely, fix $T\colon \mathcal{D}_T \longrightarrow \mathcal{H}$ a densely defined operator, and consider

$$\mathcal{D}_{T^*} := \{v \in \mathcal{H} : u \longmapsto \langle Tu, v\rangle \text{ is bounded on } \mathcal{D}_T\}.$$

It is a subspace of $\mathcal{H}$. For $v \in \mathcal{D}_{T^*}$, $u \longmapsto \langle Tu, v\rangle$ can be uniquely extended to a bounded linear functional on $\mathcal{H}$. By Riesz representation theorem, there is a unique $v^* \in \mathcal{H}$ so that $\langle Tu, v\rangle = \langle u, v^*\rangle$ for all $u \in \mathcal{D}_T$. We let $T^*v := v^*$, and thus

$$\langle Tu, v\rangle = \langle u, T^*v\rangle, u \in \mathcal{D}_T, v \in \mathcal{D}_{T^*}.$$

Now we prove that T^* is linear. Let $v_1, v_2 \in \mathcal{D}_{T^*}$ and $\lambda_1, \lambda_2 \in \mathbb{C}$. We have

$$\begin{aligned}\langle Tu, \lambda_1 v_1 + \lambda_2 v_2\rangle &= \overline{\lambda_1}\langle Tu, v_1\rangle + \overline{\lambda_2}\langle Tu, v_2\rangle \\ &= \overline{\lambda_1}\langle u, T^* v_1\rangle + \overline{\lambda_2}\langle u, T^* v_2\rangle \\ &= \langle u, \lambda_1 T^* v_1 + \lambda_2 T^* v_2\rangle\end{aligned}$$

for any $u \in \mathcal{D}_T$. As $T^*(\lambda_1 v_1 + \lambda_2 v_2)$ is the unique element of $\mathcal{H}$, so that

$$\langle Tu, \lambda_1 v_1 + \lambda_2 v_2\rangle = \langle u, T^*(\lambda_1 v_1 + \lambda_2 v_2)\rangle$$

for all $u \in \mathcal{D}_T$, this forces $T^*(\lambda_1 v_1 + \lambda_2 v_2) = \lambda_1 T^* v_1 + \lambda_2 T^* v_2$. Hence T^* is linear.

Remark 3.6. Note that if $\mathcal{D} \subset \mathcal{H}$ is dense, then $w = 0$ if and only if $\langle w, u\rangle = 0$ for all $u \in \mathcal{D}$. Indeed, if $w = 0$ then clearly $\langle w, u\rangle = 0$ for any $u \in \mathcal{D}$, and conversely, if this condition holds, take a sequence $(w_n) \subset \mathcal{D}$ converging to w to compute

$$\|w\|^2 = \langle w, w\rangle = \lim_{n\to\infty} \langle w, w_n\rangle = 0$$

whence $w = 0$.

Lemma 3.7. If T is densely defined, then $\mathrm{Ker}(T^*) = \mathrm{Im}(T)^\perp$.

Proof. We have the equivalences

$$\begin{aligned}v \in \mathrm{Ker}(T^*) &\iff T^* v = 0 \\ &\iff \forall u \in \mathcal{D}_T,\ \langle u, T^* v\rangle = 0 \\ &\iff \forall u \in \mathcal{D}_T,\ \langle Tu, v\rangle = 0 \\ &\iff v \in \mathrm{Im}(T)^\perp\end{aligned}$$

using Remark 3.6 for the second equivalence, that we may apply since $\mathcal{D}_T$ is dense in $\mathcal{H}$. This proves the claim. □

The following proposition gives a fundamental property of the adjoint.

Proposition 3.8. If T is densely defined, then T^* is closed.

Proof. Let $(u_n, T^* u_n) \subset G_{T^*}$, and assume that this sequence converges to $(u, v) \in \mathcal{H} \times \mathcal{H}$. This implies that u_n converges to u in $\mathcal{H}$ and $T^* u_n$ converges to v in $\mathcal{H}$. Now we have

$$\langle Tw, u\rangle = \lim_{n\to\infty} \langle Tw, u_n\rangle = \lim_{n\to\infty} \langle w, T^* u_n\rangle = \langle w, v\rangle$$

for all $w \in D_T$, so the linear functional $w \longmapsto \langle Tw, u\rangle$ is continuous on D_T. This implies that $u \in \mathcal{D}_{T^*}$, and since

$$\langle w, T^*u\rangle = \langle Tw, u\rangle = \langle w, v\rangle$$

it follows that $\langle w, T^*u - v\rangle = 0$ for all $w \in \mathcal{D}_T$. As $\mathcal{D}_T$ is dense in $\mathcal{H}$, we have $T^*u = v$ by Remark 3.6. Hence G_{T^*} is closed in $\mathcal{H} \times \mathcal{H}$, which means that T^* is closed. □

Here is two exercises to manipulate these definitions.

Exercise 3.9. Let $\phi \in L^\infty(\mathbb{R})$ and consider the multiplication operator $T\colon L^2(\mathbb{R}) \longrightarrow L^2(\mathbb{R})$ defined as $(Tu)(x) := \phi(x)u(x)$, $x \in \mathbb{R}$. Show that T is bounded and compute its norm. Find T^*, and determine under which condition T is symmetric.

Now, suppose that $\lim\limits_{x\longrightarrow+\infty} |\phi(x)| = +\infty$. Show that T is unbounded, and find its domain. Find T^*.

Exercise 3.10. Show that if T^{-1}, T^* and $(T^{-1})^*$ exist, then $(T^*)^{-1}$ also exists and $(T^*)^{-1} = (T^{-1})^*$.

3.3 Operator graphs

We define two operators $U, V\colon \mathcal{H} \oplus \mathcal{H} \longrightarrow \mathcal{H} \oplus \mathcal{H}$ by setting

$$U(u,v) := (v,u),\ V(u,v) := (v,-u)$$

for any $(u,v) \in \mathcal{H} \oplus \mathcal{H}$. These operators have nice properties.

Exercise 3.11. Show that U, V are both bounded, unitary and satisfy $U^2 = \mathrm{Id}_{\mathcal{H}\oplus\mathcal{H}} = -V^2$. Show that U preserves the inner product on $\mathcal{H} \oplus \mathcal{H}$, and that for any subspace $X \subset \mathcal{H} \oplus \mathcal{H}$, we have $V(X^\perp) = V(X)^\perp$.

The next lemma is fundamental.

Lemma 3.12. Let $T\colon D_T \longrightarrow \mathcal{H}$ be densely defined. Then $G_{T^*} = (V(\overline{G_T}))^\perp$, or equivalently $(G_{T^*})^\perp = V(\overline{G_T})$.

Proof. First, note that T^* is well-defined as T is densely defined. For all $u \in D_T$, $v \in D_{T^*}$, we have $\langle Tu, v\rangle = \langle u, T^*v\rangle$, which can be written as

$$\begin{aligned}\langle Tu, v\rangle - \langle u, T^*v\rangle = 0 &\iff \langle (v, T^*v), (Tu, -u)\rangle_{\mathcal{H}\oplus\mathcal{H}} = 0\\ &\iff \langle (v, T^*v), V(u, Tu)\rangle = 0.\end{aligned}$$

This already shows that G_{T^*} is orthogonal to $V(G_T)$. Now we show both inclusions. Let $(w, T^*w) \in G_{T^*}$. Pick $(u,v) \in \overline{G_T}$ and a sequence $(u_n, v_n) \subset G_T$ so that $(u_n, v_n) \longrightarrow (u,v)$. Then we have

$$\langle (w, T^*w), V(u,v)\rangle = \lim_{n\longrightarrow\infty} \langle (w, T^*w), V(u_n, v_n)\rangle = 0$$

by the continuity of the inner product and the fact that $V(G_T)$ is orthogonal to the graph of T^*. This proves that $G_{T^*} \subset (V(\overline{G_T}))^\perp$.

Conversely, fix $(u,v) \in (V(\overline{G_T}))^\perp$. For all $w \in D_T$, we have

$$0 = \langle V(w,Tw),(u,v)\rangle = \langle (Tw,-w),(u,v)\rangle = \langle Tw,u\rangle - \langle w,v\rangle$$

so the linear functional $w \longmapsto \langle Tw,u\rangle$ equals the linear functional $w \longmapsto \langle w,v\rangle$, which is clearly bounded. We deduce $u \in D_{T^*}$, and that

$$\langle w,v\rangle = \langle Tw,u\rangle = \langle w,T^*u\rangle$$

for all $w \in D_T$. Hence $\langle w, v - T^*v\rangle = 0$ for all $w \in D_T$. As $D_T \subset \mathcal{H}$ is dense, Remark 3.6 ensures that $v - T^*u = 0$, whence $v = T^*u$. Thus $(u,v) = (u,T^*u) \in G_{T^*}$, and the inclusion

$$(V(\overline{G_T}))^\perp \subset G_{T^*}$$

is established, as well as the equality $G_{T^*} = (V(\overline{G_T}))^\perp$. Taking orthogonal of both sides and using Exercise 1.38, we see that

$$G_{T^*}^\perp = \overline{V(\overline{G_T})}.$$

Lastly, observing that V is a continuous linear bijection between two Banach spaces, V is open, and hence also closed (as it is a bijection). We conclude that $V(\overline{G_T})$ is closed in $\mathcal{H} \oplus \mathcal{H}$, whence $G_{T^*} = V(\overline{G_T})$. ◻

We use this result to prove that in the unbounded case, under suitable assumptions, the correspondance $T \longmapsto T^*$ is an involution, as in the bounded case.

Theorem 3.13. Let $T\colon D_T \longrightarrow \mathcal{H}$ be densely defined and closed. Then D_{T^*} is dense and $T^{**} := (T^*)^*$ equals T.

