OPERATORS ON HILBERT SPACES

1. Hilbert spaces

1.1. Generalities.

Let X be a vector space over $\mathbb{C}$. An *inner product* on X is a map $\langle\cdot,\cdot\rangle\colon X\times X\to\mathbb{C}$ such that:

(1) $\langle x,y\rangle=\overline{\langle y,x\rangle}$, for all x, $y\in X$.
(2) $\langle \alpha x+\beta y,z\rangle=\alpha\langle x,z\rangle+\beta\langle y,z\rangle$, for all x, y, $z\in X$ and all α y β in $\mathbb{C}$.
(3) $\langle x,x\rangle\geq 0$ for all $x\in X$.
(4) $\langle x,x\rangle=0$ if and only if $x=0$.

The *norm associated to the inner product* is defined as

$$\|x\|=\langle x,x\rangle^{1/2}\qquad\text{for all } x\in X.$$

It is in fact a norm because it satisfies, for x, $y\in X$, and $\alpha\in\mathbb{C}$:

(1) $\|x\|\geq 0$. $\|x\|=0$ if and only if $x=0$.
(2) $\|x+y\|\leq\|x\|+\|y\|$.
(3) $\|\alpha x\|=|\alpha|\cdot\|x\|$.

A *Hilbert space* is a vector space H with an inner product such that it is a complete metric space for the distance associated to the norm associated to the inner product. That is, it is Banach space whose norm is associated to an inner product. Along these notes H will always denote a Hilbert space.

Examples.
1) The Euclidean space $\mathbb{C}^n$ is a Hilbert space for usual inner product given by:

$$\langle x,y\rangle=x_1\overline{y_1}+x_2\overline{y_2}+\cdots+x_n\overline{y_n}$$

for $x=(x_1,x_2,\ldots,x_n)$ and $y=(y_1,y_2,\ldots,y_n)$ in $\mathbb{C}^n$.

2) We denote by ℓ^2 the space of sequences $x=(x_n)_{n\geq1}$ such that $\sum_n|x_n|^2<+\infty$. This is a Hilbert space for the inner product

$$\langle x,y\rangle=\sum_{n=1}^{\infty}x_n\overline{y_n}\,,\qquad x=(x_n)_n,\ y=(y_n)_n\in\ell^2\,.$$

3) Let (Ω,Σ,μ) be a measure space. We denote by $L^2(\mu)$ the space of (classes of) measurable functions $f\colon\Omega\to\mathbb{C}$ such that $\int_\Omega|f|^2\,d\mu<+\infty$. We identify two functions when they are equal almost everywhere. $L^2(\mu)$ is a Hilbert space for the inner product

$$\langle f,g\rangle=\int f\overline{g}\,d\mu\,,\quad f,g\in L^2(\mu)\,.$$

The inner product and the norm of a Hilbert space satisfy:

- Cauchy–Bunyakowski–Schwarz Inequality:

$$|\langle x, y\rangle| \leq \|x\|\|y\| \,, \qquad \text{for all } x, y \in H.$$

- Parallelogram Law:

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2 \,, \qquad \text{for all } x, y \in H.$$

We recall that a subset A of a vector space is *convex* if $tx + (1-t)y$ belongs to A whenever x, $y \in A$ and $t \in [0,1]$.

Theorem 1.1. *Every non empty closed convex set $E \subset H$ contains an unique element x_0 of minimal norm. We have $\Re\langle x - x_0, x_0\rangle \geq 0$, for every $x \in E$.*

Proof. You can see the proof of the first part in [Co, 2.6] or [Ru, 12.3]. For the second part, by convexity we have $x_0 + t(x - x_0) \in E$, for all $t \in [0,1]$. The function

$$f(t) = \|x_0 + t(x-x_0)\|^2 = \|x_0\|^2 + 2t\Re\langle x - x_0, x_0\rangle + t^2\|x - x_0\|^2 \,, \quad t \in [0,1],$$

must have a non negative derivative at $t = 0$. It follows $\Re\langle x - x_0, x_0\rangle \geq 0$. □

Consequently given a closed convex set E and any $a \in H$, there is a unique $x_0 \in E$ which minimizes the distance to a; that is,

$$d(a, E) = \inf\{\|x - a\| : x \in E\} = \|x_0 - a\| \,.$$

1.2. **Orthogonality.**

We say that two vectors x, y in a Hilbert space H are *orthogonal* if $\langle x, y\rangle = 0$. Sometimes this will be denoted by $x \perp y$.

Given a subset A of H, the *orthogonal complement* of A is the set

$$A^\perp = \{x \in H : \langle x, a\rangle = 0 \,, \text{for all } a \in A\} \,.$$

It is easy to see that $A^\perp$ is a closed linear subspace of H.

When M is a closed linear subspace of H, and $x \in H$, we denote by $P_M x$ the element in M which minimizes the distance to x. One can see that $x - P_M x$ is in $M^\perp$. In fact this yields (see [Co, 2.6] or [Ru, 12.4])

Theorem 1.2. *If M is a closed subspace of H, then $H = M \oplus M^\perp$.*

Therefore P_M is a linear map called the *orthogonal projection of H onto M*. As a corollary we have $(M^\perp)^\perp = M$, whenever M is a closed linear subspace.

Another consequence of the previous theorem is the Riesz Representation Theorem. We denote by H^* the *dual space of H*, that is, the space of all continuous linear functionals $\Lambda \colon H \to \mathbb{C}$. This is a Banach space for the (dual) norm:

$$\|\Lambda\| = \sup\{|\Lambda x| : x \in H, \|x\| \leq 1\} = \inf\{c > 0 : |\Lambda x| \leq c\|x\| \text{ for all } x \in H\} \,.$$

Proposition 1.3 (Riesz Representation Theorem)**.** *There exists a conjugate-linear isometry $y \mapsto \Lambda_y$ from H onto H^* given by*

$$\Lambda_y x = \langle x, y\rangle \,, \qquad \textit{for all } x \in H.$$

Proof. See [Co, 3.4] or [Ru, 12.5]. □

An *orthogonal system* is a subset A of H formed by pairwise orthogonal elements; that is,

$$x, y \in A\,, \quad x \neq y \quad \Longrightarrow \quad x \perp y.$$

If A is an orthogonal system such that $\|x\| = 1$, for every $x \in A$, then we will say that A is an *orthonormal system*. We will deal mainly with orthogonal and orthonormal sequences. If $\{x_1, x_2, \ldots, x_n\}$ is an orthogonal system, we have:

- Pytagorean Theorem:

$$\|x_1 + x_2 + \cdots + x_n\|^2 = \|x_1\|^2 + \|x_2\|^2 + \cdots + \|x_n\|^2\,.$$

Proposition 1.4. *If $\{x_n\}$ is an orthogonal sequence, the following facts are equivalent:*

(a) *The series $\sum_n x_n$ is convergent (in the norm topology of H).*
(b) $\sum_n \|x_n\|^2 < +\infty$.
(c) *The series $\sum_n \langle x_n, y\rangle$ is convergent in $\mathbb{C}$, for every $y \in H$.*

Proof. See [Ru, 12.6]. □

Given an orthonormal sequence $\{e_n\}_n$. The Fourier coefficients of an element $x \in H$ is the sequence $(\langle x, e_n\rangle)_n$. This sequence is in ℓ^2, in fact we have:

Proposition 1.5 (Bessel's Inequality)**.** *Let $\{e_n\}$ be an orthonormal sequence in H, and $x \in H$. Then*

$$\sum_n |\langle x, e_n\rangle|^2 \leq \|x\|^2\,.$$

Proof. See [Co, 4.8]. □

Theorem 1.6. *Let $\{e_n\}$ be an orthonormal sequence in H. The following facts are equivalent:*

(a) *The linear span of $\{e_n : n \in \mathbb{N}\}$ is dense in H.*
(b) *For every $x \in H$, the fact $\langle x, e_n\rangle = 0$, for every n, implies $x = 0$.*
(c) $x = \sum_n \langle x, e_n\rangle e_n$, *for every $x \in H$.*
(d) $\langle x, y\rangle = \sum_n \langle x, e_n\rangle \overline{\langle y, e_n\rangle}$, *for all x, $y \in H$.*
(e) $\|x\|^2 = \sum_n |\langle x, e_n\rangle|^2$, *for every $x \in H$.* *(Parseval's Identity)*

Proof. See [Co, 4.13]. □

An orthonormal sequence $\{e_n\}_n$ (finite or infinite) whose linear span is dense in H is called an *orthonormal basis* of H. Every separable Hilbert space has an orthonormal basis.

Examples.
1) **The trigonometric system.** Consider the interval $[0, 2\pi]$ with the normalized Lebesgue measure, and denote

$$e_m(t) = e^{imt}\,, \quad t \in [0, 2\pi]\,,\; m \in \mathbb{Z}\,.$$

In $L^2[0, 2\pi]$ the system $\{e_m; m \in \mathbb{Z}\}$ is an orthonormal basis. This is the trigonometric system.

2) **The Rademacher sequence.** For every $n \in \mathbb{N}$, and $t \in [0,1]$ define

$$r_n(t) = \operatorname{sig}\left(\sin(2^n \pi t)\right),$$

where sig is the sign function given by $\operatorname{sig}(x) = 1$, if $x \geq 0$, and $\operatorname{sig}(x) = -1$, if $x < 0$. The Rademacher sequence $\{r_n\}_{n\in\mathbb{N}}$ is an othonormal system in $L^2[0,1]$ which is not an orthonormal basis.

