Vorlesung aus dem Wintersemester 2011/12

Functional Analysis 2

Prof. Thomas Østergaard Sørensen, Ph. D.

geTEXt von Viktor Kleen & Florian Stecker

Contents

0 Motivation and repetition

Recall the notion of a Banach space: A *Banach space* X is a vector space (mostly over $\mathbb{C}$) with a norm $\|\cdot\|$ such that the metric space (X, d), where $d(x, y) = \|x - y\|$, is a complete metric space. A *Hilbert space* H is a Banach space, where the norm comes from a scalar product: $\|x\| = \sqrt{\langle x, x\rangle}$. For example $L^2(\mathbb{R})$, with

$$\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)g(x)\,\mathrm{d}x$$

is a Hilbert space. So is $\ell_2(\mathbb{N})$, where

$$\ell_2(\mathbb{N}) = \Big\{(x_1, \dots) \in \mathbb{C}^{\mathbb{N}} \colon \sum_{i=1}^{\infty} |x_i|^2 < \infty\Big\}$$

Note that there are norms that do not arise from an inner product. For example ℓ_p, $p \neq 2$, is a Banach space but not a Hilbert space.

The space $B(X, Y)$ is the set of bounded linear operators $X \to Y$, i.e. $T \in B(X, Y)$ if and only if $T \colon X \to Y$ is linear and there exists $C > 0$ such that $\|Tx\|_Y \leq C\|x\|_X$ for all $x \in X$. This definition only requires normed vector space X and Y. We will write $B(X)$ for $B(X, X)$. There is a norm on $B(X, Y)$ such that

$$\|T\| = \sup_{\substack{x\in X\\ \|x\|\leq 1}} \|Tx\|_X.$$

If Y is a Banach space, so is $B(X, Y)$ with this norm.

If $Y = \mathbb{K}$, i.e. $Y = \mathbb{C}$ or $Y = \mathbb{R}$ depending on the ground field of X and Y, we call elements of $B(X, \mathbb{K})$ *bounded linear functionals* and write $X' = B(X, \mathbb{K})$. X' is called the *dual* of X. It always is a Banach space. Note that for Hilbert spaces H, "$H' = H$", in the sense that there exists an antilinear isometry $\Phi \colon H \to H', y \mapsto \langle y, -\rangle$. One has "$X \subseteq X''$" in the sense that there is a canonical embedding $\iota \colon X \to X''$ that embeds X isometrically in X''. If $X = X''$ (i.e. $\iota(X) = X''$), then X is called *reflexive*. In particular all Hilbert spaces are reflexive. So are ℓ_p and L^p for $1 < p < \infty$.

Definition 0.1. Let $T \in B(X)$ with a $\mathbb{C}$-Banach space X. We define the *resolvent set* $\rho(T)$ of T by

$$\rho(T) = \{\lambda \in \mathbb{C} \colon \ \mathrm{N}(T - \lambda I) = 0 \text{ and } \mathrm{R}(T - \lambda I) = X\} \subseteq \mathbb{C}.$$

We define the *spectrum* $\sigma(T)$ of T by $\sigma(T) = \mathbb{C} \setminus \rho(T)$.

The spectrum of T can be split in three types. There is the *point spectrum*

$$\sigma_p(T) = \{\lambda \in \mathbb{C} \colon \ \mathrm{N}(T - \lambda I) \neq 0\} \subseteq \sigma(T),$$

the *continuous spectrum*

$$\sigma_c(T) = \{\lambda \in \sigma(T) \colon \ \mathrm{N}(T - \lambda I) = 0 \text{ and } \overline{\mathrm{R}(T - \lambda I)} = X\} \subseteq \sigma(T),$$

and the *rest (residual) spectrum*

$$\sigma_r(T) = \{\lambda \in \sigma(T) \colon \ \mathrm{N}(T - \lambda I) \text{ and } \overline{\mathrm{R}(T - \lambda I)} \neq X\} \subseteq \sigma(T).$$

Remark 0.2.

1. Note that $\lambda \in \rho(T)$ if and only if $T - \lambda I \colon X \to X$ is bijective. This is equivalent to the existence of $R_\lambda(T) = (T - \lambda I)^{-1} \in B(X)$, called the *resolvent* of T at λ. The map $\rho(T) \ni \lambda \mapsto R_\lambda(T) \in B(X)$ is called the *resolvent map.*

2. If $\lambda \in \sigma_p(T)$ then there exists $x \neq 0$ such that $Tx = \lambda x$. Then λ is called an *eigenvalue*, and x an *eigenvector* of T. (However, if the space X is a space of functions, x is often called an *eigenfunction*). In this case $\mathrm{N}(T - \lambda I)$ is called the *eigenspace* of T corresponding to λ. It is a T-invariant subspace of X.

Proposition 0.3. *Let $T \in B(X)$. Then $\rho(T) \subseteq \mathbb{C}$ is an open set, hence $\sigma(T)$ is a closed subset of $\mathbb{C}$, and the resolvent map $\lambda \mapsto R_\lambda(T)$ is a complex analytic map, with*

$$\|R_\lambda(T)\|^{-1} \leq \operatorname{dist}(\lambda, \sigma(T)).$$

Here, "complex analytic" means for all $\lambda_0 \in \rho(T)$ there exist $r > 0$ and $c_j \in B(X)$ such that

$$R_\lambda(T) = \sum_{j=0}^{\infty} c_j (\lambda - \lambda_0)^j$$

for all $\lambda \in B_r(\lambda_0)$.

The aim of the course is (mainly) to study the spectrum and its properties for various classes or types of operators and to prove theorems "in analogy" to the spectral theorem of linear algebra concerning diagonalization of symmetric/self adjoint/Hermitian matrices. The theory here is, however, much more rich.

Definition 0.4. The *compact (linear) operators* from X to Y (normed vectorspaces) are defined by

$$K(X, Y) = \{T \in B(X, Y) \colon \overline{T(B_1(0))} \subseteq Y \text{ is compact}\}.$$

Remark 0.5.

1. As for $B(X)$, we write $K(X) = K(X, X)$.

2. If Y is Banach, then "$\overline{T(B_1(0))}$ is compact" can be replaced by "$T(B_1(0))$ is precompact".

3. That is, $T \in K(X, Y)$ iff T maps bounded sequences into sequences which has a convergent subsequence.

4. Given $k \in C(I^2)$, $I = [0, 1]$, define

$$(T_k f)(x) = \int_0^1 k(x, y) f(y) \,\mathrm{d}y, \quad x \in I, f \in C(I).$$

   T_k is called an *integral operator*, k is called the *integral kernel.* Then $T_k \colon X \to X$, $X = C(I)$, is a bounded linear operator which is compact.

Definition 0.6 (Banach space adjoints)**.** Let $T \in B(X, Y)$, X, Y Banach spaces, and define, for $y' \in Y'$,

$$(T'y')(x) = y'(Tx).$$

Then $T'\colon Y' \to X'$ is linear and bounded. T' is called the *(Banach space) adjoint* of T. Furthermore, $\|T'\| = \|T\|$. In fact, $-'\colon B(X, Y) \to B(Y', X')$ is a (linear) isometric embedding. It may be not surjective.

Definition 0.7 (Hilbert space adjoint)**.** Let H be a Hilbert space and let $\Phi\colon H \to H'$ be the map $y \mapsto \langle y, -\rangle$ identifying H with H', and let $T \in B(H)$. Then

$$T^* = \Phi^{-1} T' \Phi$$

is called the *Hilbert space adjoint* of T. It satisfies $\langle T^* x, y\rangle = \langle x, Ty\rangle$ for all $x, y \in H$. T is called *self-adjoint* iff $T = T^*$.

The “programme” of the course will consisct of the spectral theory for compact operators, the spectral theory for self-adjoint bounded operators, unbounded operators — in particular symmetric operators and quadratic forms — and the spectral theory for self-adjoint unbounded operators. We will also talk about the Fourier transform.

Lemma 0.8 (Algebraic properties of the adjoint)**.**

(1) $(\alpha T_1 + T_2)' = \alpha T_1' + T_2'$ *for* $T_1, T_2 \in B(X, Y)$ *and* $\alpha \in \mathbb{K}$.

(1)* $(\alpha T_1 + T_2)^* = \overline{\alpha} T_1^* + T_2^*$ *for* $T_1, T_2 \in B(H)$ *and* $\alpha \in \mathbb{K}$.

(2) $I' = I$ *for* $I \in B(X)$, $I\colon X \to X, x \mapsto x$.

(3) *For* $T_1 \in B(X, Y)$, $T_2 \in B(Y, Z)$, $(T_2 T_1)' = T_1' T_2'$.

(3)* *For* $S, Y \in B(H)$, $(ST)^* = T^* S^*$.

(4) *With* $\iota_X\colon X \to X''$ *and* $\iota_Y\colon Y \to Y''$ *the canonical embeddings and* $T \in B(X, Y)$, *we have* $T'' \iota_X = \iota_Y T$.

(4)* *For* $T \in B(H)$, $T^{**} = T$.

Proposition 0.9. *Let X and Y be Banach spaces and $T \in B(X, Y)$. Then $T^{-1} \in B(Y, X)$ exists if and only if $(T')^{-1} \in B(X', Y')$ exists and in this case, $(T^{-1})' = (T')^{-1}$ (or, in the case of $X = Y$ being Hilbert spaces $(T^*)^{-1} = (T^{-1})^*$).*

Proof. Assume T is invertible. Then (by 0.8) $I_{X'} = (I_X)' = (T^{-1}T)' = T'(T^{-1})'$ and $I_{Y'} = (I_Y)' = (TT^{-1})' = (T^{-1})'T'$.