Proof. To prove $\overline{D_{T^*}} = \mathcal{H}$, we show equivalently that its orthogonal is $\{0\}$. Let then $h \in (\overline{D_{T^*}})^\perp$. As T is closed, G_T is closed, whence

$$\begin{aligned} G_T &= \overline{G_T} \\ &= V^2(\overline{G_T}) \\ &= V(V(\overline{G_T})) \\ &= V((G_{T^*})^\perp) \\ &= V(G_{T^*})^\perp \end{aligned}$$

using $V^2 = -\mathrm{Id}_{\mathcal{H}\oplus\mathcal{H}}$ for the second equality, Lemma 3.12 for the fourth equality, and Exercise 3.11 for the last one. Now, using once again Exercise 1.38 and the closedness of V, we get

$$(G_T)^\perp = (V(G_{T^*})^\perp)^\perp = \overline{V(G_{T^*})} = V(G_{T^*}).$$

Here we used that G_{T^*} is closed in $\mathcal{H} \oplus \mathcal{H}$, which is guaranteed by Proposition 3.8. Thus by the orthogonal decomposition theorem, which we may apply since G_T is closed, it follows that

$$\mathcal{H} \oplus \mathcal{H} = G_T \oplus (G_T)^\perp = G_T \oplus V(G_{T^*}).$$

Using this decomposition for $(0,h) \in \mathcal{H} \oplus \mathcal{H}$, we deduce there exists $u_h \in D_T$, $v_h \in D_{T^*}$ so that

$$(0,h) = (u_h, Tu_h) + (T^* v_h, -v_h)$$

whence $u_h + T^* v_h = 0$ and $h = Tu_h - v_h$. As $h \in (\overline{D_{T^*}})^\perp \subset (D_{T^*})^\perp$, we have $\langle h, v_h \rangle = 0$, which leads to

$$0 = \langle Tu_h - v_h, v_h \rangle = \langle Tu_h, v_h \rangle - \|v_h\|^2 = \langle u_h, T^* v_h \rangle - \|v_h\|^2 = -\langle T^* v_h, T^* v_h \rangle - \|v_h\|^2.$$

We conclude that $\|T^* v_h\|^2 = \|v_h\|^2 = 0$, so $v_h = T^* v_h = 0$, and thus $u_h = 0$ as well. Hence $h = 0$, as announced, and D_{T^*} is dense in $\mathcal{H}$. This ensures that T^{**} is well-defined, and using once again Lemma 3.12, one gets

$$G_{T^{**}} = (V(\overline{G_{T^*}}))^\perp = (V(G_{T^*}))^\perp = V(G_{T^*}^\perp) = V(V(\overline{G_T})) = \overline{G_T} = G_T$$

using the closedness of both G_T and G_{T^*} (as T and T^* are closed), the fact that V commutes with orthogonal, and that $V^2 = -\mathrm{Id}_{\mathcal{H} \oplus \mathcal{H}}$. As two operators are equal if and only their graphs are equal, we just proved that $T^{**} = T$. □

The following theorem will also be important throughout the rest of the course.

Theorem 3.14. Let T be closed and densely defined. Then

$$B := (\mathrm{Id}_{\mathcal{H}} + T^*T)^{-1}, \; C := T(\mathrm{Id}_{\mathcal{H}} + TT^*)^{-1}$$

are well-defined and bounded linear operators on $\mathcal{H}$, with $\|B\|, \|C\| \leq 1$. Furthermore, B is positive.

Proof. From Theorem 3.13 and its proof, we have $T^{**} = T$ and the splitting

$$\mathcal{H} \oplus \mathcal{H} = G_T \oplus V(G_{T^*}).$$

We use this splitting to define two operators B and C, and we will show they must coincide with the ones given in the statement. For any $h \in \mathcal{H}$, according to the above decomposition, we find $u_h \in D_T$ and $v_h \in D_{T^*}$ so that

$$(h, 0) = (u_h, Tu_h) + (T^* v_h, -v_h).$$

Define then $Bh := u_h, Ch := v_h$. Then both B and C are well-defined (because u_h, v_h are unique) and linear, and

$$\begin{cases} h = Bh + T^*Ch \\ 0 = TBh - Ch \end{cases}$$

for all $h \in \mathcal{H}$. This implies that $B + T^*C = \mathrm{Id}_{\mathcal{H}}$ and $C = TB$. Hence

$$\mathrm{Id}_{\mathcal{H}} = B + T^*C = B + T^*TB = (\mathrm{Id}_{\mathcal{H}} + T^*T)B.$$

This identity in particular provides $\mathrm{Im}(\mathrm{Id}_{\mathcal{H}} + T^*T) = \mathcal{H}$. Furthermore, for any $u \in D_{T^*T}$, it holds that

$$\langle (\mathrm{Id}_{\mathcal{H}} + T^*T)u, u\rangle = \langle u, u\rangle + \langle T^*Tu, u\rangle \geq \|u\|^2$$

and thus $\mathrm{Id}_{\mathcal{H}} + T^*T$ is injective. We conclude it is invertible, with inverse B. Additionally, since (u_h, Tu_h) and $(T^*v_h, -v_h)$ are orthogonal in $\mathcal{H} \oplus \mathcal{H}$, it follows that

$$\begin{aligned}\|h\|^2 &= \|(h,0)\|^2 \\ &= \|u_h\|^2 + \|Tu_h\|^2 + \|T^*v_h\|^2 + \|v_h\|^2 \\ &\geq \|Bh\|^2 + \|Ch\|^2\end{aligned}$$

so $\|B\|, \|C\| \leq 1$. To finish, we also have

$$\langle Bu, u\rangle = \langle Bu, (\mathrm{Id}_{\mathcal{H}} + T^*T)Bu\rangle = \|Bu\|^2 + \|TBu\|^2 \geq 0$$

for all $u \in \mathcal{H}$. Hence B is positive, and the proof is over. □

A useful corollary of this result will be stated in the next section.

3.4 Symmetric and self-adjoint operators

Definition 3.15. A densely defined operator $T\colon D_T \longrightarrow \mathcal{H}$ is called symmetric if $T \subset T^*$, *i.e.* if $D_T \subset D_{T^*}$ and $T^*u = Tu$ for all $u \in D_T$.

Here the condition we check in practice to prove a given operator is symmetric.

Lemma 3.16. $T\colon D_T \longrightarrow \mathcal{H}$ is symmetric if and only if $\overline{D_T} = \mathcal{H}$ and

$$\langle Tu, v\rangle = \langle u, Tv\rangle$$

for all $u, v \in D_T$.

Proof. Suppose first that T is symmetric. Then D_T is dense by definition. Moreover, for any $u, v \in D_T$, we have

$$\langle Tu, v\rangle = \langle u, T^*v\rangle = \langle u, Tv\rangle$$

because T^* equals T on D_T.

Conversely, suppose T is densely defined and that $\langle Tu, v\rangle = \langle u, Tv\rangle$ whenever $u, v \in D_T$. We must show the two conditions in Definition 3.15. First, fix $v \in D_T$. The linear functional $u \longmapsto \langle Tu, v\rangle$ equals the linear functional $u \longmapsto \langle u, Tv\rangle$ by assumption, and this functional is clearly bounded. Hence $v \in D_{T^*}$, and the first condition is fulfilled. For the second, we just notice that if $v \in D_T$, then

$$\langle u, Tv\rangle = \langle Tu, v\rangle = \langle u, T^*v\rangle$$

for all $u \in D_T$, whence $\langle u, Tv - T^*v\rangle = 0$ for all $u \in D_T$. As $D_T \subset \mathcal{H}$ is dense, Remark 3.6 gives that $Tv = T^*v$, and this holds for any $v \in D_T$. Thus T is symmetric, and this concludes the proof. ☐

Definition 3.17. $T\colon D_T \longrightarrow \mathcal{H}$ is closable if $\overline{G_T}$ is the graph of an operator. In this case, this operator is the closure of T and is denoted $\overline{T}$.

We immediately note that symmetric operators are closable.

Remark 3.18. Suppose that T is symmetric. Then $\mathcal{H} = \overline{D_T} \subset \overline{D_{T^*}}$, so D_{T^*} is dense in $\mathcal{H}$, and T^{**} is well-defined. Furthermore, by Lemma 3.12 we have

$$G_{T^{**}} = (V(G_{T^*}))^{\perp} = V(G_{T^*}^{\perp}) = V^2(\overline{G_T}) = \overline{G_T}$$

whence T is closable, and $\overline{T} = T^{**}$. This shows that T^{**} is an extension of T. Additionally, this extension is symmetric, as

$$T \subset T^* \Longrightarrow T^{**} \subset T^* = (T^*)^{**} = (T^{**})^*.$$

Here the first equality is Theorem 3.13. Hence, for T symmetric, T^{**} is a closed symmetric extension of T.

Exercise 3.19. Let $T\colon D_T \longrightarrow \mathcal{H}$ be densely defined.
(i) Show that T is closable if and only if T^* is densely defined, and that in this case $\overline{T} = T^{**}$.
(ii) Show that if T is densely defined and closable, then $(\overline{T})^* = T^*$.

We can now properly introduce self-adjoint operators.

Definition 3.20. $T\colon D_T \longrightarrow \mathcal{H}$ is self-adjoint if D_T is dense in $\mathcal{H}$ and if $T = T^*$.

Note that if T is self-adjoint, then it is closed and symmetric.

Remark 3.21. By Remark 3.18, a symmetric operator always has a closed symmetric extension. However, a symmetric operator (even closed) may have no self-adjoint extensions. Such operators are called *maximal symmetric*. Observe that any self-adjoint operator T is maximal symmetric. Indeed, if $T \subset S$ with S symmetric, then

$$S \subset S^* \subset T^* = T$$

whence $S = T$.

Additionally, a symmetric operator T is called *essentially self-adjoint* if its closure $\overline{T}$ (which exists by Exercise 3.19(i)) is self-adjoint.

Here is a nice characterization of essentially self-adjoint operators. However, we omit the proof here.

Theorem 3.22. A symmetric operator $T\colon D_T \longrightarrow \mathcal{H}$ is essentially self-adjoint if and only if $\operatorname{Im}(T + i\mathrm{Id}_{\mathcal{H}})$, $\operatorname{Im}(T - i\mathrm{Id}_{\mathcal{H}})$ are dense in $\mathcal{H}$.

We conclude this section with the following corollary of Theorem 3.14.

Corollary 3.23. Let A be self-adjoint. Then the operators $B = (I + A^2)^{-1}$ and $C = AB$ have the following properties:

(i) $B(\mathcal{D}_A) = \mathcal{D}_{A^3}$.

(ii) $BA \subset AB$.

(iii) $BC = CB$.

(iv) Any $T \in \mathcal{B}(\mathcal{H})$ with $AT \subset TA$ satisfies $TB = BT$.

Proof. (i) Let $u \in \mathcal{D}_A$. As $(\mathrm{Id}_{\mathcal{H}} + A^2)B = \mathrm{Id}_{\mathcal{H}}$ and $\mathcal{D}_{A-C} = \mathcal{D}_A$, we have

$$(A - C)u = Au - ABu = A(\mathrm{Id}_{\mathcal{H}} - B)u = A^3 Bu$$

and thus $Bu \in \mathcal{D}_{A^3}$. Conversely, if $v \in \mathcal{D}_{A^3}$, then $v \in \mathcal{D}_{A^2} = \mathcal{D}_{B^{-1}}$ and

$$u = B^{-1}v = (\mathrm{Id}_{\mathcal{H}} + A^2)v \in \mathcal{D}_A.$$

Thus $v \in B(\mathcal{D}_A)$, and (i) is settled.