3) **The Walsh system.** Let $\mathcal{P}_f(\mathbb{N})$ denote the family of all finite subset of $\mathbb{N}$, and for every $A \in \mathcal{P}_f(\mathbb{N})$ define

$$w_A(t) = \prod_{n\in A} r_n(t), \qquad t \in [0,1].$$

With the convention $w_\emptyset(t) = 1$, for all t. Thus $r_n = w_{\{n\}}$, for all n. The system $\{w_A : A \in \mathcal{P}_f(\mathbb{N})\}$ is an orthonormal basis of $L^2[0,1]$ called the Walsh system.

1.3. **Weak topology.**

The weak topology of a Hilbert space H is the locally convex topology generated by the family of seminorms $\{p_a : a \in H\}$ defined by

$$p_a(x) = |\langle x, a\rangle|, \qquad x \in H.$$

The weak topology is a linear topology on H. A base of neighbourhood of 0 for this topology is formed by the sets:

$$V(x_1, x_2, \ldots, x_n; \varepsilon) = \{y \in H : \max_{1\leq j\leq n} |\langle y, x_j\rangle| < \varepsilon\},$$

for $n \in \mathbb{N}$, x_1, x_2,. . . ,$x_n \in H$, and $\varepsilon > 0$.

The weak topology is the weakest topology on H for which all the functionals $\Lambda\colon H \to \mathbb{C}$, $\Lambda \in H^*$ become continuous functions.

A sequence $\{x_n\}_n$ in H converges to $a \in H$ for the weak topology if and only if

$$\langle a, x\rangle = \lim_{n\to\infty} \langle x_n, x\rangle \qquad \text{for all } x \in H.$$

Every closed subspace E of a Hilbert space is also a Hilbert space (for the same inner product). In this situation the weak topology of E coincides with the restriction to E of the weak topology of H.

The weak topology is weaker than the norm topology. Every weak closed set is norm closed. The reciprocal implication is true for convex sets:

Theorem 1.7 (Mazur)**.** *Let E be a convex subset of H, then $\overline{E}^{\text{norm}} = \overline{E}^{\text{weak}}$.*

Proof. This is a general fact for the weak topology in any Banach space. Let us see the proof for H Hilbert. The only inclusion we need to prove is $\overline{E}^{\text{weak}} \subset \overline{E}^{\text{norm}}$. Take $a \notin \overline{E}^{\text{norm}}$. As translations are continuous for both topologies, we can assume $a = 0$. An application of Theorem 1.1 gives us a point $x_0 \in \overline{E}^{\text{norm}}$ such that

$$\Re\langle x, x_0\rangle \geq \langle x_0, x_0\rangle = \|x_0\|^2 > 0, \quad \text{for all } x \in E.$$

The set $V = \{y \in H : \Re\langle y, x_0\rangle < \|x_0\|^2\}$ is a weak neighbourhood of $a = 0$ which does not meet E. So $a \notin \overline{E}^{\text{weak}}$. □

Theorem 1.8. *The closed unit ball of H is compact for the weak topology of H.*

Proof. This is a consequnce of Banach–Alaoglu Theorem ([Co, 3.1] or [Ru, 3.15]) and the fact that H is a reflexive Banach space. In fact, the closed unit ball of a Banach space X is weakly compact if and only if X is reflexive (see [Co, 4.2]). □

Proposition 1.9. *If H is separable, then the unit ball of H is metrizable for the weak topology.*

Proof. Take a sequence $\{a_n\}_n$ dense in the unit ball of H, and define

$$d(x, y) = \sum_{n=1}^{\infty} \frac{|\langle x - y, a_n\rangle|}{2^n}, \qquad \text{for } x, y \in H.$$

It is easy to check that d is a metric on H whose generated topology is weaker than the weak topology of H. On the unit ball B_H these topologies coincide because of the compactness of B_H for the weak topology. □

Corollary 1.10. *Every bounded sequence in H has a subsequence converging in the weak topology.*

Proof. We can assume that the sequence is contained in the unit ball B_H. In the case H is separable the result is clear since B_H is compact and metrizable for the weak topology. In the general case, one can consider a separable closed subspace H_0 of H containing the sequence, and recall that the weak topology of H_0 is the restriction to H_0 of the weak topology of H. □

2. Bounded operators

2.1. Some classes of operators.

A *bounded linear operator* between two Hilbert spaces H_1, H_2 is a continuous linear map $T\colon H_1 \to H_2$. We will denote by $\mathcal{L}(H_1, H_2)$ the space of all bounded linear operator from H_1 to H_2. $\mathcal{L}(H_1, H_2)$ is a Banach space for the operator norm

$$\|T\| = \sup\{\|Tx\|_{H_2} : x \in H_1,\ \|x\| \leq 1\}.$$

When $H_1 = H_2 = H$ we simply write $\mathcal{L}(H)$ for $\mathcal{L}(H, H)$. If we take the composition of operators as multiplication, then $\mathcal{L}(H)$ is a (usually non commutative) Banach algebra with unit (the identity operator).
In particular, $\|S \circ T\| \leq \|S\|\|T\|$ for S, $T \in \mathcal{L}(H)$.

Examples.

1) The space $\mathcal{L}(\mathbb{C}^n, \mathbb{C}^m)$ is identified with the $n \times m$ complex matrices.

2) If M is a closed subspace of H, then P_M, the orthogonal projection onto M, belongs to $\mathcal{L}(H)$. We have $\|P_M\| = 1$, unless $M = \{0\}$.

3) The *shift* or more concretely the *forward shift* $S\colon \ell^2 \to \ell^2$ is defined, for $x = (x_n)_{n\geq 1}$, by $Sx = (y_n)_{n\geq 1}$, with $y_1 = 0$ and $y_n = x_{n-1}$, for $n \geq 2$. That is,

$$S\colon (x_1, x_2, x_3, \ldots, x_n, x_{n+1}, \ldots) \mapsto (0, x_1, x_2, \ldots, x_{n-1}, x_n \ \ldots) \,.$$

S belongs to $\mathcal{L}(H)$ and $\|S\| = 1$. The *backward shift* B is defined by

$$B\colon (x_1, x_2, x_3 \ldots, x_n, x_{n+1}, \ldots) \mapsto (x_2, x_3, x_4, \ldots, x_{n+1}, x_{n+2} \ \ldots) \,.$$

We also have $\|B\| = 1$.

4) Let (Ω, Σ, μ) be a σ-finite measure space. For every $g \in L^\infty(\mu)$, we define the operator of *multiplication by* g as

$$M_g\colon L^2(\mu) \to L^2(\mu)\,, \qquad M_g f = g \cdot f\,, \quad f \in L^2(\mu).$$

M_g is a bounded operator and

$$\|M_g\| = \|g\|_{L^\infty} = \inf\{C > 0 : |f| \leq C \ \mu\text{–almost everywhere}\}\,.$$

A particular case of multiplication operators is the case of diagonal operators on ℓ^2. Given a bounded sequence $\alpha = (\alpha_n)_n$ of complex numbers, we define $M_\alpha\colon \ell^2 \to \ell^2$ by

$$M_\alpha\colon (x_1, x_2, x_3, \ldots, x_n, , \ldots) \mapsto (\alpha_1 x_1, \alpha_2 x_2, \alpha_3 x_3, \ldots, \alpha_n x_n, , \ldots).$$

We have $\|M_\alpha\| = \|\alpha\|_\infty = \sup_n |\alpha_n|$.

As we are dealing only with complex Hilbert spaces, we have:

Lemma 2.1. *If $T \in \mathcal{L}(H)$ and $\langle Tx, x\rangle = 0$, for every $x \in H$, then $T = 0$.*

Proof. For all x, $y \in H$ we have

$$0 = \langle T(x + y), x + y\rangle = \langle Tx, x\rangle + \langle Ty, y\rangle + \langle Tx, y\rangle + \langle Ty, x\rangle = \langle Tx, y\rangle + \langle Ty, x\rangle.$$

Changing y for iy we have then $-i\langle Tx, y\rangle + i\langle Ty, x\rangle = 0$, and therefore $\langle Tx, y\rangle - \langle Ty, x\rangle = 0$. Summing up the two equalities we get

$$\langle Tx, y\rangle = 0\,, \qquad \text{for all } x, y \in H.$$

This yields $T = 0$. □

The definition of the adjoint operator is given in the following proposition:

Proposition 2.2. *Let $T\colon H_1 \to H_2$ be a bounded operator between two Hilbert spaces. There exists a unique operator $T^*\colon H_2 \to H_1$, called the* adjoint operator *of T, such that*

$$\langle Tx, y\rangle_{H_2} = \langle x, T^* y\rangle_{H_1}\,, \qquad \text{for } x \in H_1,\ y \in H_2.$$

We have the following facts about adjoint operators:

- $(T + S)^* = T^* + S^*$, for $S, T \in \mathcal{L}(H_1, H_2)$.
- $(\alpha T)^* = \overline{\alpha} T^*$, for $\alpha \in \mathbb{C}$ and $T \in \mathcal{L}(H_1, H_2)$.
- $(S \circ T)^* = T^* \circ S^*$, for $T \in \mathcal{L}(H_1, H_2)$ and $S \in \mathcal{L}(H_2, H_3)$.
- $\|T\| = \|T^*\|$, for $T \in \mathcal{L}(H_1, H_2)$.

Definition 2.3. An operator $T \in \mathcal{L}(H)$ is said to be:

(1) *nomal* if $TT^* = T^*T$,
(2) *self–adjoint* or *hermitian* if $T^* = T$,
(3) *unitary* if $TT^* = T^*T = I$, where I is the identity on H,
(4) *idempotent* or a *projection* if $T^2 = T$.