Assume now that T' is invertible. Then T'' is invertible. In particular, T'' is a homeomorphism, hence it maps closed sets into closed sets. Recall, that $T''\iota_X = \iota_Y T$ and that ι_X and ι_Y are isometries. Hence $\mathrm{R}(\iota_Y T) = \mathrm{R}(T''\iota_X) = T''(\mathrm{R}(\iota_X))$. This is a closed subspace of Y'', since $\mathrm{R}(\iota_X)$ is closed [take a convergent sequence $\{\iota_X(x_n)\}$ in $\mathrm{R}(\iota_X)$; this is Cauchy, so the corresponding sequence $\{x_n\}$ in X is also Cauchy, hence convergent $x_n \to x$, so $\iota_X(x_n) \to \iota(x) \in \mathrm{R}(\iota_X)$]. Hence, $\mathrm{R}(T) = \iota_Y^{-1}(\mathrm{R}(\iota_Y T))$ is closed. Since T'

is injective, it follows that $0 = \mathrm{N}(T') = \mathrm{Ann}(\mathrm{R}(T))$ by the next lemma. From this and the following proposition (a consequence of Hahn-Banach) we have $Y = \overline{\mathrm{R}(T)} = \mathrm{R}(T)$. Hence, T is surjective. Since T'' is injective, $\mathrm{N}(\iota_Y T) = \mathrm{N}(T''\iota_X) = 0$, hence T is injective. Hence, T is a continuous linear bijection, hence $T^{-1} \in B(Y, X)$ by the open mapping theorem. □

Lemma 0.10. *For a linear subspace $Z \subseteq W$, define the* annihilator

$$\mathrm{Ann}(Z) = \{w' \in W' \colon \forall z \in Z.\ w'(z) = 0\}.$$

For $T \in B(X, Y)$, $\mathrm{N}(T') = \mathrm{Ann}(\mathrm{R}(T))$.

Proof. We have: $y' \in \mathrm{N}(T')$ iff $T'y' = 0$ iff $\forall x \in X.\ T'y'(x) = 0$ iff $\forall x \in X.\ y'(Tx) = 0$ iff $y' \in \mathrm{Ann}(\mathrm{R}(T))$. □

Proposition 0.11. *For $Z \subseteq W$, W a normed space, Z a closed linear subspace, and $u \notin Z$, there exists a linear functional $w' \in W'$ such that $w'|_Z = 0$, $\|w'\| = 1$ and $w'(u) = \mathrm{dist}(u, Z)$.*

Remark. For $T \in B(X)$ we have that $T^{-1} \in B(X)$ exists if and only if $0 \in \rho(T)$, i.e. invertibility of T is a "spectral question" related to $0 \in \rho(T)$ or $0 \in \sigma(T)$. The proposition says (for $T \in B(X)$) that $0 \in \sigma(T)$ if and only if $0 \in \sigma(T')$.

Proposition 0.12. *Let X be a Banach space, $T \in B(X)$, with $\|T\| < 1$. Then $(I - T)^{-1} \in B(X)$ and*

$$(I - T)^{-1} = \sum_{n=0}^{\infty} T^n$$

in $B(X)$.

Remark 0.13.

1. The series in proposition 0.12 is called the *Neumann-series* for T.

2. Of course,

$$(T - 1 \cdot I)^{-1} = (T - I)^{-1} = -(I - T)^{-1} = -\sum_{n=0}^{\infty} T^n \in B(X)$$

   if $\|T\| < 1$. Hence, if $\|T\| < 1$, $1 \in \rho(T)$.

3. The proposition is a "perturbation result" in the following sense: The identity $I \colon X \to X$ is invertible. If one adds something "not too big" , then the result is also invertible ("perturbing I a little preserves invertibility").

Corollary 0.14. *Let X and Y be Banach spaces. Then the subset of invertible operators in $B(X, Y)$ is an open set. More precisely, if $S, T \in B(X, Y)$ such that T us invertible and $\|T - S\| < \|T^{-1}\|^{-1}$, then S is invertible.*

Remark 0.15 (Functions of an operator).

1. For $T \in B(X)$, we have the obvious definition $T^0 = I$, $T^{n+1} = TT^n$, $T^{-1} = T^{-1}$ if it exists. Also $T^{-n} = (T^{-1})^n$ is the inverse of T^n.

2. More generally, if $p\colon \mathbb{C} \to \mathbb{C}$ is a polynomial, $p(z) = \sum_{i=0}^n a_i z^i$, then we can define, for $T \in B(X)$, $p(T) = \sum_{i=0}^n a_i T^i \in B(X)$.

3. We have already seen other "functions of T", namely the resolvent of T at λ, $R_\lambda(T) = (T - \lambda I)^{-1}$, and, if $\|T\| < 1$, the Neumann-series of T.

4. More generally, let $f(z) = \sum_{n=0}^\infty a_n z^n$ be a power series in $\mathbb{K}$ with radius of convergence $r > 0$. Then, by the same type of argument as for the Neumann series above, $\sum_{n=0}^\infty a_n T^n$ converges in $B(X)$, if $\|T\| < r$. Hence, one can make sense of $f(T)$ for such f.

In conclusion, we can define $f(T)$ ($T \in B(X)$) for functions f which can be expanded into a power series. We will enlarge the class of functions f for which we can give meaning to $f(T)$ considerably (possibly at the price of not making sense for all $T \in B(X)$). We will also study the relationship between the spectrum of T and that of $f(T)$.

Example 0.16.

1. For all $T \in B(X)$, we have the exponential function
$$\exp(T) = e^T = \sum_{n=0}^\infty \frac{1}{n!} T^n \in B(X).$$
For $S, T \in B(X)$, $e^{S+T} = e^S e^T$ if $ST = TS$, otherwise this may fail.

2. For $T \in B(X)$, define
$$A(s) = e^{sT} \in B(X), \quad s \in \mathbb{R}.$$
This gives a map $A\colon \mathbb{R} \to B(X)$, in fact $A \in C^\infty(\mathbb{R}, B(X))$ with $\frac{\mathrm{d}}{\mathrm{d}s} A(s) = TA(s) = A(s)T$ and $A(0) = I$, where $\frac{\mathrm{d}}{\mathrm{d}s} A(s) = \lim_{h \to 0} \frac{A(s+h) - A(s)}{h}$.

3. For $T \in B(X)$, with $\|I - T\| < 1$, define
$$\log(T) = -\sum_{n=1}^\infty \frac{1}{n} (I - T)^n \in B(X).$$

4. For $T \in B(X)$, with $\|T\| < 1$, $|s| \le 1$, define $A(s) = \log(I - sT)$. Then $\frac{\mathrm{d}}{\mathrm{d}s} A(s) = -T(I - sT)^{-1} = -(I - sT)^{-1} T$ and $\exp(A(s)) = I - sT$.

1 Spectral theory for compact operators

For a while, we shall be interested in the point spectrum $\sigma_p(T)$ for operators $T \in B(X)$, i.e. we look at the eigenvalue problem for T: Given $y \in X$ we seek all solutions $\lambda \in \mathbb{C}$, $x \in X$ to the equation $(T - \lambda I)x = y$. If $\lambda \in \rho(T)$, so that $T - \lambda I$ is invertible, then this equation has a unique solution x_0 given by $x_0 = (T - \lambda I)^{-1}y \in X$. If, on the other hand, $\lambda \in \sigma_p(T)$, then a solution x_0 to this equation (if such a thing exists) is not unique: If $x \in \mathrm{N}(T - \lambda I) \neq 0$, then also $x + x_0$ is a solution, since $(T - \lambda I)(x + x_0) = (T - \lambda I)x + (T - \lambda I)x_0 = 0 + y = y$. So, in this case, the number of solutions — the "number of degrees of freedom" — is given by $\dim \mathrm{N}(T - \lambda I)$. On the other hand, for a solution x_0 to exist, we need y to belong to $\mathrm{R}(T - \lambda I)$. This can be thought of as a "constraint" (or several) on y. If we had a scalar product and worked in a finite dimensional space, then y would have to be in the orthogonal complement to $U = \mathrm{R}(T - \lambda I)^\perp$, which means $\langle y, u_i \rangle = 0$ for a basis $\{u_i\}$ of U; hence the number of constraints on y would be $\dim U$. An important class of operators are those where both these numbers — "the number of degrees of freedom" and "the number of constraints" — are finite.

Definition 1.1. An operator $A \in B(X, Y)$ is called a *Fredholm operator* if

1. $\dim \mathrm{N}(A) < \infty$.
2. $\mathrm{R}(A)$ is closed.
3. $\operatorname{codim} \mathrm{R}(A) < \infty$.

The *index* of a Fredholm operator A is then given by $\operatorname{ind}(A) = \dim \mathrm{N}(A) - \operatorname{codim} \mathrm{R}(A)$.

Remark 1.2. That $\operatorname{codim} \mathrm{R}(A) < \infty$ means that $Y = \mathrm{R}(A) \oplus Y_0$ with $Y_0 \subseteq Y$ a linear subspace and $\dim Y_0 < \infty$. In this case $\operatorname{codim} \mathrm{R}(A) = \dim Y_0$ is independent of the choice of Y_0 such that $Y = \mathrm{R}(A) \oplus Y_0$, i.e if also $Y = \mathrm{R}(A) \oplus Y_1$ with a linear subspace $Y_1 \subseteq Y$ and $\dim Y_1 < \infty$, then $\dim Y_1 = \dim Y_0$. In a Hilbert space $Y = \mathrm{R}(A) \oplus \mathrm{R}(A)^\perp$, hence $\operatorname{codim} \mathrm{R}(A) = \dim \mathrm{R}(A)^\perp$.

A large and important class (the important) class of Fredholm operators is when $X = Y$ and A is a "compact perturbation" of the identity:

Theorem 1.3. *Let $T \in K(X)$. Then $A = I - T$ is a Fredholm operator with index* 0. *In particular,* $\dim \mathrm{N}(A) < \infty$ *(1),* $\mathrm{R}(A)$ *is closed (2),* $\mathrm{N}(A) = 0$ *implies* $\mathrm{R}(A) = X$ *(3),* $\operatorname{codim} \mathrm{R}(A) \leq \dim \mathrm{N}(A)$ *(4) and* $\dim \mathrm{N}(A) \leq \operatorname{codim} \mathrm{R}(A)$ *(5).*

Theorem 1.4 (Spectral theorem for compact operators, Riesz-Schauder)**.** *For every compact operator $T \in B(X)$, where X is a $\mathbb{C}$-Banach space, one has*

1. *$\sigma(T) \setminus \{0\}$ consists of countably many* eigenvalues*, with* 0 *as the only possible accumulation point. If $\sigma(T)$ contains infinitely many points then it follows that* $\overline{\sigma(T)} = \sigma_p(T) \cup \{0\}$.