(ii) Let $u \in \mathcal{D}_A$. By (i), $Bu \in \mathcal{D}_{A^3}$ and

$$ABu = B(\mathrm{Id}_{\mathcal{H}} + A^2)ABu = BA(\mathrm{Id}_{\mathcal{H}} + A^2)Bu = BAu$$

whence $BA \subset AB$.

(iii) This follows directly from (ii) since

$$BC = B(AB) \subset (AB)B = CB$$

and as B, C are defined everywhere, we conclude $BC = CB$.

(iv) Lastly, let $T \in \mathcal{B}(\mathcal{H})$ with $AT \subset TA$. Then also

$$A^2T = A(AT) \subset A(TA) = (AT)A \subset (TA)A = TA^2$$

and thus also $B^{-1}T \subset TB^{-1}$. Let now $u \in \mathcal{D}_A$ be arbitrary. Then

$$TBu = BB^{-1}TBu = BTB^{-1}Bu = BTu.$$

Since TB and BT are bounded (*i.e.* continuous) and agree on the dense subset $\mathcal{D}_A$, we conclude that in fact $TB = BT$, as claimed. $\square$

Exercise 3.24. Let $\mathcal{H} = L^2(\mathbb{R})$, and $H\colon \mathcal{D}_H \longrightarrow \mathcal{H}$, $\mathcal{D}_H := C_0^\infty(\mathbb{R})$, $H := -\frac{\mathrm{d}^2}{\mathrm{d}x^2}$.
(i) Prove that H is symmetric.
(ii) Prove that $H^* = -\frac{\mathrm{d}^2}{\mathrm{d}x^2}$ on the domain

$$\mathcal{D}_{H^*} = \{v \in \mathcal{H} : v \in C^1(\mathbb{R}),\ v' \in AC[a,b] \text{ for any } -\infty < a < b < +\infty,\ v'' \in L^2(\mathbb{R})\}.$$

Hint: To prove the inclusion of $\mathcal{D}_{H^*}$ into the right-hand side, think to Du-Bois Reymond's lemma.
(iii) Is H self-adjoint? essentially self-adjoint?

3.5 Integration with respect to a spectral family

The goal of this subsection is to extend the definition of spectral families to unbounded self-adjoint operators. We then define integration with respect to such spectral families for a wide class of functions, using the Lebesgue-Stieltjes integral.

Definition 3.25. A spectral family is a mapping $E\colon \mathbb{R} \longrightarrow \mathcal{B}(\mathcal{H})$, denoted $(E_\lambda)_{\lambda\in\mathbb{R}}$, so that

(i) E_λ is a projection for all $\lambda \in \mathbb{R}$.

(ii) If $\lambda < \mu$, then $E_\lambda \le E_\mu$.

(iii) E is strongly left-continuous, *i.e.*

$$\forall u \in \mathcal{H}, \forall \mu \in \mathbb{R},\ \lim_{\lambda\to\mu^-} E_\lambda u = E_\mu u.$$

(iv) For all $u \in \mathcal{H}$, $\lim\limits_{\lambda\to-\infty} E_\lambda u = 0$ and $\lim\limits_{\lambda\to\infty} E_\lambda u = u$.

The next lemma follows immediately from our proof of the spectral theorem for bounded operators.

Lemma 3.26. Let $(E_\lambda)_{\lambda\in\mathbb{R}}$ be a family satisfying (i) and (ii) of the above definition. Then, for all $\mu \in \mathbb{R}$, there exists two projections E_{μ^+}, E_{μ^-} so that

$$\lim_{\lambda\to\mu^+} E_\lambda u = E_{\mu^+} u,\ \lim_{\lambda\to\mu^-} E_\lambda u = E_{\mu^-} u$$

for all $u \in \mathcal{H}$.

Here is a key observation. If $(E_\lambda)_{\lambda\in\mathbb{R}}$ is a spectral family, and $u \in \mathcal{H}$, the function

$$F_u\colon \mathbb{R} \longrightarrow \mathbb{R}$$

$$\lambda \longmapsto \|E_\lambda u\|^2$$

is increasing, left-continuous, bounded, and

$$\lim_{\lambda \to -\infty} F_u(\lambda) = 0, \ \lim_{\lambda \to \infty} F_u(\lambda) = \|u\|^2.$$

Thus there is a well-defined and unique associated Lebesgue-Stieltjes measure $\mu_{\|E_\lambda u\|^2}$, so that

$$\mu_{\|E_\lambda u\|^2}([a,b)) = F_u(b) - F_u(a)$$

for any $a < b \in \mathbb{R}$.

Definition 3.27. Let $(E_\lambda)_{\lambda \in \mathbb{R}}$ be a spectral family.
We say that $f\colon \mathbb{R} \longrightarrow \mathbb{C}$ is $E-$measurable if f is $\mu_{\|E_\lambda u\|^2}-$measurable, for all $u \in \mathcal{H}$.

Note that any Lebesgue-measurable function is $E-$measurable.

To define the integral with respect to a spectral family, we start with a step function

$$t = \sum_{k=0}^{n} c_k \mathbf{1}_{I_k}$$

where $c_0, \ldots, c_n \in \mathbb{C}$ and $I_0, \ldots, I_n$ are non-empty disjoint intervals of the form

$$(a,b), (a,b], [a,b) \text{ or } [a,b].$$

The integral of t with respect to $(E_\lambda)_{\lambda \in \mathbb{R}}$ is now defined as

$$\int_{\mathbb{R}} t(\lambda)\ \mathrm{d}E_\lambda := \sum_{k=0}^{n} c_k E_\lambda(I_k)$$

where $E_\lambda((a,b)) := E_b - E_{a^+}$, $E_\lambda((a,b]) := E_{b^+} - E_{a^+}$, $E_\lambda([a,b)) := E_b - E_a$ and $E_\lambda([a,b]) = E_{b^+} - E_a$.

For any $u \in \mathcal{H}$, we compute that

$$\begin{aligned}
\left\| \left(\int_{\mathbb{R}} t(\lambda)\ \mathrm{d}E_\lambda \right) u \right\|^2 &= \left\langle \sum_{i=0}^{n} c_i E_\lambda(I_i) u, \sum_{j=0}^{n} c_j E_\lambda(I_j) u \right\rangle \\
&= \sum_{i=0}^{n} \sum_{j=0}^{n} c_i \overline{c_j} \langle E_\lambda(I_i) u, E_\lambda(I_j) u \rangle \\
&= \sum_{i=0}^{n} \sum_{j=0}^{n} c_i \overline{c_j} \langle E_\lambda(I_j) E_\lambda(I_i) u, u \rangle \\
&= \sum_{i=0}^{n} |c_i|^2 \langle E_\lambda(I_i)^2 u, u \rangle
\end{aligned}$$

$$= \sum_{i=0}^{n} |c_i|^2 \|E_\lambda(I_i)u\|^2$$
$$= \int_{\mathbb{R}} |t(\lambda)|^2 \, \mathrm{d}\mu_{\|E_\lambda u\|^2}$$

where the last integral is the Lebesgue-Stieltjes integral of $|t|^2$ with respect to $\mu_{\|E_\lambda u\|^2}$. This identity allows us to define the integral of any E−measurable function with respect to $(E_\lambda)_{\lambda\in\mathbb{R}}$. Indeed, if $f\colon \mathbb{R} \longrightarrow \mathbb{C}$ is such a function, pick a sequence $(t_n)_{n\in\mathbb{N}}$ that converges to f in L^2. In particular, $(t_n)_{n\in\mathbb{N}}$ is a Cauchy sequence, and the identity

$$\left\| \left(\int_{\mathbb{R}} t_n(\lambda) \, \mathrm{d}E_\lambda \right) u - \left(\int_{\mathbb{R}} t_m(\lambda) \, \mathrm{d}E_\lambda \right) u \right\|^2 = \int_{\mathbb{R}} |t_n(\lambda) - t_m(\lambda)|^2 \, \mathrm{d}\mu_{\|E_\lambda u\|^2}$$

shows that the sequence $\left(\left(\int_{\mathbb{R}} t_n(\lambda) \, \mathrm{d}E_\lambda \right) u\right)_{n\in\mathbb{N}}$ is Cauchy in $\mathcal{H}$, and we can therefore set

$$\left(\int_{\mathbb{R}} f(\lambda) \, \mathrm{d}E_\lambda \right) u := \lim_{n\to\infty} \left(\int_{\mathbb{R}} t_n(\lambda) \, \mathrm{d}E_\lambda \right) u.$$

This does not depend on the chosen sequence $(t_n)_{n\in\mathbb{N}}$. Letting

$$\mathcal{D}_{E(f)} := \{ u \in \mathcal{H} : f \in L^2(\mathbb{R}, \mu_{\|E_\lambda u\|^2}) \}$$

we have thus defined a mapping

$$\begin{aligned} E(f)\colon \mathcal{D}_{E(f)} &\longrightarrow \mathcal{H} \\ u &\longmapsto \left(\int_{\mathbb{R}} f(\lambda) \, \mathrm{d}E_\lambda \right) u. \end{aligned}$$

We denote the operator $E(f)$ by $\int_{\mathbb{R}} f(\lambda) \, \mathrm{d}E_\lambda$ and we call it the integral of f with respect to the spectral family $(E_\lambda)_{\lambda\in\mathbb{R}}$.

This integration procedure enjoys the following main properties.

Theorem 3.28. Let $(E_\lambda)_{\lambda\in\mathbb{R}}$ be a spectral family and $f\colon \mathbb{R} \longrightarrow \mathbb{C}$ be E−measurable. Then

(i) $u \in \mathcal{D}_{E(f)} \iff \|E(f)u\|^2 = \int_{\mathbb{R}} |f|^2 \, \mathrm{d}\mu_{\|E_\lambda u\|^2} < \infty$.

(ii) If f is bounded, then $E(f) \in \mathcal{B}(\mathcal{H})$, $\mathcal{D}_{E(f)} = \mathcal{H}$, and

$$\|E(f)\| \le \operatorname*{ess\,sup}_{\lambda\in\mathbb{R}} |f(\lambda)|.$$

(iii) If $f(\lambda) = 1$ for all $\lambda \in \mathbb{R}$, then $E(f) = \mathrm{Id}_{\mathcal{H}}$.