Examples.

1) Orthogonal projections are hermitian. We have $P_M^* = P_M$.

2) The adjoint of the backward shift is the forward shift. As they do not commute, the shift is not a normal operator.

3) For $g \in L^\infty(\mu)$, the adjoint of M_g is $M_{\overline{g}}$. Consequently M_g is a normal operator. If g is almost everywhere real valued, then M_g is self–adjoint.

Proposition 2.4. *An operator $T \in \mathcal{L}(H)$ is normal iff $\|Tx\| = \|T^*x\|$, for all $x \in H$.*

Proof. Observing that $\|Tx\|^2 = \langle Tx, Tx\rangle = \langle T^*Tx, x\rangle$, we have $\|Tx\| = \|T^*x\|$, for all $x \in H$ if and only if $\langle (T^*T - TT^*)x, x\rangle$, for all $x \in H$, and, by Lemma 2.1, if and only if $T^*T = TT^*$. □

Proposition 2.5. *If $T \in \mathcal{L}(H)$ is an hermitian operator, then*

$$\|T\| = \sup\{|\langle Tx, x\rangle| : x \in H,\ \|x\| \le 1\}.$$

Proof. Let us write $\alpha = \sup\{|\langle Tx, x\rangle| : x \in H,\ \|x\| \le 1\}$. It is plain that $\alpha \le \|T\|$. For the reverse inequality, take x, y in the unit ball B_H. Since T is self–adjoint, we have $\langle Tx, y\rangle + \langle Ty, x\rangle = \langle Tx, y\rangle + \langle y, Tx\rangle = 2\Re(\langle Tx, y\rangle)$ By the definition of α, we have

$$\alpha\|x \pm y\|^2 \ge |\langle Tx \pm Ty, x \pm y\rangle| = | \pm \langle Tx, x\rangle \pm \langle Ty, y\rangle + 2\Re(\langle Tx, y\rangle)|$$

Summing the two inequalities without the modules, we have

$$\alpha(\|x + y\|^2 + \|x - y\|^2) \ge 4\Re(\langle Tx, y\rangle)\,, \qquad \text{for all } x, y \in B_H\,.$$

By the parallelogram law, we deduce

$$\alpha \ge \Re(\langle Tx, y\rangle)\,, \qquad \text{for all } x, y \in B_H\,.$$

Taking $y = Tx/\|Tx\|$, we have $\|Tx\| \le \alpha$, for all $x \in B_H$, and $\|T\| \le \alpha$. □

Proposition 2.6. *If $U \in \mathcal{L}(H)$, the following facts are equivalent:*

(1) *U is unitary.*
(2) *$U(H) = H$ and $\langle Ux, Uy\rangle = \langle x, y\rangle$ for all x, $y \in H$.*
(3) *$U(H) = H$ and $\|Ux\| = \|x\|$ for all $x \in H$.*

Proof. See [Ru, 12.13]. □

Proposition 2.7. *If $P \in \mathcal{L}(H)$ is a projection, then the space $P(H)$ is closed and $P(H) = \mathrm{Ker}\,(I - P)$.*

Proof. If $x \in P(H)$, then there exists $y \in H$ such that $x = Py$. Therefore $x = Py = P(Py) = Px$ and $x \in \mathrm{Ker}\,(I - P)$. Conversely, if $x \in \mathrm{Ker}\,(I - P)$, then $Px = x$ and $x \in P(H)$. □

Proposition 2.8. *Let $P \in \mathcal{L}(H)$ be a projection and $M = P(H)$, then the following facts are equivalent:*

(1) *P is self-adjoint.*
(2) *P is normal.*
(3) *$M^\perp = \mathrm{Ker}\,(P)$.*
(4) *$\langle Px, x\rangle = \|Px\|^2$ for all $x \in H$.*
(5) *P is the orthogonal projection over M.*

Proof. See [Ru, 12.14]. □

As a consequence, when P and Q are orthogonal projections, then $P(H) \perp Q(H)$ if and only if $0 = \langle Px, Qy\rangle = \langle QPx, y\rangle$ for all x, $y \in H$ if and only if $QP = 0 = PQ$.

2.2. **Compact operators.**

Let us denote by B_H the closed unit ball of H. An operator $T \in \mathcal{L}(H_1, H_2)$ is called *compact* if $\overline{T(B_{H_1})}$ is a compact subset of H_2. This is the general definition of compact operators between Banach spaces. In our case we can say that T is compact iff $T(B_{H_1})$ is compact. Indeed every $T \in \mathcal{L}(H_1, H_2)$ is continuous from the weak topology of H_1 to the weak topology of H_2, and then $T(B_{H_1})$ is weakly compact and therefore weakly closed and norm closed.

We will denote by $\mathcal{K}(H_1, H_2)$ the space of all compact operators from H_1 to H_2. $\mathcal{K}(H_1, H_2)$ is a closed linear subspace of $\mathcal{L}(H_1, H_2)$. We have that the composition $T \circ S$ of two bounded operators T and S is compact as soon as one of them is compact (this is the ideal property).

Finite rank bounded operators are compact. In fact the finite rank bounded operators from H_1 to H_2 form a dense subset of $\mathcal{K}(H_1, H_2)$.

A bounded operator T is compact if and only if T^* is compact.

Lemma 2.9. *Let $T \in \mathcal{L}(H)$ be a self-adjoint compact operator $T \neq 0$. Then there exists $x \in H$, with $\|x\| = 1$ and $Tx = \lambda x$, for $\lambda = \|T\|$ or $\lambda = -\|T\|$.*

Proof. By the Proposition 2.5, there exist a sequence $\{x_n\}$ in B_H such that

$$\|T\| = \lim_{n\to\infty} |\langle Tx_n, x_n\rangle| \,. \tag{1}$$

Passing to a subsequence if necessary, we can suppose that $(\langle Tx_n, x_n\rangle)_n$ is convergent to λ in $\mathbb{C}$, and $(x_n)_n$ is convergent in the weak topology to $x \in B_H$. As T is compact we have that $(Tx_n)_n$ tends to Tx in the norm topology. Therefore

$$\lambda = \lim_{n\to\infty} \langle Tx_n, x_n\rangle = \langle Tx, x\rangle \,.$$

As T is hermitian, $\lambda = \langle Tx, x\rangle = \langle x, Tx\rangle$ is real. By (1), we conclude $\|T\| = |\lambda|$, and $\lambda = \|T\|$ or $\lambda = -\|T\|$. Since

$$|\lambda| = |\langle Tx, x\rangle| \le \|T\|\|x\|\|x\| \le \|T\| = |\lambda| \,.$$

All the inequalities are equalities and $\|x\| = 1$. If $z = Tx - \lambda x = Tx - \langle Tx, x\rangle x$, we have $z \perp x$ and, by Pytagorean theorem,

$$\|Tx\|^2 = \|z\|^2 + |\lambda|^2\|x\|^2 = \|z\|^2 + \|T\|^2 \,.$$

In consequence $z = 0$, and $Tx = \lambda x$. □

Theorem 2.10. *Let $T \in \mathcal{L}(H)$ be a self-adjoint compact operator. Then, for $N \in \mathbb{N}$ or $N = \infty$, there exist an orthonormal sequence $\{e_n\}_{n=1}^N$ in H and a sequence of real numbers $\{\lambda_n\}_{n=1}^N$ such that:*

- (i) *The sequence $\{|\lambda_n|\}$ is decreasing, and if $N = \infty$, $\lim_{n\to\infty} \lambda_n = 0$.*
- (ii) *For every $x \in H$ we have $Tx = \sum_{n=1}^N \lambda_n \langle x, e_n\rangle e_n$.*

Proof. We assume $\|T\| > 0$. By Lemma 2.9, there exist $e_1 \in H$, with $\|e_1\| = 1$ and $Te_1 = \lambda_1 e_1$ for $\lambda_1 = \|T\|$ or $\lambda_1 = -\|T\|$. Define $H_2 = \{e_1\}^\perp$. For $x \in H_2$, we have

$$\langle Tx, e_1\rangle = \langle x, Te_1\rangle = \lambda_1\langle x, e_1\rangle = 0,$$

and so $Tx \in H_2$. We can then consider the operator $T|_{H_2}$ as a self–adjoint operator on H_2, and, if $\|T|_{H_2}\| > 0$, apply again Lemma 2.9. This gives us $e_2 \in H_2$ such that $Te_2 = \lambda_2 e_2$ and $\lambda_2 = \|T|_{H_2}\|$ or $\lambda_2 = -\|T|_{H_2}\|$. We continue with $H_3 = \{e_1, e_2\}^\perp$ and so on. There is two possibilities: either there exist $N \in \mathbb{N}$ such that $T|_{H_{N+1}} = 0$ and we stop, or we have $\|T|_{H_n}\| > 0$, for every n, and we put then $N = \infty$.
In both cases, we can define an orthonormal sequence $\{e_n\}_{n\le N}$, and a sequence $\{\lambda_n\}_{n\le N}$ of real numbers such that

$$Te_n = \lambda_n e_n \,, \quad |\lambda_n| = \|T|_{H_n}\|, \quad \text{and} \quad H_{n+1} = \{e_1, \dots, e_n\}^\perp \quad \text{for all } n \le N \,.$$