2. *For $\lambda \in \sigma(T) \setminus \{0\}$ one has*

$$1 \le n_\lambda := \max\Big\{n \in \mathbb{N}\colon\ \mathrm{N}\big((T-\lambda I)^{n-1}\big) \ne \mathrm{N}\big((T-\lambda I)^n\big)\Big\} < \infty$$

   n_λ is the order *(or index) of λ (as an eigenvalue) and the dimension of* $\mathrm{N}(T-\lambda I)$ *is called the* multiplicity *of λ.*

3. *There is a so called* Riesz decomposition*: For $\lambda \in \sigma(T) \setminus \{0\}$,*

$$X = \mathrm{N}\big((T-\lambda I)^{n_\lambda}\big) \oplus \mathrm{R}\big((T-\lambda I)^{n_\lambda}\big)$$

   Both subspaces are closed and T-invariant and $\dim \mathrm{N}\big((T-\lambda I)^{n_\lambda}\big) < \infty$ *("generalised eigenspace").*

4. $\sigma\big(T|_{\mathrm{R}((T-\lambda I)^{n_\lambda})}\big) \setminus \{\lambda\} = \sigma(T) \setminus \{\lambda\}$.

5. *Let, for $\lambda \in \sigma(T) \setminus \{0\}$, E_λ be the projection on* $\mathrm{N}\big((T-\lambda I)^{n_\lambda}\big)$*. Then $E_\lambda E_\mu = \delta_{\lambda\mu} E_\lambda$.*

Proof.

1. Let $\lambda \notin \sigma_p(T)$, $\lambda \ne 0$. Then $\mathrm{N}(I - T/\lambda) = \mathrm{N}(\lambda I - T) = 0$. Hence, $\mathrm{R}(T-\lambda I) = \mathrm{R}(I - T/\lambda) = X$. So $(T-\lambda I)^{-1} \in B(X)$ exists, so $\lambda \in \rho(T)$. Hence, $\sigma(T) \setminus \{0\} \subseteq \sigma_p(T)$. If $\sigma(T) \setminus \{0\}$ is not finite, choose $\lambda_n \in \sigma(T) \setminus \{0\}$, $n \in \mathbb{N}$, $\lambda_n \ne \lambda_m$, and eigenvectors e_n corresponding to λ_n and define $X_n = \operatorname{span}\{e_1, \dots, e_n\}$. We claim, that the e_n are linearly independent: Assume that $\{e_1, \dots, e_{n-1}\}$ are linearly independent, but that there exist $\alpha_1, \dots, \alpha_{n-1} \in \mathbb{C}$ such that $e_n = \sum_{k=1}^{n-1} \alpha_k e_k$. Then $0 = Te_n - \lambda_n e_n = \sum_{k=1}^{n-1} \alpha_k T e_k - \sum_{k=1}^{n-1} \alpha_k \lambda_n e_n = \sum_{k=1}^{n-1} \alpha_k (\lambda_k - \lambda_n) e_k$ which implies $\alpha_k = 0$ since $\lambda_k \ne \lambda_n$ for all $1 \le k \le n-1$ by assumption. Hence, $e_n = 0$ which is impossible since e_n is an eigenvector.

   Hence, $X_{n-1} \subsetneq X_n$ is a proper closed subspace of X_n. Then, by Riesz's lemma, there exists an $x_n \in X_n$ with $\|x_n\| = 1$ and $\operatorname{dist}(x_n, X_{n-1}) \ge \frac{1}{2}$. Note, that all the X_k are T-invariant. Also, there exists $\alpha_n \in \mathbb{C}$ and $\tilde{x}_n \in X_{n-1}$ such that $x_n = \alpha_n e_n + \tilde{x}_n$. Then $Tx_n - \lambda_n x_n = \alpha_n \lambda_n e_n + T\tilde{x}_n - \alpha_n \lambda_n e_n - \lambda_n \tilde{x}_n = T\tilde{x}_n - \lambda_n \tilde{x}_n$, i.e. $(T - \lambda_n I)x_n \in X_{n-1}$. So, for $m < n$,

$$\|T(x_n/\lambda_n) - T(x_m/\lambda_m)\| = \left\| x_n + \frac{1}{\lambda_n}(Tx_n - \lambda_n x_n) - \frac{1}{\lambda_m} T x_m \right\| \ge \frac{1}{2}$$

   by the choice of the x_n. Hence, the sequence $\{T(x_n/\lambda_n)\}$ has no accumulation points. Since T is compact, it maps bounded sequences into sequences which have a convergent subsequence. So, the sequence $\{x_n/\lambda_n\}$ cannot be bounded, hence $1/|\lambda_n| = \|x_n/\lambda_n\| \to \infty$, i.e. $\lambda_n \to 0$. This proves that 0 is the only possible accumulation point of $\sigma(T) \setminus \{0\}$. In particular, $\sigma(T) \setminus B_r(0)$ is finite for all $r > 0$ (if not, it would have an accumulation point), so $\sigma(T) \setminus \{0\}$ is countable (for this we need that $\sigma(T)$ is bounded for any $T \in B(X)$).

2. Let $A = \lambda I - T$. Then $\mathrm{N}(A^{n-1}) \subseteq \mathrm{N}(A^n)$ for all n (always). Assume $\mathrm{N}(A^{n-1}) \neq \mathrm{N}(A^n)$ for all $n \geq 1$. Note that $\mathrm{N}(B)$ is always closed when $B \in B(X)$. So we can choose $x_n \in \mathrm{N}(A^n)$ such that $\|x_n\| = 1$ and $\operatorname{dist}(x_n, \mathrm{N}(A^{n-1})) \geq \frac{1}{2}$. Then, for $m < n$, $\|Tx_n - Tx_m\| = \|\lambda x_n - (Ax_n + \lambda x_m - Ax_m)\|$ and $Ax_n + \lambda x_m - Ax_m \in \mathrm{N}(A^{n-1})$, since $x_m, \lambda x_m \in \mathrm{N}(A^m)$ by construction and $\mathrm{N}(A^m) \subseteq \mathrm{N}(A^{n-1})$. Also, $A^{n-1}(Ax_m) = A(A^{n-1}(x_m)) = A0 = 0$ and $A^{n-1}(Ax_n) = A^n x_n = 0$. So, $\|Tx_n - Tx_m\| = |\lambda| \|x_n - 1/\lambda(Ax_n + \lambda x_m - Ax_m)\| \geq |\lambda|/2 > 0$. But, $\|x_n\| = 1$ for all $n \in \mathbb{N}$, so $\{x_n\}$ is a bounded sequence such that $\{Tx_n\}$ has no convergent subsequence which contradicts the compactness of T. Hence, there exists $n \in \mathbb{N}$ such that $\mathrm{N}(A^{n-1}) = \mathrm{N}(A^n)$. Then for all $m > n$, if $x \in \mathrm{N}(A^m)$, $A^{m-n}x \in \mathrm{N}(A^n) = \mathrm{N}(A^{n-1})$, i.e. $A^{m-1}x = A^{n-1}(A^{m-n}x) = 0$, so $x \in \mathrm{N}(A^{m-1})$. Continuing in this fashion, one sees $\mathrm{N}(A^n) = \mathrm{N}(A^m)$ and $n_\lambda < \infty$. Also $n_\lambda \geq 1$ since $\mathrm{N}(T - \lambda I) \neq 0$.

3. Write again $A = \lambda I - T = \lambda(I - T/\lambda)$. This A is Fredholm. We claim that $\mathrm{N}(A^{n_\lambda}) \oplus \mathrm{R}(A^{n_\lambda}) \subseteq X$, i.e. that $\mathrm{N}(A^{n_\lambda}) \cap \mathrm{R}(A^{n_\lambda}) = 0$. For this let $x \in \mathrm{N}(A^{n_\lambda})$, $x = A^{n_\lambda} y$, for some y. Then $A^{2n_\lambda} y = A^{n_\lambda} A^{n_\lambda} y = A^{n_\lambda} x = 0$, so $y \in \mathrm{N}(A^{2n_\lambda}) = \mathrm{N}(A^{n_\lambda})$ and $x = A^{n_\lambda} y = 0$. Note, that
$$A^{n_\lambda} = (\lambda I - T)^{n_\lambda} = \lambda^{n_\lambda} I + \sum_{k=1}^{n_\lambda} \binom{n_\lambda}{k} \lambda^{n_\lambda - k} (-T)^k$$
is Fredholm by 1.3 since the second summand is compact. Hence, $\mathrm{R}(A^{n_\lambda})$ is closed and $\operatorname{codim} \mathrm{R}(A^{n_\lambda}) \leq \dim \mathrm{N}(A^{n_\lambda}) < \infty$. Hence, we have $X = \mathrm{N}(A^{n_\lambda}) \oplus \mathrm{R}(A^{n_\lambda})$. Clearly, both these subspaces are T-invariant.

4. Write $T_\lambda = T|_{\mathrm{R}\left((T-\lambda I)^{n_\lambda}\right)}$ and take $\mu \in \mathbb{C} \setminus \{\lambda\}$. Then $\mu I - T$ is injective on $\mathrm{N}(A^{n_\lambda})$, because if $x \in \mathrm{N}(A^{n_\lambda}) \cap \mathrm{N}(\mu I - T)$, i.e. $(\lambda - \mu)x = Ax$ and $A^{n_\lambda x} x = 0$, then $(\mu - \lambda) A^{n_\lambda - 1} x = A^{n_\lambda - 1}(\lambda - \mu)x = A^{n_\lambda} x = 0$ whence $A^{n_\lambda} x = 0$. Hence by induction $x = 0$. Moreover $\mathrm{N}(A^{n_\lambda})$ is finite dimensional, thus $(\mu - T)\colon \mathrm{N}(A^{n_\lambda}) \to \mathrm{N}(A^{n_\lambda})$ is a bijection. Hence, wether or not $\mu - T$ is invertible on X is equivalent to wether or not $\mu - T$ is invertible on $\mathrm{R}(A^{n_\lambda})$ in view of the decomposition $X = \mathrm{N}(A^{n_\lambda}) \oplus \mathrm{R}(A^{n_\lambda})$. Therefore $\mu \in \rho(T)$ iff $\mu \in \rho(T_\lambda)$. Equivalently $\sigma(T) \setminus \{\lambda\} = \sigma(T_\lambda) \setminus \{\lambda\}$.