(iv) For any $u \in \mathcal{D}_{E(f)}$, one has

$$\langle E(f)u, u \rangle = \int_{\mathbb{R}} f(\lambda) \, \mathrm{d}\mu_{\|E_\lambda u\|^2}.$$

(v) For any $a, b \in \mathbb{C}$ and $E-$measurable function $g\colon \mathbb{R} \longrightarrow \mathbb{C}$, we have $\mathcal{D}_{E(f)+E(g)} = \mathcal{D}_{E(|f|+|g|)}$ and

$$aE(f) + bE(g) \subset E(af + bg).$$

(vi) For all $\mu \in \mathbb{R}$, $E_\mu E(f) \subset E(f)E_\mu$ with equality if f is bounded.

(vii) For any $E-$measurable function $g\colon \mathbb{R} \longrightarrow \mathbb{C}$, we have $\mathcal{D}_{E(fg)} = \mathcal{D}_{E(f)} \cap \mathcal{D}_{E(g)}$ and

$$E(f)E(g) \subset E(fg).$$

(viii) $\mathcal{D}_{E(f)}$ is dense in $\mathcal{H}$.

(ix) $E(\overline{f}) = E(f)^*$ and $\mathcal{D}_{E(f)^*} = \mathcal{D}_{E(f)}$.

(x) $E(f)$ is normal, *i.e.* $E(f)E(f)^* = E(f)^*E(f)$.

Proof. (i) Choose a sequence $(t_n)_{n\in\mathbb{N}}$ converging to f in L^2. Then, for all $u \in \mathcal{H}$, we have

$$\|E(f)u\|^2 = \lim_{n\to\infty} \|E(t_n)u\|^2 = \lim_{n\to\infty} \int_{\mathbb{R}} |t_n(\lambda)|^2 \,\mathrm{d}\mu_{\|E_\lambda u\|^2} = \int_{\mathbb{R}} |f(\lambda)|^2 \,\mathrm{d}\mu_{\|E_\lambda u\|^2}$$

and this integral is finite if and only if $f \in L^2(\mathbb{R}, \mu_{\|E_\lambda u\|^2})$, *i.e.* if and only if $u \in \mathcal{D}_{E(f)}$.

(ii) Suppose f is bounded, and let $M := \operatorname{ess\,sup}_{\lambda\in\mathbb{R}} |f(\lambda)|$. For all $u \in \mathcal{H}$, we have

$$\int_{\mathbb{R}} |f|^2 \,\mathrm{d}\mu_{\|E_\lambda u\|^2} \leq M^2 \int_{\mathbb{R}} 1 \,\mathrm{d}\mu_{\|E_\lambda u\|^2} = M^2\|u\|^2 < \infty$$

whence $u \in \mathcal{D}_{E(f)}$ by (i), and $\|E(f)u\|^2 \leq M^2\|u\|^2$, so $\|E(f)\| \leq M$. In particular, $E(f) \in \mathcal{B}(\mathcal{H})$.

(iii) For all $u \in \mathcal{H}$, $E(1)u = \left(\int \mathrm{d}E_\lambda\right)u = u$.

(iv) We start by checking the equality for t a step function. Write then

$$t = \sum_{k=0}^{n} c_k \mathbf{1}_{I_k}$$

where $c_0, \dots, c_n \in \mathbb{C}$ and $I_0, \dots, I_n$ are pairwise disjoint non-empty intervals. Then it follows

$$\begin{aligned}\langle E(t)u, u\rangle &= \Big\langle \sum_{k=0}^{n} c_k E_\lambda(I_k)u, u\Big\rangle \\ &= \sum_{k=0}^{n} c_k \langle E_\lambda(I_k)u, u\rangle\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{n} c_k \mu_{\|E_\lambda u\|^2}(I_k) \\
&= \int_{\mathbb{R}} t(\lambda)\, \mathrm{d}\mu_{\|E_\lambda u\|^2}
\end{aligned}
$$

where the third equality must be checked independently, into four cases. For instance, if $I_k = [a_k, b_k)$, then indeed

$$
\langle E_\lambda(I_k)u, u\rangle = \langle E_{b_k}u, u\rangle - \langle E_{a_k}u, u\rangle = \|E_{b_k}u\|^2 - \|E_{a_k}u\|^2 = \mu_{\|E_\lambda u\|^2}(I_k)
$$

by definition of $\mu_{\|E_\lambda u\|^2}$. The three other cases are similar. Now, let $f \in L^2(\mathbb{R}, \mu_{\|E_\lambda u\|^2})$, and choose a sequence $(t_n)_{n\in\mathbb{N}}$ of step functions converging to f. We get

$$
\begin{aligned}
\langle E(f)u, u\rangle &= \lim_{n\to\infty} \langle E(t_n)u, u\rangle \\
&= \lim_{n\to\infty} \int_{\mathbb{R}} t_n(\lambda)\, \mathrm{d}\mu_{\|E_\lambda u\|^2} \\
&= \int_{\mathbb{R}} f(\lambda)\, \mathrm{d}\mu_{\|E_\lambda u\|^2}
\end{aligned}
$$

where the last equality follows from Cauchy-Schwarz and the convergence of $(t_n)_{n\in\mathbb{N}}$ to f, as

$$
\begin{aligned}
&\left| \int_{\mathbb{R}} t_n(\lambda)\, \mathrm{d}\mu_{\|E_\lambda u\|^2} - \int_{\mathbb{R}} f(\lambda)\, \mathrm{d}\mu_{\|E_\lambda u\|^2} \right| \\
&\le \int_{\mathbb{R}} |t_n(\lambda) - f(\lambda)|\, \mathrm{d}\mu_{\|E_\lambda u\|^2} \\
&\le \left(\int_{\mathbb{R}} |t_n(\lambda) - f(\lambda)|^2\, \mathrm{d}\mu_{\|E_\lambda u\|^2} \right)^{1/2} \left(\int_{\mathbb{R}} 1\, \mathrm{d}\mu_{\|E_\lambda u\|^2} \right)^{1/2} \\
&= \|t_n - f\|_2 \|u\|
\end{aligned}
$$

and as $\|t_n - f\| \longrightarrow 0$ as $n \to \infty$.

(vi) Let $\mu \in \mathbb{R}$, and let $u \in \mathcal{D}_{E(f)}$. Consider a sequence of step functions $(t_n)_{n\in\mathbb{N}}$ so that $t_n \to f$ in $L^2(\mathbb{R}, \mu_{\|E_\lambda u\|^2})$. Write

$$
t_n = \sum_{k=0}^{m_n} c_k^{(n)} \mathbf{1}_{I_k^{(n)}}.
$$

By the boundedness of E_μ, we have

$$
\begin{aligned}
E_\mu E(f)u &= \lim_{n\to\infty} E_\mu \sum_{k=0}^{m_n} c_k^{(n)} E(I_k^{(n)})u \\
&= \lim_{n\to\infty} \sum_{k=0}^{m_n} c_k^{(n)} E_\mu E(I_k^{(n)})u
\end{aligned}
$$

$$= \lim_{n\to\infty} \sum_{k=0}^{m_n} c_k^{(n)} E(I_k^{(n)}) E_\mu u$$
$$= E(f) E_\mu u$$

whence $u \in \mathcal{D}_{E(f)E_\mu}$. This proves that $E_\mu E(f) \subset E(f)E_\mu$ as claimed. Additionally, if f is bounded then $E(f)$ is bounded and defined on $\mathcal{H}$, so $E_\mu E(f)$ and $E(f)E_\mu$ agree on $\mathcal{D}_{E(f)E_\mu} = \mathcal{H}$. □

3.6 The spectral theorem for self-adjoint operators

We start with the following lemma.

Lemma 3.29. Let $\mathcal{H}_1, \mathcal{H}_2, \dots$ be a sequence of closed, pairwise orthogonal, subspaces of $\mathcal{H}$, so that
$$\mathcal{H} = \bigoplus_{i\geq 1} \mathcal{H}_i.$$
Consider a sequence $A_1, A_2, \dots$ of operators on $\mathcal{H}$ so that the restriction $A_i|_{\mathcal{H}_i}$ is a bounded symmetric operator mapping $\mathcal{H}_i$ to itself, for all $i \geq 1$. Then there exists a unique self-adjoint operator $A\colon \mathcal{D}_A \longrightarrow \mathcal{H}$ which coincides with A_i on $\mathcal{H}_i$ for all $i \geq 1$. The domain of A is given by
$$\mathcal{D}_A := \{u \in \mathcal{H} : \sum_{i\geq 1} \|A_i u_i\|^2 < \infty\}$$
and, for all $u \in \mathcal{D}_A$, we have
$$Au := \sum_{i\geq 1} A_i u_i.$$

Proof. First of all, observe that the map A defined in the statement is linear, and that its domain is dense, as if $u \in \mathcal{H}$ and $\varepsilon > 0$, there exists $N \geq 1$ so that
$$\left\| u - \sum_{i=1}^{N} u_i \right\| < \varepsilon$$
and clearly $\sum_{i=1}^N u_i \in \mathcal{D}_A$. Moreover, for $u, v \in \mathcal{D}_A$, we have
$$\langle Au, v\rangle = \sum_{i=1}^{\infty} \langle A_i u_i, v_i\rangle = \sum_{i=1}^{\infty} \langle u_i, A_i v_i\rangle = \langle u, Av\rangle$$

using the symmetry of A_i, $i \geq 1$, the continuity of the inner product and the pairwise orthogonality of the subspaces $\mathcal{H}_i$, $i \geq 1$. Hence A is symmetric, and to prove it is self-adjoint, it only remains to show $\mathcal{D}_{A^*} \subset \mathcal{D}_A$. Let thus $v \in \mathcal{D}_{A^*}$, so that

$$\langle Au, v\rangle = \langle u, A^* v\rangle$$

for all $u \in \mathcal{D}_A$. This can be written as

$$\sum_{i=1}^{\infty} \langle A_i u_i, v_i\rangle = \sum_{i=1}^{\infty} \langle u_i, (A^* v)_i\rangle.$$

In particular, for $j \geq 1$ and $u = u_j \in \mathcal{H}_j$, this reduces to $\langle A_j u_j, v_j\rangle = \langle u_j, (A^* v)_j\rangle$. On the other hand, $\langle A_j u_j, v_j\rangle = \langle u_j, A_j v_j\rangle$ by symmetry of A_j, whence

$$(A^* v)_j = A_j v_j$$

for all $j \geq 1$. We deduce that

$$\sum_{i=1}^{\infty} \|A_i v_i\|^2 = \sum_{i=1}^{\infty} \|(A^* v)_i\|^2 = \|A^* v\|^2 < \infty$$

from Pythagora's theorem. Hence $v \in \mathcal{D}_A$, as announced.