Observe that $\{|\lambda_n|\}$ is decreasing, and $z = x - \sum_{k=1}^n \langle x, e_k\rangle e_k$ belongs to H_{n+1} for every $x \in H$ and every $n \le N$, with $\|z\| \le \|x\|$. In consequence

$$\left\| Tx - \sum_{k=1}^n \lambda_k\langle x, e_k\rangle e_k \right\| \le \|T|_{H_{n+1}}\|\|x\| = |\lambda_{n+1}|\|x\| \,. \tag{2}$$

This gives directly (ii) in the case $N \in \mathbb{N}$.
In the case $N = \infty$, write $\beta = \lim_n |\lambda_n|$. As we have

$$\|Te_n - Te_m\|^2 = |\lambda_n|^2 + |\lambda_m|^2 \ge 2\beta^2 \,, \qquad \text{if } n \ne m \,,$$

and $\{Te_n\}$ has a convergent subsequence, necessarily we have $\beta = 0$. Taking limit in (2) when $n \to \infty$, we obtain (ii). $\square$

Theorem 2.11. *Let $T \in \mathcal{L}(H_1, H_2)$ be a compact operator. Then, for $N \in \mathbb{N}$ or $N = \infty$, there exist an orthonormal sequence $\{a_n\}_{n=1}^N$ in H_1, an orthonormal sequence $\{b_n\}_{n=1}^N$ in H_2, and a sequence of positive numbers $\{\tau_n\}_{n=1}^N$ such that:*

(i) *The sequence $\{\tau_n\}$ is decreasing, and if $N = \infty$, $\lim_{n\to\infty} \tau_n = 0$.*
(ii) *For every $x \in H_1$ we have $Tx = \sum_{n=1}^N \tau_n \langle x, a_n\rangle b_n$.*

Proof. Apply Theorem 2.10 to the self–adjoint compact operator $T^*T \in \mathcal{L}(H_1)$. We get a sequence $\{\lambda_n\}_{n\le N}$ of real numbers and an orthonormal sequence $\{e_n\}_{n\le N}$ in H_1 such that

$$T^*Tx = \sum_{n=1}^{N} \lambda_n \langle x, e_n\rangle e_n\,, \qquad \text{for all } x \in H_1\,. \tag{3}$$

We have $\lambda_n = \langle T^*Te_n, e_n\rangle = \langle Te_n, Te_n\rangle > 0$, for every $n \le N$. If $n \ne m$, we have $0 = \langle T^*Te_m, e_n\rangle = \langle Te_m, Te_n\rangle$, and then $Te_m \perp Te_n$.
Define $a_n = e_n$, and $b_n = Te_n/\|Te_n\|$, for all $n \le N$. The sequences $\{a_n\}_{n\le N}$ and $\{b_n\}_{n\le N}$ are orthonormal and $Ta_n = \tau_n b_n$ for $\tau_n = \sqrt{\lambda_n}$. If $x \in \{a_n : n \le N\}^\perp$, by (3) we have $T^*Tx = 0$, $0 = \langle T^*Tx, x\rangle = \|Tx\|^2$, and $Tx = 0$. All these facts allow to deduce (ii). $\square$

Remark 2.12. In the previous theorem $N \in \mathbb{N}$ only in the case that T has finite rank. The number τ_n coincides with the n'th approximation number of T defined by

$$a_n(T) = \inf\{\|T - R\| : R \in \mathcal{L}(H_1, H_2), \text{rank of } R < n\}\,.$$

2.3. **Spectral Theory. Analytic functional calculus.**

An operator $T \in \mathcal{L}(H)$ is invertible if there exists an operator $S \in \mathcal{L}(H)$ such that $TS = ST = I$. S is called the inverse of T and denoted by T^{-1}. By the Open Mapping Theorem, for $T \in \mathcal{L}(H)$ be invertible is necessary and sufficient that T be bijective from H onto H.
$T \in \mathcal{L}(H)$ is invertible if and only if T^* is invertible.

Definition 2.13. Let T be in $\mathcal{L}(H)$, the *spectrum* of T is the set $\sigma(T)$ of complex numbers λ such that $T - \lambda I$ is not invertible.
$\lambda \in \mathbb{C}$ is an *eigenvalue* of T if there exists $x \in H \setminus \{0\}$ such that $Tx = \lambda x$.
$\lambda \in \mathbb{C}$ is an *approximate eigenvalue* of T if there exists a sequence $\{x_n\}$ in H such that $\|x_n\| = 1$, for every n, and $\lim_{n\to\infty} Tx_n - \lambda x_n = 0$.

Observe that $\lambda \in \sigma(T)$ if and only if $\overline{\lambda} \in \sigma(T^*)$.

Lemma 2.14. *If $T \in \mathcal{L}(H)$ and $\|T\| < 1$, then $I - T$ is an invertible operator and $(I - T)^{-1} = \sum_{n=0}^\infty T^n$.*

Proof. The series converges absolutely since $\sum_n \|T^n\| \le \sum \|T\|^n < +\infty$. Define $S_n = \sum_{k=0}^n T^k$. It is easy to check

$$S_n(I - T) = (I - T)S_n = I - T^{n+1}.$$

Taking limit the lemma follows since $\|T^n\| \to 0$. ☐

Proposition 2.15. *Let $\mathcal{G}$ denote the subset of $\mathcal{L}(H)$ formed by the invertible operators. Then $\mathcal{G}$ is an open subset of $\mathcal{L}(H)$, and the application $T \mapsto T^{-1}$ is differentiable on $\mathcal{G}$.*

Proof. Let $T_0 \in \mathcal{G}$, take $R \in \mathcal{L}(H)$ with $\|R\| < 1/\|T_0^{-1}\|$. Then, by Lemma 2.14, $T_0 + R = (I - -RT_0^{-1})T_0$ is invertible since $\| - RT_0^{-1}\| < 1$. We obtain

$$(T_0+R)^{-1} = T_0^{-1}\big(I+(-RT_0^{-1})+(RT_0^{-1}RT_0^{-1})+\cdots\big) = T_0^{-1}-T_0^{-1}RT_0^{-1}+O(\|R\|^2).$$

This gives that the open ball of center T_0 and radius $1/\|T_0^{-1}\|$ is contained in $\mathcal{G}$ and the differentiability of the application $T \mapsto T^{-1}$ at T_0. The derivative at T_0 is the linear continuous map $R \mapsto -T_0^{-1}RT_0^{-1}$ from $\mathcal{L}(H)$ to $\mathcal{L}(H)$. ☐

Proposition 2.16. *The spectrum of $T \in \mathcal{L}(H)$ is a compact subset of $\mathbb{C}$. In fact we have $|\lambda| \le \|T\|$, for every $\lambda \in \sigma(T)$.*

Proof. By Proposition 2.15 and the fact that $\lambda \mapsto \lambda I - T$ is continuous, the set $\sigma(T)$ is closed. It is easy to check that

$$(\lambda I - T)^{-1} = \sum_{n=0}^{\infty} \lambda^{-n-1}T^n, \qquad \text{if } |\lambda| > \|T\|. \tag{4}$$

Then $|\lambda| \le \|T\|$, for every $\lambda \in \sigma(T)$. It follows that $\sigma(T)$ is compact. ☐

Proposition 2.17. *The map $\lambda \mapsto (\lambda I - T)^{-1}$ is an $\mathcal{L}(H)$-valued holomorphic function defined on $\mathbb{C} \setminus \sigma(T)$.*

Proof. Our map is differentiable on $\mathbb{C} \setminus \sigma(T)$ because it is the composition of the map $\lambda \mapsto \lambda I - T$ which is affine, and $T \mapsto T^{-1}$ which is a differentiable map by Proposition 2.15. ☐

Corollary 2.18. *For every $T \in \mathcal{L}(H)$ we have $\sigma(T) \neq \emptyset$.*

Proof. If $\sigma(T) = \emptyset$ we have, by Proposition 2.17, that $\lambda \mapsto (T-\lambda I)^{-1}$ is holomorphic on $\mathbb{C}$, an entire function. But this function is bounded since, by (4), we have $\lim_{\lambda\to\infty} \|(\lambda I - T)^{-1}\| = 0$. By Liouville's theorem this function is constant and this constant has to be 0. This is a contradiction. ☐

Theorem 2.19 (Spectral radius formula)**.** *For every $T \in \mathcal{L}(H)$ we have*

$$\sup\{|\lambda| : \lambda \in \sigma(T)\} = \inf_{n\ge 1} \|T^n\|^{1/n} = \lim_{n\to\infty} \|T^n\|^{1/n}.$$

Proof. For every sequence $(\alpha_n)_n$ of positive numbers satisfying $\alpha_{n+m} \le \alpha_n\alpha_m$, it can be proved that

$$\limsup_{n\to\infty} \alpha_n^{1/n} = \inf_{n\ge 1} \alpha_n^{1/n}.$$

This gives the second equality of the statement.