5. Let $\lambda, \mu \in \sigma(T) \setminus \{0\}$, $\lambda \neq \mu$. Let $A_\lambda = \lambda - T$, $A_\mu = \mu - T$. Then, taking any $x \in \mathrm{N}(A_\mu^{n_\mu})$, write $x = z + y$ with $z \in \mathrm{N}(A_\lambda^{n_\lambda})$ and $y \in \mathrm{R}(A_\lambda^{n_\lambda})$. Then $0 = A_\mu^{n_\mu} x = A_\mu^{n_\mu} z + A_\mu^{n_\mu} y$. Both $\mathrm{N}(A_\lambda^{n_\lambda})$ and $\mathrm{R}(A_\lambda^{n_\lambda})$ are T-invariant, hence $A_\mu^{n_\mu} z \in \mathrm{N}(A_\lambda^{n_\lambda})$ and $A_\mu^{n_\mu} y \in \mathrm{R}(A_\lambda^{n_\lambda})$. Therefore $A_\mu^{n_\mu} z = 0$. Since $A_\mu\colon \mathrm{N}(A_\lambda^{n_\lambda}) \to \mathrm{N}(A_\lambda^{n_\lambda})$ is a bijection, $z = 0$, so $z = y \in \mathrm{R}(A_\lambda^{n_\lambda})$ and therefore $\mathrm{N}(A_\mu^{n_\mu}) \subseteq \mathrm{R}(A_\lambda^{n_\lambda})$. Therefore $\mathrm{R}(E_\mu) \subseteq \mathrm{N}(E_\lambda)$. □

Theorem 1.5. *Let X be a normed space, $E \subseteq X$ an n-dimensional subspace with basis $\{e_1, \dots, e_n\}$ and $Y \subseteq X$ a closed subspace such that $Y \cap E = 0$. Then there exist $e_1', \dots, e_n' \in X'$ such that $e_i'|_Y = 0$ and $e_i'(e_j) = \delta_{ij}$. Moreover there exists a continuous projection P onto E with $Y \subseteq \mathrm{N}(P)$.*

Lemma 1.6. *For all finite dimensional subspaces $E \subseteq X$, if $Y \subseteq X$ is a closed subspace with $Y \cap E = 0$, we have $Y \oplus E \subseteq X$ is also a closed subspace.*

Lemma 1.7. *If X is a normed space, $Y \subseteq X$ a closed subspace, $x_0 \notin Y$, then there exists $x' \in X'$ such that $x'|_Y = 0$, $\|x'\| = 1$ and $x'(x_0) = \operatorname{dist}(x_0, Y)$.*

Proof of theorem 1.3. We will only prove (3). Assume $\mathrm{N}(A) = 0$ and for contradiction assume there is $x \in X \setminus \mathrm{R}(A)$. Then $A^n x \in \mathrm{R}(A^n) \setminus \mathrm{R}(A^{n+1})$ for all $n \geq 0$, for otherwise if $A^n x = A^{n+1} y$ for some $y \in X$, so $A^n(x - Ay) = 0$. Since $\mathrm{N}(A) = 0$, this would imply $x = Ay \in \mathrm{R}(A)$. Moreover, $\mathrm{R}(A^{n+1})$ is closed (by part 2). By Riesz' lemma there exist $x_n \in \mathrm{R}(A^n)$, $\|x_n\| = 1$, such that $\operatorname{dist}(x_n, \mathrm{R}(A^{n+1})) = \frac{1}{2}$. Then, for $m > n$, $\|Tx_n - Tx_m\| = \|x_n - (Ax_n + x_m - Ax_m)\| \geq \frac{1}{2}$, whence no subsequence of $\{Tx_n\}$ can be Cauchy. □

Proposition 1.8 (Fredholm alternative)**.** *For $T \in K(X)$ and $\lambda \neq 0$, either the equation $Tx - \lambda x = y$ has a unique solution, or the equation $Tx - \lambda x = 0$ has nontrivial solutions.*

Proposition 1.9. *Let $T \in K(X)$, $\lambda \in \sigma(T) \setminus \{0\}$. Then the resolvent map $\rho(T) \ni \mu \mapsto R_\mu(T) = (T - \mu)^{-1}$ has an isolated pole at $\mu = \lambda$ of order n_λ (see. 1.4 (2)), that is, the map $\mu \mapsto (\mu - \lambda)^{n_\lambda} R_\mu(T)$ can be continued at the point λ to an analytic map and $(\mu - \lambda)^{n_\lambda} R_\mu(T) \neq 0$.*

Proof. We have $X = \mathrm{N}((\lambda - T)^{n_\lambda}) \oplus \mathrm{R}((\lambda - T)^{n_\lambda})$. Write E_λ for the projection on $\mathrm{N}((\lambda - T)^{n_\lambda})$ and $T_0 = T|_{\mathrm{R}(E_\lambda)}$, $T_1 = T|_{\mathrm{N}(E_\lambda)}$. Since λ is an isolated point in $\sigma(T)$, there exists $r > 0$ such that $B_{\lambda,r} \setminus \{\lambda\} \subseteq \rho(T)$, whence $B_{\lambda,r} \setminus \{\lambda\} \subseteq \rho(T_0)$ and $B_{\lambda,r} \setminus \{\lambda\} \subseteq \rho(T_1)$ because of part (4) of theorem 1.4. Moreover, for every $z \in \mathbb{C} \setminus \{0\}$, $R_{\lambda+z}(T) = R_{\lambda+z}(T_0)E_\lambda + R_{\lambda+z}(T_1)(I - E_\lambda)$. Also, $z \mapsto R_{\lambda+z}(T_1)$ is analytic. Setting

$$S(z) = \sum_{k=1}^{n_\lambda} \frac{1}{z^k}(T_0 - \lambda)^{k-1}$$

one sees that

$$S(z)((\lambda + z)I - T_0) = \sum_{k=1}^{n_\lambda} z^{1-k}(T_0 - \lambda)^{k-1} - \sum_{k=1}^{n_\lambda} z^{-k}(T_0 - \lambda)^k = I - z^{n_\lambda}(T_0 - \lambda)^{n_\lambda} = I.$$

Analogously $((\lambda + z)I - T_0)S(z) = I$, whence $S(z) = R_{\lambda+z}(T_0)$. Now it's clear that $R_{\lambda+z}(T_0)$ has a pole at $z = 0$ of order n_λ. □

Remark. So far the decomposition in theorem 1.4 (3) looks very impressive since it gives us a very fine decomposition of X into a direct sum of generalised eigenspaces. But the theorem cannot even guarantee the existence of a closed invariant subspace, let alone a direct sum decomposition.

Proposition 1.10 (finite-dimensional case)**.** *Let $\dim X < \infty$, $T \colon X \to X$ linear. Then there exist distinct $\lambda_1, \dots, \lambda_m$, $m \leq \dim X$, such that $\sigma(T) = \sigma_p(T) = \{\lambda_1, \dots, \lambda_m\}$. Each λ_j has multiplicity n_{λ_j} with the properties 1.4 (2)–(5), that is*

$$X = \bigoplus_{j=1}^{m} \mathrm{N}((T - \lambda_j)^{n_{\lambda_j}}).$$

Sketch of a proof. Equip X with any norm. Then X is Banach and T, I and $T-\lambda$ are compact. Set $T_\mu = T-\mu$ for any $\mu\in\mathbb{C}$ and apply theorem 1.4 to T_0 and T_1. $\square$

Proposition 1.11 (Jordan normal form)**.** *Let $T\in K(X)$, $\lambda\in\sigma_p(T)$. For $A=T-\lambda$ one has*

1. *For $n=1,\dots,n_\lambda$, there exist subspaces $E_n\subseteq \mathrm{N}(A^n)$ such that $E_n\cap\mathrm{N}(A^{n-1})=0$ and $\mathrm{N}(A^{n_\lambda})=\bigoplus_{k=1}^{n_\lambda} N_k$ where $N_k=\bigoplus_{l=0}^{k-1}A^l(E_k)$.*
2. *The subspaces N_k are T-invariant and $d_k=\dim A^l(E_k)$ is independent of $l\in\{0,\dots,k-1\}$.*
3. *If $\{e_{k,j}\colon j=1,\dots,d_k\}$ is a basis for E_k, then $\{A^l e_{k,j}\colon 0\le l<k\le n_j, 1\le j\le d_k\}$ is a basis for $\mathrm{N}(A^{n_\lambda})$ and with*

$$x=\sum_{k,j,l}\alpha_{k,j,l}A^l e_{k,j}\quad\text{and}\quad y=\sum_{k,j,l}\beta_{k,j,l}A^l e_{k,j}$$

the equation $Tx=y$ is equivalent to

$$\begin{pmatrix}\beta_{k,j,0}\\ \vdots\\ \beta_{k,j,k-1}\end{pmatrix}=\begin{pmatrix}\lambda & 1 & & \\ & \ddots & \ddots & \\ & & \lambda & 1\\ & & & \lambda\end{pmatrix}\begin{pmatrix}\alpha_{k,j,0}\\ \vdots\\ \alpha_{k,j,k-1}\end{pmatrix}$$

Sketch of a proof. If E is a subspace with $\mathrm{N}(A^{n-1})\oplus E\subseteq\mathrm{N}(A^n)$, then $\mathrm{N}(A^{n-1-l})\oplus A^l(E)\subseteq\mathrm{N}(A^{n-l})$ for $l=0,\dots,n$ and A^l is injective on E. Choose E_n for $n=n_\lambda,\dots,1$ such that $\mathrm{N}(A^n)=\mathrm{N}(A^{n-1})\oplus\bigoplus_{l=0}^{n_\lambda-n}A^l(E_{n+l})$ from which the claim follows. $\square$

Theorem 1.12 (Schauder)**.** *We have $T\in K(X,Y)$ if and only if $T'\in K(Y',X')$.*

Proof. First assume $T\in K(X,Y)$, i.e. $\overline{T(B_1(0))}$ is compact in Y. For $y'\in Y'$ we have $\|T'y'\|_{X'}=\sup_{\|x\|\le1}|\langle T'y',x\rangle|=\sup_{\|x\|\le1}|\langle y',Tx\rangle|=\sup_{y\in\overline{TB_1(0)}}|\langle y',y\rangle|=\|y'\|_{C(\overline{TB_1(0)})}$, form which we learn that $B_1'(0)$ is isometric to the set $\mathscr{A}=\{y'=\tilde y'|_{\overline{TB_1(0)}}\colon \tilde y'\in Y',\|\tilde y'\|\le1\}$. Therefore the precompactness of $T'B_1(0)$ is equivalent to precompactness of $\mathscr{A}$ in $C(\overline{TB_1(0)})$ equipped with $\|\cdot\|_\infty$. The latter follows from Ascoli-Arzelà: $\mathscr{A}$ is uniformly bounded, because $\|y'\|_\infty=\|T'y'\|\le\|T'\|\|y'\|\le\|T\|$. $\mathscr{A}$ is an equicontinuous family since

$$|\langle y',y_1\rangle-\langle y',y_2\rangle|=|\langle y',y_1-y_2\rangle|\le\|y'\|\|y_1-y_2\|\le\|y_1-y_2\|$$

for every $y_1,y_2\in\overline{TB_1(0)}$.