Lastly, suppose A' is another self-adjoint operator that coincides with A_i on $\mathcal{H}_i$, for any $i \geq 1$. Since A' is self-adjoint, it is closed, and well-defined at each $u \in \mathcal{H}$ for which the series

$$\sum_{i=1}^{\infty} A' u_i$$

is convergent. Furthermore, for all such u, the series converges to $A'u$. But $A'u_i = Au_i$ for all $i \geq 1$ and since all terms are pairwise orthogonal, the above series converges if and only if $\sum_{i=1}^{\infty} \|A_i u_i\|^2 < \infty$. Hence $\mathcal{D}_A \subset \mathcal{D}_{A'}$ and $A'u = Au$ for $u \in \mathcal{D}_A$. That is, $A \subset A'$. As A is self-adjoint, it is maximal symmetric, and thus $A = A'$. The proof is complete. □

We have found a way to construct unbounded self-adjoint operators from bounded symmetric ones. In fact, any self-adjoint operator can be decomposed in that way.

Lemma 3.30. Let A be a self-adjoint operator. Then there exists a sequence

$$\mathcal{H}_1, \mathcal{H}_2, \ldots$$

of closed, pairwise orthogonal, subspaces of $\mathcal{H}$ so that $\mathcal{H} = \bigoplus_{i\geq 1} \mathcal{H}_i$, and so that the restriction $A_{|\mathcal{H}_i}$ is a bounded symmetric operator mapping $\mathcal{H}_i$ to itself for all $i \geq 1$. Moreover, if $T \in \mathcal{B}(\mathcal{H})$ is so that $TA \subset AT$, the restriction of T to $\mathcal{H}_i$ is a bounded operator mapping $\mathcal{H}_i$ to itself, for all $i \geq 1$.

Proof. To come. □

We are finally in position to prove our Spectral Theorem II.

Theorem 3.31. Let A be a self-adjoint operator. Then there exists a unique spectral family $(E_\lambda)_{\lambda\in\mathbb{R}}$ so that

$$A = \int \lambda \, \mathrm{d}E_\lambda.$$

Moreover, any $T \in \mathcal{B}(\mathcal{H})$ with $TA \subset AT$ also satisfies $TE_\lambda = E_\lambda T$, for any $\lambda \in \mathbb{R}$.

Proof. Let A be a self-adjoint operator. By Lemma 3.30, we can find a sequence

$$\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3, \dots$$

of subspaces of $\mathcal{H}$ so that

$$\mathcal{H} = \bigoplus_{i\geq 1} \mathcal{H}_i$$

and $A_i := A|_{\mathcal{H}_i}$ is a bounded symmetric operator on $\mathcal{H}_i$. By Spectral Theorem I, there is a unique spectral family $(E_{\lambda,i})_{i\in\mathbb{R}}$ so that

$$A_i = \int_{m_i}^{M_i+\varepsilon} \lambda \, \mathrm{d}E_{\lambda,i}.$$

In particular, $E_{\lambda,i}$ is a bounded symmetric operator on $\mathcal{H}_i$ for any $\lambda \in \mathbb{R}$ and any $i \geq 1$, so Lemma 3.29 ensures that

$$E_\lambda u := \sum_{i=1}^{\infty} E_{\lambda,i} u_i$$

defines a self-adjoint operator on the domain

$$\mathcal{D}_{E_\lambda} := \{u \in \mathcal{H} : \sum_{i=1}^{\infty} \|E_{\lambda,i} u_i\|^2 < \infty\}.$$

We now proceed to show $(E_\lambda)_{\lambda\in\mathbb{R}}$ is the spectral family we are seeking. First of all, note that for any $u \in \mathcal{H}$, we have

$$\sum_{i=1}^{\infty} \|E_{\lambda,i} u_i\|^2 \leq \sum_{i=1}^{\infty} \|u_i\|^2 = \|u\|^2 < \infty$$

so $u \in \mathcal{D}_{E_\lambda}$ and $\|E_\lambda u\| \leq \|u\|$. Also

$$E_\lambda^2 u = E_\lambda\left(\sum_{i=1}^{\infty} E_{\lambda,i} u_i\right) = \sum_{i=1}^{\infty} E_{\lambda,i} E_{\lambda,i} u_i = \sum_{i=1}^{\infty} E_{\lambda,i} u_i = E_\lambda u$$

and thus E_λ is idempotent for any $\lambda \in \mathbb{R}$. Being symmetric and bounded, we deduce it is a projection.

Now let $\lambda < \mu$, and $u \in \mathcal{H}$. We compute that

$$\langle (E_\mu - E_\lambda)u, u\rangle = \left\langle (E_\mu - E_\lambda)\sum_{i=1}^\infty u_i, \sum_{j=1}^\infty u_j \right\rangle = \sum_{i=1}^\infty \langle (E_{\mu,i} - E_{\lambda,i})u_i, u_i\rangle$$

and as $E_{\mu,i} - E_{\lambda,i} \geq 0$, we deduce that $\langle (E_\mu - E_\lambda)u, u\rangle \geq 0$ as well, and $(E_\lambda)_{\lambda\in\mathbb{R}}$ is increasing.

Now, let $u \in \mathcal{H}$. Write

$$E_\lambda u = \sum_{i=1}^N E_{\lambda,i}u_i + \sum_{i=N+1}^\infty E_{\lambda,i}u_i.$$

Let $\eta > 0$ and choose $N_\eta \in \mathbb{N}$ so that

$$\left\| \sum_{i=N_\eta+1}^\infty u_i \right\| < \eta.$$

Let $m_\eta := \min(m_1, \ldots, m_{N_\eta})$, and let $\lambda < m_\eta$. Then

$$\begin{aligned} \|E_\lambda u\| &\leq \left\| \sum_{i=1}^{N_\eta} E_{\lambda,i}u_i \right\| + \left\| \sum_{i=N_\eta+1}^\infty E_{\lambda,i}u_i \right\| \\ &\leq \left\| E_\lambda \sum_{i=N_\eta+1}^\infty u_i \right\| \\ &\leq \eta. \end{aligned}$$

This shows that $\lim_{\lambda\to-\infty} E_\lambda u = 0$, and similarly one proves that $\lim_{\lambda\to\infty} E_\lambda u = u$. □

4. Applications to quantum mechanics

The spectral theory of unbounded self-adjoint operators is the adapted framework to develop basic ideas of quantum mechanics.

4.1 Representation of strongly continuous one-parameter unitary groups

In this part, we introduce the notion of *one-parameter unitary group*, which is the essential tool to describe the Hamiltonian of a quantum system. We can completely characterize such groups, through Stone's theorem.

Recall first that $U \in \mathcal{B}(\mathcal{H})$ is called unitary if $UU^* = U^*U = \mathrm{Id}_{\mathcal{H}}$.

Exercise 4.1. Check that $U \in \mathcal{B}(\mathcal{H})$ is unitary if and only if U is surjective and $\langle Uu, Uv\rangle = \langle u, v\rangle$ for all $u, v \in \mathcal{H}$. Is the surjectivity assumption really necessary? Deduce that a unitary operator has norm 1.

Let us now introduce new relevant terminologies.

Definition 4.2. A one-parameter unitary group is a mapping $U\colon \mathbb{R} \longrightarrow \mathcal{B}(\mathcal{H})$ so that

(i) $U(t)$ is unitary for all $t \in \mathbb{R}$.

(ii) $U(0) = \mathrm{Id}_{\mathcal{H}}$, and $U(t+s) = U(t)U(s)$ for all $t, s \in \mathbb{R}$.

We often denote a one-parameter unitary group by $(U_t)_{t\in\mathbb{R}}$. Moreover, such a group is called *strongly continuous* if the map $\mathbb{R} \longrightarrow \mathcal{H}$, $t \longmapsto U_t u$ is continuous for any $u \in \mathcal{H}$, and *weakly continuous* if the map $\mathbb{R} \longrightarrow \mathcal{H}$, $t \longmapsto \langle U_t u, v\rangle$ is continuous for any $u, v \in \mathcal{H}$.

In fact, these two properties are the same.

Exercise 4.3. Show that a one-parameter unitary group $(U_t)_{t\in\mathbb{R}}$ is strongly continuous if and only if it is weakly continuous.

Note that if $(U_t)_{t\in\mathbb{R}}$ is a one-parameter unitary group, then

$$U_{-t} = U_t^* = (U_t)^{-1}$$

for any $t \in \mathbb{R}$.

Exercise 4.4. For $a \in \mathbb{R}$, let $U_a\colon L^2(\mathbb{R}) \longrightarrow L^2(\mathbb{R})$, $(U_a f)(x) := f(x-a)$. Show that $(U_a)_{a\in\mathbb{R}}$ is a strongly continuous one-parameter unitary group.

Here is the central object of the study of one-parameter unitary groups.

Definition 4.5. Let $(U_t)_{t\in\mathbb{R}}$ be a strongly continuous one-parameter unitary group. The infinitesimal generator of $(U_t)_{t\in\mathbb{R}}$ is the operator $G\colon \mathcal{D}_G \longrightarrow \mathcal{H}$ defined on

$$\mathcal{D}_G := \left\{u \in \mathcal{H} : \lim_{t\to 0} \frac{1}{t}(U_t - \mathrm{Id}_{\mathcal{H}})u \text{ exists}\right\}$$

by

$$Gu := \lim_{t\to 0} \frac{1}{t}(U_t - \mathrm{Id}_{\mathcal{H}})u.$$

The next result is the first step towards characterization of strongly continuous one-parameter unitary groups.

Theorem 4.6. Let A be a self-adjoint operator on $\mathcal{H}$, and let $(E_\lambda)_{\lambda\in\mathbb{R}}$ be its spectral family. Then the family $(U_t)_{t\in\mathbb{R}}$ defined as

$$U_t = e^{itA} = \int e^{it\lambda}\, \mathrm{d}E_\lambda$$

is a strongly continuous one-parameter unitary group, and iA is its infinitesimal generator. Moreover, if $u \in \mathcal{D}_A$, then $U_t u \in \mathcal{D}_A$ for all $t \in \mathbb{R}$.

Proof. The fact that $(U_t)_{t\in\mathbb{R}}$ is a one-parameter unitary group is a consequence of Theorem 3.28. As $f(\lambda) = e^{it\lambda}$ is bounded on $\mathbb{R}$, U_t is a bounded operator on $\mathcal{H}$, and if $t = 0$ then

$$U_0 = \int 1\, \mathrm{d}E_\lambda = \mathrm{Id}_{\mathcal{H}}.$$

Additionally, for $s, t \in \mathbb{R}$, one has

$$U_t U_s = \int e^{it\lambda}\, \mathrm{d}E_\lambda \int e^{is\lambda}\, \mathrm{d}E_\lambda = \int e^{it\lambda} e^{is\lambda}\, \mathrm{d}E_\lambda = \int e^{i(t+s)\lambda}\, \mathrm{d}E_\lambda = U_{t+s}$$

by Theorem 3.28(vii), and by point (ix) of the same result, it follows

$$U_t^* = \int \overline{e^{it\lambda}}\, \mathrm{d}E_\lambda = \int e^{-it\lambda}\, \mathrm{d}E_\lambda = U_{-t}$$

for all $t \in \mathbb{R}$, whence $U_t U_t^* = U_t U_{-t} = \mathrm{Id}_{\mathcal{H}} = U_{-t} U_t$, and U_t is unitary for any $t \in \mathbb{R}$.