If $|\lambda| > \limsup_{n\to\infty} \|T^n\|^{1/n}$ the series in (4) is absolutely convergent, it yields the inverse of $\lambda I - T$, and $\lambda \notin \sigma(T)$. Therefore

$$\sup\{|\lambda| : \lambda \in \sigma(T)\} \leq \limsup_{n\to\infty} \|T^n\|^{1/n}.$$

Let $R = \sup\{|\lambda| : \lambda \in \sigma(T)\}$. The function $\Psi\colon z \mapsto (\frac{1}{z}I - T)^{-1}$ is holomorphic on the open disk of centre 0 and radius $1/R$, and its Taylor series on 0 is, by (4),

$$\Psi(z) = \sum_{n=1}^{\infty} z^n T^{n-1}\,.$$

For every continuous linear functional $\tau\colon \mathcal{L}(H) \to \mathbb{C}$, $\tau \circ \Psi$ is holomorphic in this disk, and the radius of convergence of

$$\sum_{n=1}^{\infty} z^n \tau(T^{n-1})\,.$$

is at least $1/R$. Consequently, for every $z \in \mathbb{C}$, with $|z| < 1/R$, the sequence $\{\tau(z^nT^{n-1})\}_n$ is bounded, for every τ continuous linear functional. By the uniform boundedness principle, the sequence $\{z^nT^{n-1}\}_n$ is bounded in $\mathcal{L}(H)$. Therefore, there exists $M > 0$ such that

$$M \geq |z|^{n-1}\|T^n\|\,, \qquad \text{for all } n \in \mathbb{N}.$$

Take n'th root and limit when $n \to \infty$ and you get

$$1 \geq |z| \lim \|T^n\|^{1/n}\,,$$

whenever $|z| < 1/R$. Therefore $R \geq \lim \|T^n\|^{1/n}$. □

Examples.

1) If M is a nontrivial closed subspace of H ($H \neq M \neq \{0\}$), then $\sigma(P_M) = \{0, 1\}$.

2) If S is the shift; then $\sigma(S) = \overline{\mathbb{D}} = \{z \in \mathbb{C} : |z| \leq 1\}$. Considering the backward shift B, as we have $B = S^*$ and $\sigma(S) = \{\overline{\lambda} : \lambda \in \sigma(B)\}$, it is enough to see $\sigma(B) = \overline{\mathbb{D}}$. As $\|B\| = 1$, we have $\sigma(B) \subset \overline{\mathbb{D}}$. Every $\alpha \in \mathbb{D}$ is an eigenvalue of B since

$$Bx = \alpha x\,, \qquad \text{for } x = (\alpha^n)_{n\in\mathbb{N}} \in \ell^2\,.$$

Consequently $\mathbb{D} \subset \sigma(B)$ and as the spectrum is closed, $\overline{\mathbb{D}} \subset \sigma(B)$.

3) If (Ω, Σ, μ) is a σ-finite measure space and $g \in L^\infty(\mu)$, then M_g is invertible if and only if there exists $\varepsilon > 0$ such that $|g| \geq \varepsilon$ μ-almost everywhere; that is, if and only if there exists $\varepsilon > 0$ such that $\mu\big(g^{-1}(B(0,\varepsilon))\big) = 0$.
The spectrum of M_g is the essential range of g:

$$\operatorname{ess\,rg} g = \{z \in \mathbb{C} : \mu\big(g^{-1}(B(z,\varepsilon))\big) > 0 \text{ for all } \varepsilon > 0\}.$$

It is not difficult to see that $\operatorname{ess\,rg} g$ is compact and that $g(\omega) \in \operatorname{ess\,rg} g$ for μ–almost every $\omega \in \Omega$.

The functional calculus or symbolic calculus for operators tries to give meaning in a "reasonable way" to the expression $f(T)$ when f is a complex function defined on (a subset of) $\mathbb{C}$. This reasonable way should satisfy some algebraic conditions such as $(f+g)(T) = f(T) + g(T)$, $(\alpha f)(T) = \alpha f(T)$, and $(fg)(T) = f(T) \circ g(T)$, and some convergence conditions too.

The easiest case is consider polynomials. If f is the polynomial

$$f(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n \,,$$

and $T \in \mathcal{L}(H)$, then we just define

$$f(T) = a_0 I + a_1 T + a_2 T^2 + \cdots + a_n T^n \,.$$

We extend this definition to entire functions $f \in \mathcal{H}(\mathbb{C})$ considering their Taylor series at the origin. If

$$f(z) = \sum_{n=0}^{\infty} a_n z^n \,, \qquad \text{for all } z \in \mathbb{C} \,,$$

we can define

$$f(T) = \sum_{n=0}^{\infty} a_n T^n, \qquad (\text{where } T^0 = I)$$

because this series is absolutely convergent in $\mathcal{L}(H)$. It is plain to check that this definition obeys the above algebraic conditions.

If f is a rational function with poles out of $\sigma(T)$, we can write

$$f(z) = \frac{P(z)}{(z-\lambda_1)(z-\lambda_2)\cdots(z-\lambda_n)}, \qquad \text{for } z \in \mathbb{C} \setminus \{\lambda_1, \lambda_2, \ldots, \lambda_n\};$$

where λ_1,λ_2,. . . ,$\lambda_n \notin \sigma(T)$ and they can be repeated. We define then

$$f(T) = P(T) \circ (T - \lambda_1 I)^{-1} \circ (T - \lambda_2 I)^{-1} \circ \cdots \circ (T - \lambda_n I)^{-1} \,.$$

Oberve that the different components of the above expression are commuting operators, since when $ST = TS$ and S is invertible, we also have $S^{-1}T = S^{-1}TSS^{-1} = S^{-1}STS^{-1} = TS^{-1}$.

We are going to extend the above definitions and give meaning to $f(T)$ when f is an holomorphic function defined on an open set containing the spectrum $\sigma(T)$. We need the following lemma.

Lemma 2.20. *Suppose* $T \in \mathcal{L}(H)$, $\alpha \in \mathbb{C} \setminus \sigma(T)$, *and* Γ *is a cycle in* $\mathbb{C} \setminus (\{\alpha\} \cup \sigma(T))$ *such that* $\operatorname{Ind}_\Gamma(\lambda) = 1$ *for every* $\lambda \in \sigma(T)$ *and* $\operatorname{Ind}_\Gamma(\alpha) = 0$. *Then*

$$\frac{1}{2\pi i} \oint_\Gamma (\alpha - z)^m (zI - T)^{-1} \, dz = (\alpha I - T)^m \,, \qquad m = 0, \pm 1, \pm 2, \pm 3, \ldots$$

Proof. It can be found in [Ru, 10.24] in the setting of Banach algebras. □

Let $T \in \mathcal{L}(H)$, if f is an holomorphic function defined on an open set Ω containing the spectrum of T, we can define $f(T)$ in the following way: take any cycle Γ in $\Omega \setminus \sigma(T)$ with the property that $\operatorname{Ind}_\Gamma(\lambda) = 1$ for every $\lambda \in \sigma(T)$ and $\operatorname{Ind}_\Gamma(z) = 0$, for every $z \in \mathbb{C} \setminus \Omega$, and put

$$f(T) = \frac{1}{2\pi i} \oint_\Gamma f(z)(zI - T)^{-1}\, dz\,.$$

This definition does not depend of the choice of Γ and, thanks to Lemma 2.20, when f is a polynomial, an entire function or a rational function with poles out of $\sigma(T)$ coincides with the one given previously.

Proposition 2.21. *Let Ω be an open set containing $\sigma(T)$, $\alpha \in \mathbb{C}$, and f, f_n, $g \in \mathcal{H}(\Omega)$. Then:*

(1) $(\alpha f + g)(T) = \alpha f(T) + g(T)$.
(2) *If $f_n \to f$ uniformly on compact sets of Ω, then $\|f(T) - f_n(T)\| \to 0$.*
(3) $(f \cdot g)(T) = f(T) \circ g(T) = g(T) \circ f(T)$.
(4) *$f(T)$ is invertible iff $f(\lambda) \neq 0$, for every $\lambda \in \sigma(T)$.*
(5) $\sigma\big(f(T)\big) = f\big(\sigma(T)\big) = \{f(\lambda) : \lambda \in \sigma(T)\}$. (Spectral mapping theorem)

Proof. (1) is consequence of the linearity of the integral.
(2). Fix a cycle Γ in $\Omega \setminus \sigma(T)$ with the required properties. The trace Γ^* of Γ is a compact set in Ω. Then there exists $M > 0$ such that $\|(zI - T)^{-1}\| \le M$ for every $z \in \Gamma^*$. Let $\|f\|_A$ denote the supremum of $|f|$ on the set A. We have

$$\|f(T) - f_n(T)\| \le \frac{1}{2\pi} \oint_\Gamma |f(z) - f_n(z)| M\, |dz| \le \frac{M}{2\pi} \|f - f_n\|_{\Gamma^*} \text{length}(\Gamma)\,.$$

As $f_n \to f$ uniformly on Γ^*, we get $\|f(T) - f_n(T)\| \to 0$.
(3) is plain when f and g are rational functions with poles out of $\sigma(T)$. The general case can be deduced from this applying Runge's theorem. For given f, $g \in \mathcal{H}(\Omega)$, there exist two sequences $(f_n)_n$ and $(g_n)_n$ of rational functions with poles out of Ω such that $f_n \to f$ and $g_n \to g$ uniformly on Γ^*. As $(f_n \cdot g_n)(T) = f_n(T) \circ g_n(T) = g_n(T) \circ f_n(T)$, for every n, by (2), we get (3).
(4). If $f(\lambda) = 0$, for $\lambda \in \sigma(T)$, then there exists $h \in \mathcal{H}(\Omega)$ such that $f(z) = (z - \lambda)h(z)$. By (3) we have $f(T) = (T - \lambda I) \circ h(T) = h(T) \circ (T - \lambda I)$. As $T - \lambda I$ is not invertible, then $f(T)$ is not invertible. In the opposite direction, suppose $f(\lambda) \neq 0$, for all $\lambda \in \sigma(T)$. Then $h = 1/f$ is holomorphic in a neighbourhood of $\sigma(T)$, and, by (3), $h(T)$ is the inverse of $f(T)$.
(5). Given $\alpha \in \mathbb{C}$, by (4) we have $f(T) - \alpha I$ is not invertible iff there exists $\lambda \in \sigma(T)$ such that $f(\lambda) - \alpha = 0$. That is, $\alpha \in \sigma\big(f(T)\big)$ iff $\alpha \in f\big(\sigma(T)\big)$. □

Proposition 2.22. *Suppose f is holomorphic in a neighbourhood of $\sigma(T)$ and g is holomorphic in a neighbourhood of $f\big(\sigma(T)\big)$. Then*

$$(g \circ f)(T) = g\big(f(T)\big)\,.$$

Proof. It can be found in [Ru, 10.29] in the setting of Banach algebras. □

2.4. Spectral Theory for normal operators.

Lemma 2.23. *If T is a normal operator, then:*

(1) *Every $\lambda \in \sigma(T)$ is an approximate eigenvalue of T.*
(2) $\|T\| = \sup\{|\lambda| : \lambda \in \sigma(T)\}$.