Conversely, if $T'\in K(Y',X')$, by the above $T''\in K(X'',Y'')$. Then exploit the canonical embedding $X\hookrightarrow X''$: $T(B_1(0))\subseteq\overline{T''(B_1(0))}$ and the latter is compact. $\square$

Remark. If X is a Hilbert space, the proof of theorem 1.12 is much shorter since one has approximation of compact operators by finite rank operators.

The last step of this chapter on the spectral theory of compact operators is to investigate the special case of compact and *normal* operators. We shall see that the Riesz-Schauder decomposition results on this case in an orthonormal decomposition of eigenspaces.

Definition 1.13. For a Hilbert space H, $T \in B(H)$ is called *normal* if $T^*T = TT^*$.

Lemma 1.14. *For $T \in B(H)$, $[T, T^*] = 0$, i.e. T is normal, if and only if $\|Tx\| = \|T^*x\|$ for all $x \in H$.*

Proof. We have $\|Tx\|^2 - \|T^*x\|^2 = \langle Tx, Tx\rangle - \langle T^*x, T^*x\rangle = \langle Tx, Tx\rangle - \langle TT^*x, x\rangle = \langle Tx, Tx\rangle - \langle T^*Tx, x\rangle = \langle Tx, Tx\rangle - \langle Tx, Tx\rangle = 0$, if T is normal. □

Theorem 1.15. *Let $T \in B(X)$, where $X \neq 0$ is a $\mathbb{C}$-Banach space. Then $\sigma(T) \subseteq \mathbb{C}$ is compact and nonempty, and*

$$\sup_{\lambda\in\sigma(T)} |\lambda| = \lim_{m\to\infty} \|T^m\|^{1/m} \leq \|T\|.$$

We call $r(T) = \sup_{\lambda\in\sigma(T)} |\lambda|$ the spectral radius *of T.*

Proof. By Proposition 0.3, $\rho(T) \subseteq \mathbb{C}$ is open, hence $\sigma(T)$ is closed. By Proposition 0.12, $I - T/\lambda$ is invertible, if $\|T/\lambda\| < 1$, that is, if $|\lambda| > \|T\|$, and in this case

$$R_\lambda(T) = (T - \lambda I)^{-1} = -\frac{1}{\lambda}\left(I - \frac{T}{\lambda}\right)^{-1} = -\sum_{k=0}^{\infty} \frac{1}{\lambda^{n+1}} T^n.$$

Hence, $r(T) \leq \|T\|$. So, $\sigma(T)$ is bounded, hence compact by Heine-Borel. Note that $T^m - \lambda^m I = (T - \lambda I)p_m(T) = p_m(T)(T - \lambda I)$ for

$$p_m(T) = \sum_{i=0}^{m-1} \lambda^{m-1-i} T^i.$$

Hence, $\lambda \in \sigma(T)$ implies $\lambda^m \in \sigma(T^m)$. Hence, $|\lambda^m| \leq \|T^m\|$, i.e. $|\lambda| \leq \|T^m\|^{1/m}$ for all m. So, $|\lambda| \leq \liminf_{m\to\infty} \|T^m\|^{1/m}$ for all $\lambda \in \sigma(T)$. So $r(T) \leq \liminf_{m\to\infty} \|T^m\|^{1/m}$. It remains to prove that $r(T) \geq \limsup_{m\to\infty} \|T^m\|^{1/m}$.

By Proposition 0.3 $\rho(T) \subseteq \mathbb{C}$ is open and the resolvent map $\lambda \mapsto R_\lambda(T) \in B(X)$ is a complex analytic map, i.e. for all $\lambda_0 \in \mathbb{C} \setminus \overline{B_{r(t)}(0)}$ there exist $c_n(\lambda_0) \in B(X)$, $n \geq 0$ and $\varepsilon(\lambda_0) > 0$ such that

$$R_\lambda(T) = \sum_{n=0}^{\infty} c_j(\lambda_0)(\lambda - \lambda_0)^n$$

for all $\lambda \in B_{\varepsilon(\lambda_0)}(\lambda_0)$. We will need some Complex Analysis:

Theorem (Cauchy's integral formula)**.** *Let $\Omega \subseteq \mathbb{C}$ be open and simply connected and assume $f\colon \Omega \to \mathbb{C}$ is analytic. Let $\gamma\colon [a, b] \to \Omega$ be any closed path. Then*

$$\oint_\gamma f(z)\,\mathrm{d}z = 0.$$

Note that "simply connected" informally means "no holes" and a "path" means a rectifiable curve, in particular it is enough for γ to be C^1.

The same holds when $f\colon \Omega \to Y$ for a complex Banach space Y. Only the linear structure and the completeness of Y where needed for the proof, once one can make sense of the Riemann integral appearing in Cauchy's formula in this setting.

Now $\lambda \mapsto R_\lambda(T)$ is complex analytic and by Cauchy's formula

$$\int_{\gamma_1} \lambda^j R_\lambda(T)\,\mathrm{d}\lambda = \int_{\gamma_2} \lambda^j R_\lambda(T)\,\mathrm{d}\lambda$$

where γ_i are circles of the same orientation around the origin of radius strictly bigger that $r(T)$. Hence, for any $j \geq 0$, $s > R(T)$, the integral

$$\frac{1}{2\pi i}\int_{\partial B_s(0)} \lambda^j R_\lambda(T)\,\mathrm{d}\lambda$$

is independent of s. Recall that $R_\lambda(T) = -\sum_{n=0}^\infty \lambda^{-n-1}T^n$ for $|\lambda| > r(T)$. Hence,

$$\begin{aligned}\frac{1}{2\pi i}\int_{\partial B_s(0)} \lambda^j R_\lambda(T)\,\mathrm{d}\lambda &= -\frac{1}{2\pi i}\int_{\partial B_s(0)} \sum_{n=0}^\infty \lambda^{j-n-1}T^n\,\mathrm{d}\lambda = \\ &= -\frac{1}{2\pi}\sum_{n=0}^\infty \left(s^{j-n}\int_0^{2\pi} e^{i\theta(j-n)}\,\mathrm{d}\theta\right) T^n = -T^j.\end{aligned}$$

Hence, for $j \geq 0$ and $s > r(T)$

$$\begin{aligned}\|T^j\| &= \frac{1}{2\pi}\left\|\int_{\partial B_s(0)} \lambda^j R_\lambda(T)\,\mathrm{d}\lambda\right\| = \frac{1}{2\pi}\left\|\int_0^{2\pi}\left(se^{i\theta}\right)^j R_{se^{i\theta}}(T)\left(ise^{i\theta}\right)\mathrm{d}\theta\right\| \leq \\ &\leq \frac{1}{2\pi}s^{j+1}\int_0^{2\pi}\|R_{se^{i\theta}}(T)\|\,\mathrm{d}\theta \leq s^{j+1}\sup_{\lambda\in\partial B_s(0)}\|R_\lambda(T)\|\end{aligned}$$

and hence $\|T^j\|^{1/j} \leq s\left(s\cdot\sup_{\lambda\in\partial B_s(0)}\|R_\lambda(T)\|\right)^{1/j} \xrightarrow{j\to\infty} s$. So, $\limsup_{m\to\infty}\|T^m\|^{1/m} \leq s$ for all $s > r(T)$. Hence, $\limsup_{m\to\infty}\|T^m\|^{1/m} \leq r(T)$.

Assume for contradicition that $\sigma(T) = \emptyset$ (then $r(T) := 0$). For $j = 0$ and any $s > 0$, $\|I\| = \|T^0\| \leq s\cdot\max_{\lambda\in\partial B_s(0)}\|R_\lambda(T)\| \leq s\cdot\max_{|\lambda|\leq s}\|R_\lambda(T)\| \xrightarrow{s\to 0} 0$, which is only possible if $X = 0$. $\square$

Proposition 1.16. *Let $H \neq 0$ be a complex Hilbert space, and $T \in B(H)$. If T is normal, $r(T) = \|T\|$.*

Proof. Let $T \neq 0$. Recall that $\lim_{m\to\infty}\|T^m\|^{1/m} \leq \|T\|$, hence it is enough to prove that $\|T^m\| \geq \|T\|^m$ for $m \geq 0$. This is clear for $m = 0, 1$.

For $m \geq 1$, and $x \in H$, we have

$$\begin{aligned}\|T^m x\|^2 &= \langle T^m x, T^m x\rangle = \left\langle T^*T^m x, T^{m-1}x\right\rangle \leq \|T^*T^m x\|\|T^{m-1}x\| = \\ &= \|T^{m+1}x\|\|T^{m-1}x\| \leq \|T^{m+1}\|\|x\|\|T\|^{m-1}\|x\| \leq \|T^{m+1}\|\|T\|^{m-1}\|x\|^2\end{aligned}$$

Hence, $\|T^m\|^2 \leq \|T^{m+1}\|\|T\|^{m-1}$. So $\|T^{m+1}\| \geq \|T^m\|^2/\|T\|^{m-1} \geq \|T\|^{2m-(m-1)} = \|T\|^{m+1}$ if $\|T^m\| \geq \|T\|^m$. Hence, by induction, the proposition follows. $\square$

Definition 1.17. Let $T \in B(H)$ be selfadjoint, i.e. $T^* = T$. Then T is called *positive semi-definite* if $\langle x, Tx\rangle \geq 0$ for all $x \in H$. We will write $T \geq 0$. We will also write $A \geq B$ for operators A, B iff $A - B \geq 0$.