Let us now check the strong continuity. Let $t^* \in \mathbb{R}$ and let $(t_n)_{n\in\mathbb{N}} \subset \mathbb{R}$ be a sequence converging to t^*. Let $u \in \mathcal{H}$. Then one has

$$\begin{aligned}\|U_{t_n} u - U_{t^*} u\|^2 &= \left\|\left(\int (e^{it_n\lambda} - e^{it^*\lambda})\, \mathrm{d}E_\lambda\right)u\right\|^2 \\ &= \int_{\mathbb{R}} |e^{it_n\lambda} - e^{it^*\lambda}|^2\, \mathrm{d}\mu_{\|E_\lambda u\|^2}\end{aligned}$$

$$= 4\int_{\mathbb{R}} \left|\sin\left(\frac{t_n - t^*}{2}\right)\right|^2 \, \mathrm{d}\mu_{\|E_\lambda u\|^2}.$$

The quantity under the integral tends to 0 as $n \to \infty$, and is bounded by 1 which is $\mu_{\|E_\lambda u\|^2}$–integrable. It follows from the dominated convergence theorem that

$$\begin{aligned}
\lim_{n\to\infty} \|U_{t_n}u - U_{t^*}u\|^2 &= \lim_{n\to\infty} 4\int_{\mathbb{R}} \left|\sin\left(\frac{t_n - t^*}{2}\right)\right|^2 \, \mathrm{d}\mu_{\|E_\lambda u\|^2} \\
&= 4\int_{\mathbb{R}} \lim_{n\to\infty} \left|\sin\left(\frac{t_n - t^*}{2}\right)\right|^2 \, \mathrm{d}\mu_{\|E_\lambda u\|^2} \\
&= 0
\end{aligned}$$

whence $U_{t_n}u \to U_{t^*}u$ as $n \to \infty$. Thus $(U_t)_{t\in\mathbb{R}}$ is strongly continuous.

We turn now to prove that iA is the infinitesimal generator G of $(U_t)_{t\in\mathbb{R}}$. First, assume that $u \in \mathcal{D}_A$, and write

$$\left\|\left(\frac{1}{t}(U_t - \mathrm{Id}_{\mathcal{H}}) - iA\right)u\right\|^2 = \int_{\mathbb{R}} \left|\frac{1}{t}(e^{it\lambda} - 1) - i\lambda\right|^2 \, \mathrm{d}\mu_{\|E_\lambda u\|^2}.$$

When $t \to 0$, the integrand tends to 0, since

$$\lim_{t\to 0} \frac{1}{t}(e^{it\lambda} - 1) = \frac{\mathrm{d}}{\mathrm{d}t} e^{it\lambda}|_{t=0} = i\lambda.$$

On the other hand, by the mean-value theorem $|\frac{1}{t}(e^{it\lambda} - 1)| \le |\lambda|$, so

$$\left|\frac{1}{t}(e^{it\lambda} - 1) - i\lambda\right|^2 \le (|\lambda| + |\lambda|)^2 = 4\lambda^2$$

and this function is $\mu_{\|E_\lambda u\|^2}$–integrable, as

$$\int \lambda^2 \, \mathrm{d}E_\lambda = \|Au\|^2 < \infty.$$

Appealing once again the dominated convergence theorem, it follows that

$$\begin{aligned}
\lim_{t\to 0} \left\|\left(\frac{1}{t}(U_t - \mathrm{Id}_{\mathcal{H}}) - iA\right)u\right\|^2 &= \lim_{t\to 0} \int_{\mathbb{R}} \left|\frac{1}{t}(e^{it\lambda} - 1) - i\lambda\right|^2 \, \mathrm{d}\mu_{\|E_\lambda u\|^2} \\
&= \int_{\mathbb{R}} \lim_{t\to 0} \left|\frac{1}{t}(e^{it\lambda} - 1) - i\lambda\right|^2 \, \mathrm{d}\mu_{\|E_\lambda u\|^2} \\
&= 0
\end{aligned}$$

and we conclude that if $u \in \mathcal{D}_A$, the limit

$$\lim_{t\to 0} \frac{1}{t}(U_t - \mathrm{Id}_{\mathcal{H}})u$$

exists and equals iAu. This already shows $iA \subset G$.

Conversely, let $u \in \mathcal{D}_G$. Then the limit

$$\lim_{t \to 0} \frac{1}{t}(U_t - \mathrm{Id}_{\mathcal{H}})u$$

exists, and we compute that

$$\|Gu\|^2 = \lim_{t \to 0} \left\| \frac{1}{t}(U_t - \mathrm{Id}_{\mathcal{H}})u \right\|^2 = \lim_{t \to 0} \int_{\mathbb{R}} \left| \frac{1}{t}(e^{it\lambda} - 1) \right|^2 \mathrm{d}\mu_{\|E_\lambda u\|^2}.$$

Noting that $\lim\limits_{t \to 0} \left| \frac{1}{t}(e^{it\lambda} - 1) \right|^2 = \lambda^2$ and using Fatou's lemma, we end up with

$$\int_{\mathbb{R}} \lambda^2 \, \mathrm{d}\mu_{\|E_\lambda u\|^2} \leq \liminf_{t \to 0} \int_{\mathbb{R}} \left| \frac{1}{t}(e^{it\lambda} - 1) \right|^2 \mathrm{d}\mu_{\|E_\lambda u\|^2} = \|Gu\|^2 < \infty$$

ensuring that the function $f(\lambda) = \lambda$ is in $L^2(\mathbb{R}, \mu_{\|E_\lambda u\|^2})$, and thus that $u \in \mathcal{D}_A$. We conclude then that $G = iA$ is the infinitesimal generator of $(U_t)_{t \in \mathbb{R}}$.

Lastly, let $u \in \mathcal{D}_A$ and $t \in \mathbb{R}$. As $f(\lambda) = e^{it\lambda}$ is bounded, we have $E_\lambda U_t = U_t E_\lambda$ for all $\lambda \in \mathbb{R}$ by Theorem 3.28(vi), whence

$$\|E_\lambda(U_t u)\|^2 = \|U_t(E_\lambda u)\|^2 = \|E_\lambda u\|^2$$

for any $\lambda \in \mathbb{R}$. It follows

$$\int_{\mathbb{R}} \lambda^2 \, \mathrm{d}\mu_{\|E_\lambda(U_t u)\|^2} = \int_{\mathbb{R}} \lambda^2 \, \mathrm{d}\mu_{\|E_\lambda u\|^2} < \infty$$

as $u \in \mathcal{D}_A$. Hence $U_t u \in \mathcal{D}_A$ as well, and the proof is complete. □

Hence, we have a generic way of building one-parameter unitary groups from a self-adjoint operator. The next theorem, usually attributed to Stone, ensures that this is the only way of getting such a unitary group.

Theorem 4.7. Let $(U_t)_{t \in \mathbb{R}}$ be a strongly continuous one-parameter unitary group. Then there exists a unique self-adjoint operator A so that

$$U_t = e^{itA}$$

for all $t \in \mathbb{R}$. Furthermore, $U_t A \subset A U_t$ for all $t \in \mathbb{R}$.

Proof. The strategy is the following: first remark that if $A = A^*$ satisfies $U_t = e^{itA}$, then Theorem 4.6 ensures that iA is the infinitesimal generator of the group, so it is unique. Now let G be the infinitesimal generator of $(U_t)_{t \in \mathbb{R}}$ and define $A := -iG$ on

$\mathcal{D}_A = \mathcal{D}_G$. We will show that $\mathcal{D}_A$ contains a dense subset, which implies it is itself dense, that A is essentially self-adjoint, and that

$$U_t = e^{it\overline{A}}$$

for any $t \in \mathbb{R}$. Then it will follow from Theorem 4.6 that $i\overline{A}$ is the infinitesimal generator of $(U_t)_{t\in\mathbb{R}}$, so $i\overline{A} = G = iA$, and we conclude that $A = \overline{A}$ is actually self-adjoint.

Consider $\mathcal{D}$ the set of all finite linear combinations of vectors of the form

$$u_\varphi = \int_{\mathbb{R}} \varphi(t) U_t u \, \mathrm{d}t$$

where $u \in \mathcal{H}$, and $\varphi \in C_0^\infty(\mathbb{R})$. For $t \neq 0$, we have then

$$\begin{aligned}
\frac{1}{t}(U_t - I)u_\varphi &= \int_{\mathbb{R}} \frac{1}{t}(U_t - I)\varphi(s)U_s u \, \mathrm{d}s \\
&= \int_{\mathbb{R}} \frac{1}{t}\varphi(s)(U_{t+s} - U_s)u \, \mathrm{d}s \\
&= \int_{\mathbb{R}} \frac{1}{t}\varphi(s)U_{t+s}u \, \mathrm{d}s - \int_{\mathbb{R}} \frac{1}{t}\varphi(s)U_s u \, \mathrm{d}s \\
&= \int_{\mathbb{R}} \frac{1}{t}\varphi(s' - t)U_{s'}u \, \mathrm{d}s' - \int_{\mathbb{R}} \frac{1}{t}\varphi(s)U_s u \, \mathrm{d}s \\
&= \int_{\mathbb{R}} \frac{1}{t}(\varphi(s-t) - \varphi(s))U_s u \, \mathrm{d}s.
\end{aligned}$$

Now, as $t \longrightarrow 0$, the integrand converges uniformly to the function $s \longmapsto -\varphi'(s)U_s u$, and it follows that

$$\lim_{t\to 0} \frac{1}{t}(U_t - I)u_\varphi = -\int_{\mathbb{R}} \varphi'(s)U_s u \, \mathrm{d}s \in \mathcal{H}.$$

Hence $u_\varphi \in \mathcal{D}_G = \mathcal{D}_A$, and then also $\mathcal{D} \subset \mathcal{D}_A$.