Proof. (1). Let $S = T - \lambda I$ and observe that S is also a normal operator. If λ is not an approximate eigenvalue of T, then there exists $\varepsilon > 0$ such that $\|Sx\| \geq \varepsilon\|x\|$ for every $x \in H$. This implies that S is injective and $S(H)$ is closed in H. By Proposition 2.4, we also have $\|S^*x\| \geq \varepsilon\|x\|$ for every $x \in H$. Then $S(H)^\perp = \{0\}$ and $S(H)$ is dense in H. As $S(H)$ is closed and dense, S is onto. Therefore S is invertible and $\lambda \notin \sigma(T)$.
(2). As $\langle T^*Tx, x\rangle = \|Tx\|^2$, for every $x \in H$, taking supremum for x in B_H, we get $\|T\|^2 \leq \|T^*T\|$, for all $T \in \mathcal{L}(H)$. This yields, by induction, that for all T normal

$$\|T\|^{2^{n+1}} \leq \|(T^*)^{2^n} T^{2^n}\| \leq \|T^{2^n}\|^2\,, \qquad \text{for } n = 0, 1, 2, \ldots$$

This implies $\|T\| \leq \|T^{2^n}\|^{1/2^n}$, for all n. Using the spectral radius formula, we get $\|T\| \leq \sup\{|\lambda| : \lambda \in \sigma(T)\}$. The reverse inequality is true for every operator. □

Proposition 2.24. *Let $T \in \mathcal{L}(H)$:*

(1) *If T is hermitian and $\lambda \in \sigma(T)$, then $\lambda \in \mathbb{R}$.*
(2) *If T is unitary and $\lambda \in \sigma(T)$, then $|\lambda| = 1$.*

Proof. (1). Suppose $\lambda \notin \mathbb{R}$. Then $\lambda = \alpha + i\beta$, with $\beta \neq 0$. For every $x \in H$, we have $\langle Tx, x\rangle \in \mathbb{R}$ since T is self–adjoint. Then, for all $x \in H$,

$$|\langle (T - \lambda I)x, x\rangle| = |\langle Tx, x\rangle - \alpha\langle x, x\rangle - i\beta\langle x, x\rangle| \geq |\beta|\|x\|^2\,,$$

and λ is not an approximate eigenvalue. By (1) in Lemma 2.23, $\lambda \notin \sigma(T)$.
(2). If T is unitary, then $\|Tx\| = \|x\|$, for all $x \in H$. For any $\lambda \in \mathbb{C}$ we have

$$\|Tx - \lambda x\| \geq \big|\|x\| - |\lambda|\|x\|\big| = \big|1 - |\lambda|\big|\|x\|.$$

Therefore, if $|\lambda| \neq 1$, then λ is not an approximate eigenvalue, and by (1) in Lemma 2.23, $\lambda \notin \sigma(T)$. □

Definition 2.25. Let Σ be a σ-algebra of subsets of Ω and H a Hilbert space. A *resolution of the identity* or *spectral measure* is a mapping $E\colon \Sigma \to \mathcal{L}(H)$ with the properties:

(1) $E(\emptyset) = 0$, $E(\Omega) = I$.
(2) $E(A)$ is a self-adjoint projection (an orthogonal projection) for every $A \in \Sigma$.
(3) $E(A_1 \cap A_2) = E(A_1) \circ E(A_2) = E(A_2) \circ E(A_1)$, for all A_1, $A_2 \in \Sigma$.
(4) If $A_1 \cap A_2 = \emptyset$, then $E(A_1 \cup A_2) = E(A_1) + E(A_2)$.
(5) For all x, $y \in H$, if $E_{x,y}(A) = \langle E(A)x, y\rangle$, for $A \in \Sigma$, then $E_{x,y}$ is a complex measure on Σ.

Example. Let (Ω, Σ, μ) be a σ-finite measure space. For $A \in \Sigma$, let X_A be the closed subspace of $L^2(\mu)$ formed by the functions vanishing almost everywhere in $\Omega \setminus A$. Then

$$E(A) = P_{X_A} = \text{ the orthogonal projection onto } X_A \ , \quad A \in \Sigma,$$

defines a resolution of the identity in $\mathcal{L}\big(L^2(\mu)\big)$.
Observe that every projection $E(A)$ is a multiplication operator. Indeed, if χ_A denotes the characteristic function of A, then $E(A) = M_{\chi_A}$.

Suppose E is a resolution of the identity. Let us call, for every $A \in \Sigma$, X_A the range of the projection $E(A)$; that is, $E(A)$ is the orthogonal projection onto the closed subspace X_A. If A_1, $A_2 \in \Sigma$ are disjoint sets, then, by (1) and (3), we have $X_{A_1} \perp X_{A_2}$, and by (4),

$$X_{A_1 \cup A_2} = X_{A_1} \oplus X_{A_2}.$$

We have the same, for $\{A_k\}_{1 \le k \le n}$ a finite family of pairwise disjoint sets in Σ:

$$X_A = X_{A_1} \oplus X_{A_2} \oplus \cdots \oplus X_{A_n} \ , \qquad \text{for } A = \bigcup_{k=1}^{n} A_k \, .$$

As a consequence of (2) in the following proposition, the same is true for $\{A_n\}_{n \in \mathbb{N}}$ a pairwise disjoint sequence in Σ:

$$X_A = X_{A_1} \oplus X_{A_2} \oplus \cdots \oplus X_{A_n} \oplus \cdots \ , \qquad \text{for } A = \bigcup_{n=1}^{\infty} A_n \, .$$

Proposition 2.26. *Let E be a resolution of the identity in $\mathcal{L}(H)$ defined on (Ω, Σ):*

(1) *$E_{x,x}$ is a positive measure for every $x \in H$.*
(2) *For $x \in H$, the map $A \mapsto E(A)x$ is a countably additive H-valued measure.*
(3) *If $\{A_n\} \subset \Sigma$, and $E(A_n) = 0$, for all n, then $E\big(\bigcup_{n=1}^{\infty} A_n\big) = 0$.*

Proof. (1). If P is an orthogonal projection, then $Px \perp x - Px$, for every $x \in H$, and so $\langle x, Px \rangle = \langle Px, Px \rangle \ge 0$. We deduce that $E_{x,x}(A) \ge 0$, for all $A \in \Sigma$ and all $x \in H$.
(2). Take $\{A_n\}_{n \in \mathbb{N}}$ a sequence of pairwise disjoint sets in Σ. Let $A = \bigcup_{n=1}^{\infty} A_n$. By (5) in Definition 2.25, we have

$$\text{(5)} \qquad \langle E(A)x, y \rangle = \sum_{n=1}^{\infty} \langle E(A_n)x, y \rangle \ , \qquad \text{for all } y \in H \, .$$

As we said before, the sequence $\{E(A_n)x\}_n$ is orthogonal. Then using (5) and Proposition 1.4, we have that the series $\sum_n E(A_n)x$ is convergent in the norm topology of H, and its sum is $E(A)x$.

(3). Define $B_1 = A_1$, and $B_n = A_n \setminus \bigcup_{k=1}^{n-1} A_k$, for $n \geq 2$. Then $\{B_n\}_n$ are pairwise disjoint, $B_n \subset A_n$, for all n, and

$$A = \bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n \,.$$

As $E(A_n) = 0$, and the range of the projection $E(B_n)$ is included in the range of $E(A_n)$, we have $E(B_n) = 0$ too. Since $E_{x,y}$ is countably additive, we have

$$\langle E(A)x, y\rangle = \sum_{n=1}^{\infty} \langle E(B_n)x, y\rangle = 0\,, \qquad \text{for all } x, y \in H\,.$$

Therefore $E(A) = 0$. □

We define $L^\infty(E)$ as the space of (classes of) measurable functions $f\colon \Omega \to \mathbb{C}$ which are essentially bounded; that is, there exist $C > 0$ such that $E\big(\{|f| \leq C\}\big) = 0$. This is a Banach space with the essential supremum norm.