Remark. By problem 12 $\langle x, Tx\rangle \in \mathbb{R}$ if T is selfadjoint.

Proposition 1.18. *Let H be a $\mathbb{C}$-Hilbert space, $T \in B(H)$.*

1. *If $T = T^*$, then $\sigma(T) \subseteq [-\|T\|, \|T\|] \subseteq \mathbb{R}$. If also $T \in K(H)$, then $-\|T\|$ or $\|T\|$ is an eigenvalue of T.*

2. *If $T = T^*$ and $T \geq 0$, then $\sigma(T) \subseteq [0, \|T\|] \subseteq \mathbb{R}$. If also $T \in K(H)$, then $\|T\|$ is an eigenvalue of T.*

Proof. By problem 12, for $T = T^*$, $\sigma(T) \subseteq \left[\inf_{x\in H, \|x\|\leq 1} \langle x, Tx\rangle, \sup_{x\in H, \|x\|\leq 1} \langle x, Tx\rangle\right]$. Also, by 1.15, $\sigma(T) \subseteq B_{\|T\|}(0) \subseteq \mathbb{C}$. So $\sigma(T) \subseteq [-\|T\|, \|T\|]$. In fact, since s.a. implies normal, $\max_{\lambda\in\sigma(T)} |\lambda| = \|T\|$. If T is compact, then $\sigma(T) \smallsetminus \{0\}$ consists of isolated eigenvalues by the spectral theorem for compact operators on a Banach space, so $\|T\|$ or $-\|T\|$ is an eigenvalue.

For positive T, again using problem 12, $\sigma(T) \subseteq \left[0, \sup_{x\in H, \|x\|\leq 1} \langle x, Tx\rangle\right]$ and as before $\sigma(T) \subseteq B_{\|T\|}(0)$. So $\sigma(T) \subseteq [0, \|T\|] \subseteq \mathbb{R}$. If also $T \in K(H)$, then, as above $\|T\|$ is an eigenvalue. □

Remark. In particular, if $T \in K(H)$, $T = T^*$, $T \neq 0$, and $T \geq 0$, then for the largest eigenvalue λ_0 of T one has $\lambda_0 = \max_{\lambda\in\sigma(T)} \lambda = \|T\| = \sup_{x\in H, \|x\|\leq 1} \langle x, Tx\rangle = \lambda_0$. I.e. maximising $\langle x, Tx\rangle$ under the constraint $\|x\| \leq 1$ (or $\|x\| = 1$) gives the largest eigenvalue of T. In particular $\lambda_0 \geq \langle x, Tx\rangle$ for all $x \in H$ with $\|x\| = 1$. One can use this to compute a lower bound on λ_0 by choosing some x. Also, one can repeat this: for $\lambda_1 \leq \lambda_0$ the next eigenvalue, the same method works by restricting T so $(\text{span}\{x_0\})^\perp =: \widetilde{X}$, since $\widetilde{X}$ is T-invariant: for $\tilde{x} \in \widetilde{X}$ one has $\langle T\tilde{x}, x_0\rangle = \langle \tilde{x}, Tx_0\rangle = \lambda_0\langle \tilde{x}, x_0\rangle = 0$, hence $T\tilde{x} \in \widetilde{X}$. This can be used to compute the λ_j and their eigenvectors.

Example 1.19. Let H be a complex Hilbert space, and choose an ONS $\{e_j\}_{j\in N}$ for some $N \subseteq \mathbb{N}$ and choose a sequence $\{\lambda_j\}_{j\in N} \subseteq \mathbb{C}$ with $|\lambda_k| \leq r < \infty$ for all $k \in N$ and some $r > 0$. Then $Tx = \sum_{j\in N} \lambda_j \langle e_j, x\rangle e_j$ defines a bounded and normal operator. Also T is compact iff $\lambda_j \to 0$ as $j \to \infty$ (if N is infinite, otherwise T has finite rank, hence is compact). The following theorem shows that every compact, normal operator has this form.

Theorem 1.20. *Let H be a complex Hilbert space, and let $T \in K(H)$, $T \neq 0$, be normal. Then*

1. *There exists an ONS $\{e_j\}_{j\in N}$ with $N \subseteq \mathbb{N}$ and $\{\lambda_j\}_{j\in N} \subseteq \mathbb{C} \smallsetminus \{0\}$ such that $Te_k = \lambda_k e_k$, $k \in N$, and $\sigma(T) \smallsetminus \{0\} = \{\lambda_j\}_{j\in N}$. If N is infinite, then $\lambda_k \to 0$ as $k \to \infty$.*

2. *For all k, $n_{\lambda_k} = 1$ (the order of λ_k, i.e. the maximal n such that $\mathrm{N}((T-\lambda_k I)^{n-1}) \neq \mathrm{N}((T - \lambda_k I)^n)$.*

3. *One has the orthogonal decomposition* $H = \mathrm{N}(T) \oplus \overline{\operatorname{span}\{e_n : n \in N\}}$.

4. $Tx = \sum_{k \in N} \lambda_k \langle e_k, x\rangle e_k$ *for all* $x \in H$.

Proof. Apply first the spectral theorem for compact operators (1.4) on T to get that $\sigma(T) \setminus \{0\}$ consists of eigenvalues $\tilde{\lambda}_k$, $k \in \widetilde{N} \subseteq \mathbb{N}$, with $\tilde{\lambda}_k \to 0$ as $k \to \infty$ (if $\widetilde{N}$ is infinite). In this enumeration, the $\tilde{\lambda}_k$ are distinct. Moreover, the space $N_k = \mathrm{N}(T - \tilde{\lambda}_k I)$ is finite dimensional. Let $N_0 = \mathrm{N}(T)$, and $\tilde{\lambda}_0 = 0$. Note that $\mathrm{N}(T - \tilde{\lambda}_k I) = \mathrm{N}(T^* - \overline{\tilde{\lambda}_k} I)$ for $k \in \widetilde{N} \cup \{0\}$. We claim that $N_k \perp N_l$ for $k, l \in \widetilde{N} \cup \{0\}$, $k \neq l$. For $x_k \in N_k$ and $x_l \in N_l$, $k \neq l$, one has

$$\tilde{\lambda}_l \langle x_k, x_l\rangle = \left\langle x_k, \tilde{\lambda}_l x_l \right\rangle = \langle x_k, T x_l\rangle = \langle T^* x_k, x_l\rangle = \left\langle \overline{\tilde{\lambda}_k} x_k, x_l \right\rangle = \tilde{\lambda}_k \langle x_k, x_l\rangle$$

and hence, since $\tilde{\lambda}_l \neq \tilde{\lambda}_k$, $\langle x_k, x_l\rangle = 0$ as claimed. Next, we claim $H = \overline{\bigoplus_{k \in \widetilde{N} \cup \{0\}} N_k}$. Let $y \in Y := \left(\overline{\bigoplus_{k \in \widetilde{N} \cup \{0\}} N_k}\right)^{\perp}$. Since, $N_k = \mathrm{N}(T^* - \overline{\tilde{\lambda}_k} I)$, $k \in \widetilde{N} \cup \{0\}$, we have, for $x \in N_k$, $k \in \widetilde{N} \cup \{0\}$:

$$\langle x, Ty\rangle = \langle T^* x, y\rangle = \tilde{\lambda}_k \langle x, y\rangle = 0,$$

hence $Ty \in Y$. Hence, Y is a T-invariant, closed subspace of H. We now look at $T_0 = T|_Y : Y \to Y$. Then $T_0 \in K(Y)$ is normal. In case $Y \neq 0$, $\sup_{\lambda \in \sigma(T_0)} |\lambda| = \|T_0\|$ and its eigenvalues $\mu_k \to 0$ as $k \to \infty$. Then there exists a $\mu_0 \in \sigma(T_0)$ such that $|\mu_0| = \|T_0\|$. If $\|T_0\| \neq 0$, then $\mu_0 \neq 0$ is an eigenvalue for T_0 by the spectral theorem, hence μ_0 is an eigenvalue for T with corresponding eigenvector $u \in Y \subseteq H$. Hence, $\mu_0 \in \sigma(T)$ and $u \in N_k$ for some k, which is a contradiction, since $Y \perp N_k$. If $\|T_0\| = 0$, then $T_0 = 0$, so $Y \subseteq \mathrm{N}(T)$ which is equally impossible if $Y \neq 0$. Hence, we have $Y = 0$.

Let E_k, $k \in \widetilde{N} \cup \{0\}$, be the orthogonal projection on N_k. Then $x = \sum_{k \in \widetilde{N} \cup \{0\}} E_k x$, so

$$Tx = \sum_{k \in \widetilde{N} \cup \{0\}} T E_k x = \sum_{k \in \widetilde{N} \cup \{0\}} \tilde{\lambda}_k E_k x = \sum_{k \in \widetilde{N}} \tilde{\lambda}_k E_k.$$

Now the representation of T follows by choosing orthonormal bases $\{e_{k1}, \ldots, e_{kd_k}\}$, $d_k = \dim N_k$, of N_k. Then $E_k x = \sum_{j=1}^{d_k} \langle e_{kj}, x\rangle e_{kj}$ and

$$Tx = \sum_{k \in N} \lambda_k \langle e_k, x\rangle e_k$$

by relabelling the $\tilde{\lambda}_k$ and e_{kj}.

From this representation of T it follows that $\mathrm{N}((T - \lambda_k I)^2) = N_k = \mathrm{N}(T - \lambda_k I)$, for let $x \in \mathrm{N}((T - \lambda_k I)^2)$, i.e. $0 = (T - \tilde{\lambda}_k I)^2 x = \sum_{j \in \widetilde{N}} (\tilde{\lambda}_k - \tilde{\lambda}_j)^2 E_j x$. Since $H = \overline{\bigoplus_{k \in \widetilde{N} \cup \{0\}} N_k}$, this implies $E_j x = 0$ for all $j \neq k$. Hence $x = E_k x$, so $x \in N_k$. $\square$

Let (as before) E_k be the orthogonal projection on $N_k = \mathrm{N}(T - \tilde{\lambda}_k I)$, that is $E_k x = \sum_{j=1}^{d_k} \langle e_{kj}, x\rangle e_{kj}$ for an orthonormal basis $\{e_{k1}, \ldots, e_{kd_k}\}$ of N_k. The spectral theorem says that $Tx = \sum_{k \in N} \tilde{\lambda}_k E_k x$, i.e. the sum $T = \sum_{k \in N} \tilde{\lambda}_k E_k$ converges pointwise.