We now prove that $\overline{\mathcal{D}} = \mathcal{H}$. To this end, let $u \in \mathcal{H}$, and consider $(\varphi_n)_{n\in\mathbb{N}} \in C_0^\infty(\mathbb{R})$ a sequence of functions so that $\varphi_n \geq 0$, $\varphi_n(s) = 0$ if $|s| \geq \frac{1}{n}$, and $\int_{\mathbb{R}} \varphi_n = 1$, for all $n \in \mathbb{N}$. Then one gets

$$\|u_{\varphi_n} - u\| = \left\| \int_{\mathbb{R}} \varphi_n(s)(U_s - I)u \, \mathrm{d}s \right\| \leq \int_{\mathbb{R}} \varphi_n(s) \, \mathrm{d}s \sup_{t\in[-\frac{1}{n},\frac{1}{n}]} \|(U_t - I)u\|$$

and by the strong continuity of $(U_t)_{t\in\mathbb{R}}$ the last supremum goes to 0 as $n \longrightarrow \infty$. Hence $u_{\varphi_n} \longrightarrow u$ as $n \longrightarrow \infty$, and $\mathcal{D}$ is dense in $\mathcal{H}$.

Now we show that A is essentially self-adjoint. First, observe that if $u, v \in \mathcal{D}_A$, we have

$$\langle Au, v \rangle = -i\langle Gu, v \rangle$$

$$
\begin{aligned}
&= -i \lim_{t\to 0} \big\langle \frac{1}{t}(U_t - I)u, v \big\rangle \\
&= -i \lim_{t\to 0} \big\langle u, \frac{1}{t}(U_{-t} - I)v \big\rangle \\
&= i \lim_{t\to 0} \big\langle u, \frac{1}{-t}(U_{-t} - I)v \big\rangle \\
&= i\langle u, Gv\rangle \\
&= \langle u, Av\rangle
\end{aligned}
$$

using the continuity of the inner product in both variables. Symmetry of A now follows from Lemma 3.16.

Let $u \in \operatorname{Im}(A + iI)^\perp = \operatorname{Ker}(A^* - iI)$ (by Lemma 3.7). Then, for $v \in \mathcal{D}_A$, we compute that

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\langle u, U_t v\rangle &= \lim_{h\to 0} \frac{1}{h}(\langle u, U_{t+h}v\rangle - \langle u, U_t v\rangle) \\
&= \big\langle u, \lim_{h\to 0} \frac{1}{h}(U_{t+h} - U_t)v \big\rangle \\
&= \big\langle u, \lim_{h\to 0} \frac{1}{h}(U_h - U_0)U_t v \big\rangle \\
&= \langle u, GU_t v\rangle \\
&= \langle G^* u, U_t v\rangle \\
&= i\langle A^* u, U_t v\rangle \\
&= i\langle iu, U_t v\rangle \\
&= -\langle u, U_t v\rangle
\end{aligned}
$$

using again the continuity of the inner product, the definition of G and the fact that $A^* u = iu$. Thus the function $f(t) = \langle u, U_t v\rangle$ is a solution of the first order differential equation $f' = -f$, and we conclude that $f(t) = f(0)e^{-t}$, $t \in \mathbb{R}$. However, as U_t is unitary, it has norm 1, and the Cauchy-Schwarz inequality provides

$$
|f(t)| = |\langle u, U_t v\rangle| \le \|u\|\|v\|
$$

for all $t \in \mathbb{R}$, whence f is bounded. We conclude that $f(0) = 0$, *i.e.* $\langle u, v\rangle = 0$, and this holds for any $v \in \mathcal{D}_A$. Since the latter is dense, we deduce from Remark 3.6 that $u = 0$, and thus $\operatorname{Im}(A + iI)$ is dense in $\mathcal{H}$ (as its orthogonal reduces to $\{0\}$). In a similar way, we prove that $\operatorname{Im}(A - iI)^\perp = \{0\}$, and Theorem 3.22 (that we may apply since A is symmetric) implies that A is essentially self-adjoint.

The last part of the proof aims at proving that $U_t = e^{it\overline{A}}$ for any $t \in \mathbb{R}$. Hence, let us introduce $V_t := e^{it\overline{A}}$, $t \in \mathbb{R}$. As $\overline{A}$ is self-adjoint, Theorem 4.6 ensures $(V_t)_{t\in\mathbb{R}}$ is a strongly continuous one-parameter unitary group, and we are left to show $U_t = V_t$ for all $t \in \mathbb{R}$. First, for all $u \in \mathcal{D}_A \subset \mathcal{D}_{\overline{A}}$, Theorem 4.6 implies that $V_t u \in \mathcal{D}_{\overline{A}}$, and so

$$
\lim_{h\to 0} \frac{1}{h}(V_{t+h} - V_t)u = \lim_{h\to 0} \frac{1}{h}(V_h - V_0)V_t u = i\overline{A}V_t u.
$$

Therefore, letting $w(t) = (U_t - V_t)u$, we get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t} w(t) &= (GU_t - i\overline{A}V_t)u \\ &= i(AU_t - \overline{A}V_t)u \\ &= i\overline{A}(U_t - V_t)u = i\overline{A}w(t)\end{aligned}$$

for all $u \in \mathcal{D}_A$. It follows that

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\|w(t)\|^2 &= \frac{\mathrm{d}}{\mathrm{d}t}\langle w(t), w(t)\rangle \\ &= \Big\langle \frac{\mathrm{d}}{\mathrm{d}t} w(t), w(t)\Big\rangle + \Big\langle w(t), \frac{\mathrm{d}}{\mathrm{d}t} w(t)\Big\rangle \\ &= i\langle \overline{A}w(t), w(t)\rangle + \langle w(t), i\overline{A}w(t)\rangle \\ &= i\langle \overline{A}w(t), w(t)\rangle - i\langle \overline{A}w(t), w(t)\rangle \\ &= 0\end{aligned}$$

as $\overline{A}$ is self-adjoint. Thus $\|w(t)\|^2 = \|w(0)\|^2 = 0$ for all $t \in \mathbb{R}$, whence $U_t u = V_t u$ for all $u \in \mathcal{D}_A$. As the latter is dense, we conclude that $U_t = V_t$ for any $t \in \mathbb{R}$. As explained at the beginning of the proof, we conclude that in fact $A = \overline{A}$, and $U_t = e^{itA}$ for all $t \in \mathbb{R}$. □

4.2 Basic postulates of quantum mechanics

The following postulates pertain to *quantum systems*, that is, systems which are best described by the laws of quantum mechanics. We shall see that underlying lies the essential notion of *measurement apparatus*. Indeed, since we don't have any macroscopic intuition of quantum systems, the only thing the theory predicts is the result of measurements performed on the system under given experimental conditions.

Postulate I. At any given time, the *state* of the system is represented by a vector $\psi \neq 0$ of a complex separable Hilbert space $\mathcal{H}$. Furthermore, for all $c \in \mathbb{C} \setminus \{0\}$, the vector $c\psi$ represents the same state as ψ. Thus, the states of the system are in one-to-one correspondence with the rays

$$\{c\psi : c \in \mathbb{C}\},\ \psi \neq 0$$

or, equivalently, with the orthogonal projections P_ψ onto these one-dimensional subspaces.

Postulate II. Every *observable* $\mathcal{A}$ is represented by a self-adjoint operator A on $\mathcal{H}$.

Postulate III. The result of a *measurement* of the observable $\mathcal{A}$ can only be a real number λ, eigenvalue of A.

Postulate IV. If the system is in state ψ at time t, then the probability of observing the value λ when measuring the observable $\mathcal{A}$ at time t is given by

$$\mathbb{P}(\text{measurement of } \mathcal{A} \text{ yields } \lambda) = \frac{\langle \psi, P_\lambda \psi \rangle}{\langle \psi, \psi \rangle}$$

where P_λ is the projection onto the eigenspace corresponding to the eigenvalue λ.

Postulate V. The *mean value* of A, computed over a large number of systems all prepared in state ψ, is given by

$$\langle A \rangle_\psi = \frac{\langle \psi, A\psi \rangle}{\langle \psi, \psi \rangle}.$$

Postulate VI. If the system is in state ψ, then immediately after a measurement of $\mathcal{A}$ yielding the value λ, the system is in state $\phi = P_\lambda \psi$, and thus ϕ is an eigenvector of A with eigenvalue λ.

Postulate VII. There exists a self-adjoint operator H, called the *Hamiltonian* and representing the energy of the system, so that the time evolution of the system is given by the *Schrödinger equation*

$$i\hbar \partial_t \psi_t = H\psi_t$$

where ψ_t is the state of the system in time t and $\hbar = \frac{h}{2\pi}$ is the reduced Planck constant.

An important observation for what comes next is the following: consider a quantum system and let $\mathcal{H}$ be the state space. Suppose $\psi \in \mathcal{H}$ of the system at a given time is normalized (*i.e.* $\|\psi\| = 1$). Consider an observable $\mathcal{A}$ represented by a self-adjoint operator A, with spectral family $(E_\lambda)_{\lambda \in \mathbb{R}}$ so that

$$A = \int \lambda \, \mathrm{d}E_\lambda.$$

As ψ is normalized, we have

$$1 = \|\psi\|^2 = \int_{\mathbb{R}} \mathrm{d}\mu_{\|E_\lambda \psi\|^2}$$

and thus $\mu_{\|E_\lambda \psi\|^2}$ is a probability measure on $\mathbb{R}$. Then, from Theorem 3.28, the mean value of A in state ψ, as defined in Postulate V, is

$$\langle A \rangle_\psi = \langle \psi, A\psi \rangle = \int_{\mathbb{R}} \lambda \, \mathrm{d}\mu_{\|E_\lambda \psi\|^2}$$

which is exactly the expectation value of the probability measure $\mu_{\|E_\lambda \psi\|^2}$. The function $\lambda \longmapsto F_\psi(\lambda) = \|E_\lambda \psi\|^2 = \langle E_\lambda \psi, \psi \rangle$ thus represents the *distribution function* of the observable $\mathcal{A}$ in state ψ. It is indeed increasing, left-continuous, satisfies

$$F_\psi(\lambda) = \int_{-\infty}^{\lambda} \mathrm{d}\mu_{\|E_\lambda \psi\|^2}$$

and, in particular, we have

$$\lim_{\lambda \to -\infty} F_\psi(\lambda) = 0, \ \lim_{\lambda \to \infty} F_\psi(\lambda) = 1.$$

Moreover, if F_ψ is absolutely continuous on any finite interval, then the measure $\mu_{\|E_\lambda \psi\|^2}$ has a probability density φ_ψ satisfying

$$\mathrm{d}\mu_{\|E_\lambda \psi\|^2} = \varphi_\psi(\lambda)\ \mathrm{d}\lambda, \ \varphi_\psi(\lambda) = F'_\psi(\lambda) \text{ a.e. } \lambda \in \mathbb{R}.$$

We say that a quantum system on a state space $\mathcal{H}$ is *invariant under time translations* if its evolution is governed by a strongly continuous one-parameter unitary group $(U_t)_{t\in\mathbb{R}}$, in the sense that if the system is in a normalized state ψ_0 at time $t_0 = 0$, it is in state

$$\psi_t = U_t \psi_0$$

at time t.