We are going to give the definition of $\int f\,dE$. If f is a simple function and

$$f = \sum_{k=1}^{n} \alpha_k \chi_{A_k}\,, \qquad \alpha_k \in \mathbb{C}\,, \quad A_k \in \Sigma\,, \tag{6}$$

we define

$$\int f\,dE = \sum_{k=1}^{n} \alpha_k E(A_k) \quad \in \quad \mathcal{L}(H)\,.$$

It is easy to check, thanks to the finite additivity of E, that the above definition only depends on f, and not on the expression of f used in (6). We can choose in (6) the sets $\{A_k\}$ pairwise disjoint, and suppose that $E(A_k) = 0$ iff $k > l$, for certain $1 \leq l \leq n$. Then

$$\begin{aligned} \Big\| \Big(\int f\,dE \Big) x \Big\|^2 &= \sum_{k=1}^{n} |\alpha_k|^2 \|E(A_k)x\|^2 = \sum_{k=1}^{l} |\alpha_k|^2 \|E(A_k)x\|^2 \\ &\leq \big(\max_{1\leq k \leq l} |\alpha_k|^2 \big) \sum_{k=1}^{l} \|E(A_k)x\|^2 \leq \|f\|^2_{L^\infty(E)} \|x\|^2\,. \end{aligned}$$

In consequence we have $\| \int f\,dE\| \leq \|f\|_{L^\infty(E)}$. In fact, if $k \leq l$, and $x \neq 0$ belongs to the range of $E(A_k)$, then $(\int f\,dE)x = \alpha_k x$, and $\| \int f\,dE\| \leq |\alpha_k|$. Therefore $\| \int f\,dE\| = \|f\|_{L^\infty(E)}$, for every simple function f.

The map $f \mapsto \int f\,dE$ is then a continuous linear map that can be extended, by the density of simple functions in $L^\infty(E)$, to an unique continuous linear map $\Phi\colon L^\infty(E) \to \mathcal{L}(H)$. Then define $\int f\,dE = \Phi(f)$, for all $f \in L^\infty(E)$.

Observe that given $f \in L^\infty(E)$, the operator $T = \int f\,dE$ is the unique operator satisfying

$$\langle Tx, y\rangle = \int f\,dE_{x,y}\,, \qquad \text{for all } x, y \in H.$$

In fact this is trivial to check for simple functions, and follows, by density, for every $f \in L^\infty(E)$.

Proposition 2.27. *Let E be a resolution of the identity, f, $g \in L^\infty(E)$, $T = \int f\,dE$ and $S = \int g\,dE$. Then:*

1. $T^* = \int \overline{f}\,dE$.
2. *If f is essentially real-valued, then T is self-adjoint.*
3. $T \circ S = S \circ T = \int fg\,dE$.
4. $\|T\| = \|f\|_{L^\infty(E)}$.
5. *$\sigma(T)$ is the E-essential range of f; that is, $\mathbb{C} \setminus \sigma(T)$ is the biggest open subset G of $\mathbb{C}$ such that $E\big(f^{-1}(G)\big) = 0$.*

Proof. (1), (2) and (3) are easily checked if f and g are simple functions. By density they follow for all f, $g \in L^\infty(E)$.
(4) has been already proved for f simple, and follows by density for all $f \in L^\infty(E)$.
(5) We claim that T is invertible iff there exists $\varepsilon > 0$ such that $|f| \geq \varepsilon$, E-almost everywhere. As a consequence, $\lambda \in \sigma(T)$ iff $E\big[f^{-1}\big(B(\lambda, \varepsilon)\big)\big] \neq 0$, for all $\varepsilon > 0$, that is, iff λ is in the E-essential range of f.
Suppose that $|f| \geq \varepsilon$, E-almost everywhere, and $\varepsilon > 0$. Defining $h(\omega) = 1/f(\omega)$, if $|f(\omega)| \geq \varepsilon$ and $h(\omega) = 0$ elsewhere, we have $h \in L^\infty(E)$ and $fh = 1$ E-a.e.. By (3) we have that $\int h\,dE$ is the inverse of T, and T is invertible.
Suppose now that T is invertible, $0 < \varepsilon < 1/\|T^{-1}\|$, and consider $A = \{\omega \in \Omega : |f(\omega)| \leq \varepsilon\}$. It is enough to check $E(A) = 0$ for proving our claim. If $E(A) \neq 0$, take $x \neq 0$ in the range of $E(A)$, and obseve that $\|\chi_A f\|_{L^\infty(E)} \leq \varepsilon$. Then we would have

$$\|Tx\| = \|T(E(A)x)\| = \Big\|\Big(\int f\chi_A\,dE\Big)x\Big\| \leq \varepsilon\|x\| \leq \varepsilon\|T^{-1}\|\|Tx\|\,,$$

a contradiction to the fact $\varepsilon^{-1} > \|T^{-1}\|$. □

By definition, or as a consequence of (1) and (3) in the previous Proposition, we have that $\int f\,dE$ is a normal operator for every $f \in L^\infty$. Now we are going to see that every normal operator T induces a resoution of the indentity E on the Borel subsets of $\sigma(T)$ such that

$$T = \int_{\sigma(T)} z\,dE(z)\,.$$

We will use the following lemma:

Lemma 2.28. *Let T be a normal operator and $p \in \mathbb{C}[X,Y]$ a complex polynomial in two variables. Then we have*

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

Proof. The proof of this lemma is elementary in the case T is self-adjoint and more complicated in the general case. First let us see the case T self-adjoint. If $T^* = T$, then $p(T,T^*)$ becomes $q(T)$ for certain polynomial $q \in \mathbb{C}[X]$, such that $p(x,x) =$

$q(x)$, for every $x \in \mathbb{R}$. As $\sigma(T) \subset \mathbb{R}$, by (2) in Lemma 2.23 and by the spectral mapping theorem ((5) in Proposition 2.21), we conclude

$$\big\|q(T)\big\| = \sup\{|\lambda| : \lambda \in \sigma\big(q(T)\big)\} = \sup\{|q(\lambda)| : \lambda \in \sigma(T)\}\,.$$

The proof of the general case needs some elaborated facts about commutative Banach algebras. Let $\mathcal{A}_0$ be the subset of $\mathcal{L}(H)$:

$$\mathcal{A}_0 = \{p(T,T^*) : p \in \mathbb{C}[X,Y]\}\,.$$

It is easy to see that $p(T,T^*)$ is well defined thanks to the fact that T is normal. By the same reason, $\mathcal{A}_0$ is a commutative subalgebra of $\mathcal{L}(H)$. Taking its closure $\mathcal{A} = \overline{\mathcal{A}}_0$ in the norm topology we get a commutative Banach algebra $\mathcal{A}$. This algebra is closed by taking adjoint and satisfies

$$\|S^*S\| = \|S\|^2\,, \qquad \text{for all } S \in \mathcal{A}\,.$$

In fact this is true for all $S \in \mathcal{L}(H)$. So $\mathcal{A}$ is a commutative B^*-algebra (see [Ru, 11.17]). By the Gelfand-Naimark Theorem (see [Ru, 11.18]), if

$$\Delta = \{\phi\colon \mathcal{A} \to \mathbb{C} : \phi \text{ is an algebra homomorphism}\},$$

then

$$\|S\| = \sup\{|\phi(S)| : \phi \in \Delta\}\,, \quad \text{and} \quad \phi(S^*) = \overline{\phi(S)}\,, \quad \text{for all } S \in \mathcal{A}\,.$$

In particular, for $p \in \mathbb{C}[X,Y]$, as $\phi\big(p(T,T^*)\big) = p\big(\phi(T),\phi(T^*)\big)$ for all $\phi \in \Delta$,

$$\|p(T,T^*)\| = \sup\{|p(\lambda,\overline{\lambda})| : \lambda = \phi(T), \phi \in \Delta\}\,. \tag{7}$$

But, as $\mathcal{A}$ is a commutative Banach algebra, we have that $\lambda = \phi(T)$ for some $\phi \in \Delta$ if and only if $\lambda \in \sigma_{\mathcal{A}}(T)$; that is, iff $T - \lambda I$ is not invertible in $\mathcal{A}$ (see [Ru, 11.5 (e)]). But, it can be proved (see [Ru, page 321]) that, given $S \in \mathcal{A}$, S has an inverse in $\mathcal{L}(H)$ if and only if it has an inverse in $\mathcal{A}$. Then, by (7), we conclude

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

□

Theorem 2.29 (Spectral Theorem)**.** *Let $T \in \mathcal{L}(H)$ be a normal operator. There exists a unique spectral measure E defined on the σ-algebra $\mathcal{B}$ of Borel subsets of $\sigma(T)$ such that:*

$$\int_{\sigma(T)} z\, dE(z) = T\,.$$

Moreover, T commutes with the projection $E(A)$, for every Borel subset A of $\sigma(T)$.

Proof. Let $\mathcal{A}$ be now the subalgebra of $\mathcal{C}(\sigma(T))$ formed by the functions

$$z \mapsto p(z,\overline{z})\,, \qquad p \in \mathbb{C}[X,Y]\,.$$

Lemma 2.28 allows to see that the map

$$\Phi\colon p(z,\overline{z}) \mapsto p(T,T^*)$$

is a well defined linear isometry from $\mathcal{A}$ to $\mathcal{L}(H)$ that satisfies

$$\Phi(f)^* = \Phi(\overline{f}) \qquad \text{and} \qquad \Phi(fg) = \Phi(f) \circ \Phi(g)\,, \tag{8}$$

for all f, $g \in \mathcal{A}$. But $\mathcal{A}$ is a subalgebra of $\mathcal{C}(\sigma(T))$ that contains the constants, that is closed by conjugation and that separates the points. By the Stone-Weierstrass Theorem, $\mathcal{A}$ is dense in $\mathcal{C}(\sigma(T))$ and Φ can be extended to an unique linear isometry $\Phi\colon \mathcal{C}(\sigma(T)) \to \mathcal{L}(H)$ satisfying (8) for all f, $g \in \mathcal{C}(\sigma(T))$.
For every x, $y \in H$, the map

$$f \mapsto \langle \Phi(f)x, y\rangle$$

defines a continuous linear map on $\mathcal{C}(\sigma(T))$. By the Riesz Theorem, there exists a unique complex measure $\mu_{x,y}$ defined on $\mathcal{B}$ such that

$$\langle \Phi(f)x, y\rangle = \int_{\sigma(T)} f\, d\mu_{x,y}\,, \qquad \text{for all } f \in \mathcal{C}(\sigma(T))\,. \tag{9}$$

We let the reader check the following properties:

(i) $\mu_{\alpha x+z,y} = \alpha\mu_{x,y} + \mu_{z,y}$, for x, y, $z \in H$, and $\alpha \in \mathbb{C}$.
(ii) $\mu_{y,x} = \overline{\mu_{x,y}}$, for all x, $y \in H$. (Use $\Phi(f)^* = \Phi(\overline{f})$).
(iii) $\|\mu_{x,y}\| \le \|x\|\|y\|$, for all x, $y \in H$. (Use $\|\Phi(f)\| \le \|f\|$).
(iv) $\mu_{x,x} \ge 0$, and $\|\mu_{x,x}\| = \|x\|^2$, for all $x \in H$. (Use $\int 1 d\mu_{x,x} = \langle x, x\rangle = \|x\|^2$).