Corollary 1.21 (Spectral theorem for compact operators; Projection version)**.** *Let the assumptions be as in theorem 1.20. Then, in the notation of above, $\sum_{k\in N} \tilde{\lambda}_k E_k$ converges in* norm *to T.*

Proof. We have

$$\left\| T - \sum_{k=1}^{M} \tilde{\lambda}_k E_k \right\| = \left\| \sum_{k=M+1}^{\infty} \tilde{\lambda}_k E_k \right\| = \sup_{\lambda\in\sigma(\sum \tilde{\lambda}_k E_k)} |\lambda| = |\tilde{\lambda}_{M+1}| \xrightarrow{M\to\infty} 0$$

since $\sum \tilde{\lambda}_k E_k$ is normal. □

Theorem 1.22. *Let $T \in K(H)$, $T^* = T$, $T \geq 0$. Then there exists a unique, positive, self-adjoint operator $S \in K(H)$ such that $S^2 = T$. We write $T^{1/2} = S$.*

Proof. Write $Tx = \sum_{k\in N} \lambda_k \langle e_k, x\rangle e_k$ for all $x \in H$. Since $T \geq 0$, we have $\lambda_k \geq 0$. Let $Sx = \sum_{k\in N} \sqrt{\lambda_k}\langle e_k, x\rangle e_k$ for $x \in H$. This defines a compact normal operator, since $\sqrt{\lambda_k} \to 0$ as $k \to \infty$. In fact, S is self-adjoint and $S \geq 0$. Also, one computes

$$S^2 x = SSx = S \sum_{k\in N} \sqrt{\lambda_k}\langle e_k, x\rangle e_k = \sum_{k\in N} \lambda_k \langle e_k, x\rangle e_k = Tx.$$

Assume $R \in K(H)$ with $R = R^*$, $R \geq 0$ and $R^2 = T$. Using the spectral theorem on R, we have $R = \sum_k \nu_k F_k$ with ν_k the eigenvalues of R and corresponding orthogonal projections F_k. Then $T = R^2 = \big(\sum_k \nu_k F_k\big)\big(\sum_m \nu_m F_m\big) = \sum_k \nu_k^2 F_k$ and the ν_k^2 are the eigenvalues of T with corresponding orthogonal projections F_k. Hence, $R = S$, since $\nu_k \geq 0$ by $R \geq 0$. □

Let $T\colon H_1 \to H_2$ be a compact operator between two Hilbert spaces. Then $T^*T\colon H_1 \to H_1$ is positive and self-adjoint. Its unique square root is denoted $|T| := (T^*T)^{1/2}$.

Theorem 1.23 (Polar decomposition)**.** *For $T \in K(H_1, H_2)$ there exists an operator $U \in B(H_1, H_2)$ with $T = U|T|$, such that $U|_{\mathrm{N}(U)^\perp}$ is an isometry between $\mathrm{N}(U)^\perp$ and $\mathrm{R}(U)$. U is uniquely determined by the demand that $\mathrm{N}(U) = \mathrm{N}(T)$.*

Proof. $|T|$ is self-adjoint, so $\||T|x\|^2 = \langle |T|x, |T|x\rangle = \langle x, T^*Tx\rangle = \langle Tx, Tx\rangle = \|Tx\|^2$. Hence, $U(|T|x) = Tx$ defines an isometry from $\mathrm{R}(|T|)$ to $\mathrm{R}(T)$ which extends uniquely to a bounded operator $U\colon \overline{\mathrm{R}(|T|)} \to \overline{\mathrm{R}(T)}$. Let $Ux = 0$ for $x \in \mathrm{R}(|T|)^\perp = \mathrm{N}(|T|)$. This proves the existence of U with the stated properties. The uniqueness follows from $\mathrm{N}(|T|) = \mathrm{N}(T)$. □

Remark.

1. An operator with the properties of U is called a *partial isometry.*
2. $T = U|T|$ reminds of $\lambda = e^{it}|\lambda|$, however, in general, $|S+T| \not\leq |S| + |T|$.

Theorem 1.24 (Singular Value Decomposition)**.** *For any $T \in K(H_1, H_2)$ there exist orthonormal systems $\{e_i\}_{i\in N} \subseteq H_1$, $\{f_i\}_{i\in N} \subseteq H_2$ and numbers $s_1 \geq s_2 \geq \cdots \geq 0$, $s_k \to 0$ as $k \to \infty$ if $|N| = \infty$, such that*

$$Tx = \sum_{k\in N} s_k \langle e_k, x\rangle f_k, \quad \textit{for all } x \in H_1.$$

The numbers s_k^2 are the eigenvalues of T^*T, counted with multiplicities. The s_k are called the singular values *of T.*

Proof. Write $T = U|T|$ by polar decomposition, and use the spectral theorem on $|T|$ to get

$$|T|x = \sum_{k\in N} s_k \langle e_k, x\rangle e_k$$

where the e_k and the s_k have the properties in the theorem. Then let $f_k = Ue_k$ to get an orthonormal system $\{f_k\}_{k\in N}$ (since U is a partial isometry) and

$$Tx = U|T|x = \sum_{k\in N} s_k \langle e_k, x\rangle U e_k = \sum_{k\in N} s_k \langle e_k, x\rangle f_k. \qquad \square$$

Remark.

1. There is no condition that T be normal or self-adjoint!
2. SVD is important, also on matrices, in Numerics.

Much remains to be said about compact operators. They occur (and were first studied) as integral operators $(T_k f)(x) = \int_\Omega k(x,y) f(y)\,\mathrm{d}y$, $f \in X$, for various types of k, and various choices of X. We have seen the cases $X = C[0,1]$ and $k\colon [0,1]^2 \to \mathbb{C}$ a continuous function, and $X = L^2[0,1]$ and $k \in L^2([0,1]\times[0,1])$, and "weakly singular integral operators" where $k(x,y) \sim |x-y|^{-\alpha}$. Compact operators also occur as embeddings of certain spaces into certain spaces: $i\colon X \hookrightarrow Y$. We have seen this (most often) where i is bounded. Better yet, often i is even compact. Both situations are important in the study of boundary value problems; for example, the Dirichlet problem: Take an open bounded domain $\Omega \subseteq \mathbb{R}^2$ (or more generally $\Omega \subseteq \mathbb{R}^d$) and $\varphi \in C(\partial\Omega)$. The problem then is to find $u\colon \Omega \to \mathbb{C}$ such that $u \in C(\overline{\Omega})$, $u \in C^2(\Omega)$, solving $\Delta u = 0$ on Ω and $u|_{\partial\Omega} = \varphi$.

One can study compact operators according to the properties of their singular values. More precisely, let the p-th *Schatten-class* S_p be defined by

$$S_p = \Big\{ T \in K(H) \colon \|T\|_p := \Big(\sum_{n=1}^{\infty} s_n(T)^p\Big)^{1/p} < \infty \Big\}$$

with $s_n(T)$ the singular values of T. Then $\|-\|_p$ is a norm on S_p, and $(S_p, \|-\|_p)$ is Banach. These spaces have many properties in common with ℓ_p, for example $S_p' \cong S_q$ for q the Hölder conjugate of p. Of special interest are S_1, since it allows to define the *trace* of operators, and S_2, the *Hilbert-Schmidt operators.* In fact, S_2 is a Hilbert space. In particular, $T_k \in S_2(L^2[0,1])$ when $k \in L^2([0,1]\times[0,1])$.

For compact operators between Banach spaces, there is a notion of "nuclear operators": $T \in B(X,Y)$ is *nuclear* iff $Tx = \sum_{n=1}^{\infty} x_n'(x) y_n$ for all $x \in X$, with $\{x_n'\} \subseteq X'$ and $\{y_n\} \subseteq Y$, such that $\sum_n \|x_n'\| \|y_n\| < \infty$. For more on these operators, see for examples Werner, Alt.

2 Spectral theory for bounded, self-adjoint operators

As before, everything takes place on a complex Hilbert space. If $T\colon \mathbb{C}^n \to \mathbb{C}^n$ is normal, then it can be diagonalised, i.e. T is unitarily equivalent to a diagonal matrix D: $UTU^{-1} = D$ for some unitary U. The diagonal elements of D are of course the eigenvalues $\lambda_1, \ldots, \lambda_n$. Hence, $(UTU^{-1}x)_i = \lambda_i x_i$, i.e. D can be thought of as a multiplication operator. This form can be used to define $f(T)$ for certain f by $(Uf(T)U^{-1}x)_i = f(\lambda_i)x_i$. We have already seen something similar for compact operators, namely, when defined, $S = T^{1/2}$. Another way of thinking of $UTU^{-1} = D$ is $T = \sum_j \mu_j E_j$ with $\mu_j \neq \mu_k$ wehen $k \neq j$ and orthogonal projections E_j onto the eigenspaces. Then $f(T) = \sum_j f(\mu_j) E_j$. We shall elaborate on these various approaches.

Theorem 2.1 (Continuous functional calculus)**.** *Let $T \in B(H)$ be self-adjoint. Then there exists a unique map $\Phi\colon C(\sigma(T)) \to B(H)$ such that*

(a) $\Phi(t) = T$, *where* $t\colon \sigma(T) \to \mathbb{C}, t \mapsto t$, *and* $\Phi(1) = \mathrm{id}_H$, *where* $1\colon \sigma(T) \to \mathbb{C}, t \mapsto 1$.

(b) Φ *is an involution and a homomorphism of algebras, i.e.* $\Phi(\alpha f + g) = \alpha\Phi(f) + \Phi(g)$, $\Phi(fg) = \Phi(f)\Phi(g)$, *and* $\Phi(\overline{f}) = \Phi(f)^*$.

(c) Φ *is continuous, i.e. bounded from* $(C(\sigma(T)), \|-\|_\infty)$ *to* $(B(H), \|-\|)$. *In fact,* $\|\Phi(f)\| = \|f\|_\infty = \sup_{t \in \sigma(T)} |f(t)|$.