Indeed, in this case, the system "does not see time translations", as if $t \longmapsto t' = t+\tau$ is a new time frame, then $t'_0 = \tau$ and

$$\psi_{t'} = \psi_{t+\tau} = U_{t+\tau}\psi_0 = U_t U_\tau \psi_0 = U_t \psi_\tau = U_t \psi_{t'_0}$$

meaning that $(U_t)_{t\in\mathbb{R}}$ also governes the evolution of the translated system.

For such systems, Stone's theorem (Theorem 4.7) ensures the existence of a self-adjoint operator G, the infinitesimal generator of the group $(U_t)_{t\in\mathbb{R}}$. Recall that

$$\mathcal{D}_G := \{\psi \in \mathcal{H} : \lim_{t\to 0} \frac{1}{t}(U_t - \mathrm{Id}_{\mathcal{H}})\psi \text{ exists}\}$$

and that

$$G\psi := \lim_{t\to 0} \frac{1}{t}(U_t - \mathrm{Id}_{\mathcal{H}})\psi, \ \psi \in \mathcal{D}_G.$$

We define then a self-adjoint operator H by $H := i\hbar G$ on $\mathcal{D}_H = \mathcal{D}_G$. By Theorem 4.6, if $\psi \in \mathcal{D}_G = \mathcal{D}_H$ then $U_t\psi \in \mathcal{D}_H$ as well for any $t \in \mathbb{R}$ and thus, given an initial state $\psi_0 \in \mathcal{D}_H$, one has

$$\partial_t \psi_t = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}(\psi_{t+\varepsilon} - \psi_t) = \lim_{\varepsilon\to 0} \frac{1}{\varepsilon}(U_\varepsilon - \mathrm{Id}_{\mathcal{H}})\psi_t = G\psi_t = -\frac{i}{\hbar} H\psi_t$$

which is exactly the Schrödinger equation. We thus conclude that Postulate VII is equivalent to requiring that the system is invariant under time translations.

Additionally, by Stone's theorem $U_t H\psi = HU_t\psi$ for all $\psi \in \mathcal{D}_H$, whence

$$\langle H\rangle_{\psi_t} = \langle \psi_t, H\psi_t\rangle = \langle U_t\psi_0, HU_t\psi_0\rangle = \langle U_t\psi_0, U_tH\psi_0\rangle = \langle \psi_0, H\psi_0\rangle = \langle H\rangle_{\psi_0}$$

as U_t is unitary for every $t \in \mathbb{R}$. This means that the energy is preserved by the evolution of the system.

4.3 The quantum particle on $\mathbb{R}$

Let us now turn to the description of a concrete quantum system, the particle on the real line. It is characterized by the following properties:

(i) To any Borel subset $\Delta \subset \mathbb{R}$ one can associate a measurement device, *i.e.* an observable $\mathcal{P}_\Delta$ represented by a self-adjoint operator P_Δ, called *particle detector*, taking the value 0 or 1 depending on whether the particle is in Δ or not.

(ii) The set of all operators P_Δ, $\Delta \subset \mathbb{R}$, forms a family of pairwise commuting self-adjoint operators.

(iii) To every $a \in \mathbb{R}$ one can associate a translation of the detectors
$$\tau_a P_\Delta = P_{\Delta - a}, \text{ where } \Delta - a := \{q \in \mathbb{R} : q + a \in \Delta\}.$$

(iv) The only observables commuting with all the P_Δ are functions of them.

The Hilbert space representing the states of the system is $\mathcal{H} = L^2(\mathbb{R})$, and the detector $\mathcal{P}_\Delta$ is represented by the projection $P_\Delta \colon \mathcal{H} \longrightarrow \mathcal{H}$, $(P_\Delta \psi)(q) = \mathbf{1}_\Delta(q)\psi(q)$, $\psi \in \mathcal{H}$.

For any $\Delta \subset \mathbb{R}$, the probability of finding the particle in a normalized state ψ in Δ is the mean value of the observable $\mathcal{P}_\Delta$, given by
$$\langle P_\Delta \rangle_\psi = \langle \psi, P_\Delta \psi \rangle = \int_\Delta |\psi(q)|^2 \, \mathrm{d}q.$$
Therefore, the function $q \longmapsto |\psi(q)|^2$ is interpreted as the density of probability of observing the particle in state ψ.

The observable position is then naturally represented by the multiplication operator
$$(Q\psi)(q) = q\psi(q)$$
on the domain
$$\mathcal{D}_Q := \{\psi \in L^2(\mathbb{R}) : \int_{\mathbb{R}} q^2 |\psi(q)|^2 \, \mathrm{d}q < \infty\}$$
so that the mean value of the position of the particle in state ψ is exactly
$$\langle \psi, Q\psi \rangle = \int_{\mathbb{R}} q|\psi(q)|^2 \, \mathrm{d}q$$
the expectation of the probability measure $|\psi(q)|^2 \, \mathrm{d}q$. Physically, it represents the "averaged position" of the particle, when the measurement of position is performed over a large number of copies of the system in the same state ψ. The self-adjoint operator Q is therefore called the *position operator*.

Now, to each translation $a \in \mathbb{R}$, we associate a transformation of the states of the system, given by $U_a \colon L^2(\mathbb{R}) \longrightarrow L^2(\mathbb{R})$, $(U_a\psi)(q) = \psi(q-a)$. As seen in Exercise 4.4,

$(U_a)_{a\in\mathbb{R}}$ is a strongly continuous one-parameter unitary group, and thus by Stone's theorem it takes the form

$$U_a = e^{iaA}, \ a \in \mathbb{R}$$

for some self-adjoint operator A. Now, by Theorem 4.6, iA is the infinitesimal generator of $(U_a)_{a\in\mathbb{R}}$, so that

$$iA\psi = \lim_{a\to 0} \frac{1}{a}(U_a - \mathrm{Id}_{\mathcal{H}})\psi$$

for any $\psi \in \mathcal{D}_A$. Hence, for $\psi \in \mathcal{D}_A$, one gets

$$\begin{aligned}
(A\psi)(q) &= \frac{1}{i} \lim_{a\to 0} \frac{\psi(q-a) - \psi(q)}{a} \\
&= -\frac{1}{i} \lim_{a\to 0} \frac{\psi(q-a) - \psi(q)}{-a} \\
&= -\frac{1}{i} \frac{\mathrm{d}}{\mathrm{d}q}\psi(q)
\end{aligned}$$

and we thus define the *momentum operator* P on $\mathcal{D}_P = \mathcal{D}_A$ by

$$(P\psi)(q) := \frac{\hbar}{i}\frac{\mathrm{d}}{\mathrm{d}q}\psi(q)$$

so that $U_a = e^{-\frac{i}{\hbar}aP}$, for all $a \in \mathbb{R}$.

For the particle on the real line, we can push a bit further our probabilistic view exposed above. For a normalized state ψ, given an observable represented by a self-adjoint operator A, we define the *variance* of A by

$$\mathrm{Var}_\psi(A) := \langle (A - \langle A\rangle_\psi \mathrm{Id})^2\rangle_\psi = \int_{\mathbb{R}} (A - \langle A\rangle_\psi \mathrm{Id})^2 \psi(q)\overline{\psi(q)}\, \mathrm{d}q$$

and its standard deviation by

$$\Delta_\psi(A) := \sqrt{\mathrm{Var}_\psi(A)}.$$

The next theorem is fundamental in quantum mechanics.

Theorem 4.8. Let A and B be self-adjoint operators acting on $\mathcal{H} = L^2(\mathbb{R})$. Then the commutator $C = AB - BA$ satisfies

$$|\langle C\rangle_\psi| \leq 2\Delta_\psi(A)\Delta_\psi(B)$$

for all $\psi \in \mathcal{D}_C$.

Proof. Let $\psi \in \mathcal{D}_C$, and introduce $S := A - \langle A \rangle_\psi \mathrm{Id}$, $T := B - \langle B \rangle_\psi \mathrm{Id}$. Those are still self-adjoint operators and it is easy to check that $C = ST - TS$. Hence

$$\begin{aligned} |\langle C \rangle_\psi| &= |\langle (ST - TS)\psi, \psi \rangle| \\ &\leq |\langle ST\psi, \psi \rangle| + |\langle TS\psi, \psi \rangle| \\ &= |\langle T\psi, S\psi \rangle| + |\langle S\psi, T\psi \rangle| \\ &\leq 2\|S\psi\|\|T\psi\| \end{aligned}$$

by the Cauchy-Schwarz inequality, and the claim follows since

$$\|S\psi\|^2 = \langle S\psi, S\psi \rangle = \langle S^2\psi, \psi \rangle = \langle S^2 \rangle_\psi = \langle (A - \langle A \rangle_\psi \mathrm{Id})^2 \rangle_\psi = \mathrm{Var}_\psi(A)$$

and similarly $\|T\psi\|^2 = \mathrm{Var}_\psi(B)$. □

It turns out that the commutator of the position and of the momentum operators takes a simple form.

Theorem 4.9. Let P be the momentum operator and Q be the position operator for the quantum particle on $\mathbb{R}$. Then

$$QP - PQ = i\hbar I$$

where I is the identity operator on $\mathcal{D}_{QP-PQ} = \mathcal{D}_{QP} \cap \mathcal{D}_{PQ}$.

Proof. Let $\psi \in \mathcal{D}_{QP-PQ}$. Then

$$((PQ)\psi)(q) = \frac{\hbar}{i}\frac{\mathrm{d}}{\mathrm{d}q}(Q\psi)(q) = \frac{\hbar}{i}(\psi(q) + q\frac{\mathrm{d}}{\mathrm{d}q}\psi(q))$$

and on the other hand

$$((QP)\psi)(q) = q(P\psi)(q) = \frac{\hbar}{i}q\frac{\mathrm{d}}{\mathrm{d}q}\psi(q).$$

Thus $(QP - PQ)\psi = -\frac{\hbar}{i}\psi = i\hbar\psi$ and the claim follows. □

Plugging this expression into Theorem 4.8, we obtain the *Heisenberg's uncertainty principle*.

Corollary 4.10. The position and momentum operators for the quantum particle on the real line satisfy

$$\Delta_\psi(P)\Delta_\psi(Q) \geq \frac{\hbar}{2}.$$