Let $\mathcal{X}$ be the Banach space formed by all bounded measurable functions $f\colon \sigma(T) \to \mathbb{C}$ with the norm

$$\|f\|_{\mathcal{X}} = \sup\{|f(\lambda)| : \lambda \in \sigma(T)\}\,,$$

and, for $f \in \mathcal{X}$, define $\Psi(f) \in \mathcal{L}(H)$ as the unique operator such that

$$\langle \Psi(f)x, y\rangle = \int_{\sigma(T)} f\, d\mu_{x,y}\,, \qquad \text{for all } f \in \mathcal{X}\,.$$

This can be done thanks to (i), (ii) and (iii). Using (ii) we see that $\Psi(f)^* = \Psi(\overline{f})$, and that $\Psi(f)$ is self-adjoint if $f \in \mathcal{X}$ is real-valued. Then $\Psi(f)$ is a normal operator for every $f \in \mathcal{X}$. By (9) we have $\Psi(f) = \Phi(f)$, for all $f \in \mathcal{C}(\sigma(T))$.
To finish the proof we will use the following

Claim. We have $\Psi(fg) = \Psi(f) \circ \Psi(g) = \Psi(g) \circ \Psi(f)$, for all f, $g \in \mathcal{X}$.

We prove the claim later. Defining $E(A) = \Psi(\chi_A)$, for all Borel set $A \in \mathcal{B}$, we have that $E(A)$ is an orthogonal projection for $E(A)$ is self-adjoint and idempotent since χ_A is real-valued and $\chi_A = \chi_A\chi_A$. E is a resolution of the identity:

- $E(\emptyset) = \Psi(0) = 0$, and $E(\sigma(T)) = \Psi(1) = \Phi(1) = I$.
- $\chi_{A_1\cap A_2} = \chi_{A_1}\chi_{A_2}$ yields $E(A_1 \cap A_2) = E(A_1) \circ E(A_2) = E(A_2) \circ E(A_1)$, for all A_1, $A_2 \in \mathcal{B}$.
- $\chi_{A_1\cup A_2} = \chi_{A_1} + \chi_{A_2}$, when $A_1 \cap A_2 = \emptyset$, yields $E(A_1 \cup A_2) = E(A_1) + E(A_2)$.
- $E_{x,y}(A) = \langle E(A)x, y\rangle = \mu_{x,y}(A)$, for all x, $y \in H$ and all $A \in \mathcal{B}$.

As we have, for all x, $y \in H$,

$$\int_{\sigma(T)} z\, dE_{x,y} = \int_{\sigma(T)} z\, d\mu_{x,y} = \langle \Phi(z)x, y\rangle = \langle Tx, y\rangle,$$

we conclude that $\int_{\sigma(T)} z\,dE(z) = T$.
Uniqueness of E follows from the fact that, if $\int_{\sigma(T)} z\,dE(z) = T$, for a resolution of the identity E, then

$$\int_{\sigma(T)} f(z)\,dE(z) = \Phi(f),$$

for all $f \in \mathcal{A}$ by algebraic properties and, by density, for all $f \in \mathcal{C}(\sigma(T))$. Necessarily we should have $E_{x,y} = \mu_{x,y}$, for all x, $y \in H$, and this determines E. □

Proof of the Claim. Take $\varphi \in \mathcal{C}(\sigma(T))$, and let $S = \Phi(\varphi) = \Psi(\varphi)$. Then we have

$$\int \psi\varphi\,d\mu_{x,y} = \langle \Phi(\varphi\psi)x, y\rangle = \langle \Phi(\psi)Sx, y\rangle = \int \psi\,d\mu_{Sx,y}, \tag{10}$$

for every $\psi \in \mathcal{C}(\sigma(T))$ and all x, $y \in H$. In the same way we have

$$\int \psi\varphi\,d\mu_{x,y} = \langle S\Phi(\psi)x, y\rangle = \langle \Phi(\psi)x, S^*y\rangle = \int \psi\,d\mu_{x,S^*y}. \tag{11}$$

Therefore, for all x, $y \in H$, we have

$$\varphi\,d\mu_{x,y} = d\mu_{Sx,y} = d\mu_{x,S^*y}. \tag{12}$$

By integration, the equalities (12) yields $\Psi(\varphi f) = \Psi(\varphi)\circ\Psi(f) = \Psi(f)\circ\Psi(\varphi)$, for all $\varphi \in \mathcal{C}(\sigma(T))$, and all $f \in \mathcal{X}$. And this yields, by similar calculation that (10) and (11), if $S = \Psi(f)$,

$$f\,d\mu_{x,y} = d\mu_{Sx,y} = d\mu_{x,S^*y}, \tag{13}$$

for all $f \in \mathcal{X}$, and all x, $y \in H$. Integrating now g in (13) we finally obtain

$$\Psi(fg) = \Psi(f)\circ\Psi(g) = \Psi(g)\circ\Psi(f), \qquad \text{for all } f, g \in \mathcal{X}\,.$$

□

The previous theorem allows us to extend, for a normal operator T, the analytic functional calculus to bounded measurable functions $f\colon \sigma(T) \to \mathbb{C}$. We define

$$f(T) = \int_{\sigma(T)} f(z)\,dE(z)$$

for E the spectral measure associated to T.

Proposition 2.30. *Let T be a normal operator, E the spectral measure associated to T, and f, g, $f_n\colon \sigma(T) \to \mathbb{C}$ bounded measurable functions:*

1. *$(fg)(T) = f(T)\circ g(T)$ and $(f+g)(T) = f(T) + g(T)$.*
2. *$\overline{f}(T) = \big(f(T)\big)^*$.*
3. *If $f_n \to f$ uniformly on $\sigma(T)$, then $\|f(T) - f_n(T)\| \to 0$.*
4. *If G is an open subset of $\mathbb{C}$ such that $G \cap \sigma(T) \neq \emptyset$, then $E\big(G\cap\sigma(T)\big) \neq 0$.*
5. *If f is continuous, then $\sigma\big(f(T)\big) = \{f(\lambda) : \lambda \in \sigma(T)\}$.*

Proof. (1), (2) and (3) follows from the general properties of the integral with respect to spectral measures that we have seen in Proposition 2.27.
(4). There exists $\lambda \in \sigma(T)$ and $\varepsilon > 0$, such that

$$A := \{z \in \sigma(T) : |z - \lambda| < \varepsilon\} \subset G \cap \sigma(T)\,.$$

If $E(A)$ were 0, then $|z - \lambda| \geq \varepsilon$, for E-almost every $z \in \sigma(T)$, and, as in the proof of Proposition 2.27, $\lambda I - T$ would be invertible. This would be a contradiciton with the fact that $\lambda \in \sigma(T)$.
(5). From (4) we deduce that, for any continuous function f, the E-essential range of f coincides with its range $\{f(\lambda) : \lambda \in \sigma(T)\}$. The result follows using (5) in Proposition 2.27. □

Corollary 2.31. *Let T be a normal operator. Then*
1. *T is unitary iff $|\lambda| = 1$ for every $\lambda \in \sigma(T)$.*
2. *T is self-adjoint iff $\lambda \in \mathbb{R}$ for every $\lambda \in \sigma(T)$.*

Proof. Proposition 2.24 gives the implications T self-adjoint $\implies \sigma(T) \subset \mathbb{R}$, and T unitary $\implies |\lambda| = 1$ for all $\lambda \in \sigma(T)$.
Assume now that T is normal and $\sigma(T) \subset \mathbb{R}$. Let E be the spectral measure associated to T. Then the map $z \mapsto z$ is real-valued on $\sigma(T)$, and by (2) in Proposition 2.30,

$$T^* = \int_{\sigma(T)} \overline{z}\, dE(z) = \int_{\sigma(T)} z\, dE(z) = T\,.$$

In the same way, if T is normal and $|\lambda| = 1$ for all $\lambda \in \sigma(T)$. Then $z\overline{z} = 1$, for every $z \in \sigma(T)$. By (1) in Proposition 2.30,

$$T^*T = TT^* = \int z\overline{z}\, dE(z) = \int 1\, dE(z) = I$$

□

References

[Co] J. B. Conway, *A Course in Functional Analysis*, Springer-Verlag, New York, 1985.
[Ha] P. R. Halmos, *Introduction to Hilbert Space*, Chelsea, New York, 1957.
[Ru] W. Rudin, *Functional Analysis* 2nd edition, McGraw-Hill, New York, 1991.