The map Φ is called the continuous functional calculus *of T.*

Proof. We will use Weierstraß' approximation theorem. The required properties fix Φ on the polynomials, which are dense in $C(\sigma(T))$, and by continuity there exists a unique bounded extension of Φ to $C(\sigma(T))$. □

Theorem 2.2 (Weierstraß' approximation theorem)**.** *Let $f \in C[a,b]$, $\varepsilon > 0$. Then there exists a polynomial p such that $\|f - p\|_\infty < \varepsilon$. In other words, the polynomials are dense in $(C[a,b], \|-\|_\infty)$.*

Theorem 2.3 (Tietze-Urysohn extension theorem for metric spaces)**.** *Let M be a metric space and $A \subseteq M$ a closed subset. For all continuous functions $f\colon A \to [c,d]$ there exists a continuous extension $F\colon M \to [c,d]$ (i.e. F is continuous and $F|_A = f$). In particular: $\forall t_0 \notin A\colon\ \exists \varphi\colon M \to [0,1]$ continuous with $\varphi|_A = 0$, $\varphi(t_0) = 1$.*

Proof. It is enough to look at $c = 0$, $d = 1$. Let

$$F(t) = \begin{cases} f(t) & t \in A \\ \dfrac{\inf\{f(s)d(s,t)\colon s \in A\}}{\inf\{d(s,t)\colon s \in A\}} & t \notin A \end{cases}$$

where $\inf\{d(s,t)\colon s \in A\} = d(t,A) > 0$ since $t \notin A$ and A is closed. One easily (!) sees that F is continuous. (For $t_0 \notin A$: Define (for $A_0 := A \cup \{t_0\}$) $\varphi_0\colon A_0 \to [0,1]$ by $\varphi_0|_A = 0$, $\varphi_0(t_0) = 1$ and then extend φ_0 to φ, as above). $\square$

Recall: For $T \in B(H)$, $\sigma(T) \subseteq \mathbb{C}$ is compact and in particular, closed. Let $f \in C(\sigma(T))$ (i.e. $f\colon \sigma(T) \to \mathbb{C}$ is continuous). Then, by Tietze-Urysohn there exists $\tilde{f}\colon B_{r(T)}(0) \to \mathbb{C}$ continuous with $\tilde{f}|_{\sigma(T)} = f$. For T self-adjoint ($T = T^*$), we have: $\exists m, M \in \mathbb{R}\colon \sigma(T) \subseteq [m,M] \subseteq \mathbb{R}$. In other words, there exists $\tilde{f}\colon [m,M] \to \mathbb{R}$ continuous and $\tilde{f}|_{\sigma(T)} = f$. Then, by Weierstraß, for $\varepsilon > 0$ there exists a polynomial p such that $\|\tilde{f} - p\|_\infty < \varepsilon$ (here $\|g\|_\infty = \sup_{t \in [m,M]} |g(t)|$). In particular, for all $t \in \sigma(T)$: $|f(t) - p(t)| = |\tilde{f}(t) - p(t)| < \varepsilon$, so $\|f - p\|_\infty < \varepsilon$ (here $\|g\|_\infty = \sup_{t \in \sigma(T)} |g(t)|$). Hence (!) the polynomials are dense in $(C(\sigma(T)), \|\cdot\|_\infty)$.

Proof of the continuous functional calculus. Uniqueness: Last time (!). Existence: Let $p \in \mathcal{P}$, then, if $p\colon t \mapsto \sum_{i=0}^n a_i t^i$, we define $\Phi_0(p) := \sum_{i=0}^n a_i T^i \in B(H)$. Clearly (!), Φ_0 satisfies (a) and (b). It remains to prove that $\Phi_0\colon \mathcal{P} \to B(H)$ is bounded. Then, by density of $\mathcal{P}$ in $C(\sigma(T))$ we will have a unique bounded (linear) extension Φ of Φ_0 with the desired properties. That Φ satisfies (a) is clear, linearity and boundedness too, remains: Φ is involutive and multiplicative: This is done by a limit argument. For involutive ($\Phi(\overline{f}) = (\Phi(f))^*$): Let $f \in C(\sigma(T))$ then, choose a sequence $\{p_n\}_{n\in\mathbb{N}}$ of polynomials, so $\|f - p_n\|_\infty \to 0$, $n \to \infty$. Then

$$\begin{aligned}\Phi(\overline{f}) &= \Phi\left(\overline{\lim_{n\to\infty} p_n}\right) = \Phi\left(\lim_{n\to\infty} \overline{p_n}\right) = \lim_{n\to\infty} \Phi(\overline{p_n}) = \\ &= \lim_{n\to\infty} \Phi_0(\overline{p_n}) = \lim_{n\to\infty} ((\Phi_0(p_n))^*) = \left(\lim_{n\to\infty} \Phi_0(p_n)\right)^* = \\ &= \left(\lim_{n\to\infty} \Phi(p_n)\right)^* = \left(\Phi\left(\lim_{n\to\infty} p_n\right)\right)^* = (\Phi(f))^*\end{aligned}$$

To prove $\Phi_0\colon \mathcal{P} \to B(H)$ is bounded we shall in fact prove that $\|\Phi_0(p)\| = \|p\|_\infty = \sup_{\lambda \in \sigma(T)} |p(\lambda)|$. For this, we need (problem 16, week 4)

$$\begin{aligned}\sigma(\Phi_0(p)) = \sigma(p(T)) = \sigma(\textstyle\sum_{i=0}^n a_i T^i) &= \{p(\lambda)\colon \lambda \in \sigma(T)\} = \\ &= \{\textstyle\sum_{i=0}^n a_i \lambda^i\colon \lambda \in \sigma(T)\} = p(\sigma(T))\end{aligned}$$

(OK: for $T \in B(X,Y)$, X,Y Banach).

$$\begin{aligned}\|\Phi_0(p)\|^2 &= \|\Phi_0(p)^*\Phi_0(p)\| = \|\Phi_0(\overline{p})\Phi_0(p)\| = \|\Phi_0(\overline{p}p)\| = \sup\{|\lambda|\colon \lambda \in \sigma(\Phi_0(\overline{p}p))\} = \\ &= \sup\{|(\overline{p}p)(\lambda)|\colon \lambda \in \sigma(T)\} = \sup\{|p(\lambda)|^2\colon \lambda \in \sigma(T)\} = \\ &= \left(\sup\{|p(\lambda)|\colon \lambda \in \sigma(T)\}\right)^2\end{aligned}$$

since, by properties of Φ_0, $\Phi_0(\overline{p}p)$ is self-adjoint, hence, normal, so its norm equals to its spectral radius. $\square$

Theorem 2.4. *$T \in B(H)$, $T = T^*$, $f \mapsto f(T)$ is the functional calculus of T (on $C(\sigma(T))$). Then* *(a)* $\|f(T)\| = \|f\|_\infty$

(b) If $f \geq 0$ *then* $f(T)$ *is a positive semi-definite operator.*

(c) $Tx = \lambda x \Longrightarrow f(T)x = f(\lambda)x$

(d) $\sigma(f(T)) = f(\sigma(T))$ *(spectral mapping theorem)*

(e) $\mathscr{A} := \{f(T) \colon f \in C(\sigma(T))\} \subseteq B(H)$ *is a commutative algera of operators.* $f(T)$ *is normal for all* $f \in C(\sigma(T))$*.* $f(T)$ *self-adjoint* $\Longleftrightarrow f = \overline{f}$ *(*f *is real).*

Proof.

(a) Follows from the proof of theorem 2.1, since true for polynomials, and these are dense.

(b) Let $f \geq 0$ (on $\sigma(T)$!). Let $f = g^2$ with $g \in C(\sigma(T))$ and $g \geq 0$ (i.e. $g = \sqrt{f}$ on $\sigma(T)$). Then, for all $x \in H$, $\langle x, f(T)x\rangle = \langle x, g^2(T)x\rangle = \langle x, g(T)g(T)x\rangle = \langle g(T)^*, g(T)x\rangle = \langle \overline{g}(T)x, g(T)x\rangle = \langle g(T)x, g(T)x\rangle = \|g(T)x\|^2 \geq 0$.

(c) This is true for $f \in \mathcal{P}$. For general f it follows from a limiting argument: For $f \in C(\sigma(T))$, choose polynomials $\{p_n\}_{n\in\mathbb{N}}$ such that $\|p_n - f\|_\infty \to 0$ as $n \to \infty$. Then (!) $\|p_n(T) - f(T)\| \to 0$, $n \to \infty$, in particular.

$$\begin{aligned}\|f(T)x - f(\lambda)x\| &= \|(f(T)x - p_n(T)x) + (p_n(T)x - p_n(\lambda)x) + (p_n(\lambda)x - f(\lambda)x\| \leq \\ &\leq \|f(T) - p_n(T)\|\|x\| + \|p_n(T)x - p_n(\lambda)x\| + |p_n(\lambda - f(\lambda)|\|x\| \\ &\xrightarrow{n\to\infty} 0\end{aligned}$$

(d) OK if f is a polynomial (problem 16, week 4). Let $f \in C(\sigma(T))$ and $\mu \notin f(\sigma(T))$. That is, $\mu \neq f(t)\ \forall t \in \sigma(T)$, so $g(t) = (\mu - f(t))^{-1}$, $t \in \sigma(T)$ is well-defined, and $g(t)(\mu - f(t)) = (\mu - f(t))g(t) = 1\ \forall t \in \sigma(T)$. Note $g \in C(\sigma(T))$, so $g(T)(\mu - f(T)) = (\mu - f(T))g(T) = \mathrm{id}_H$, hence $\mu \in \rho(f(T))$, hence $\mu \notin \sigma(f(T))$. This shows $\sigma(f(T)) \subset f(\sigma(T))$. On the other hand: Let $\mu = f(\lambda)$ for some $\lambda \in \sigma(T)$. Let $n \in \mathbb{N}$: Choose polynomials p_n with $\|p_n - f\|_\infty \leq 1/n$, then also $|p_n(\lambda) - f(\lambda)| \leq 1/n$ for all $\lambda \in \sigma(T)$ and $\|p_n(T) - f(T)\| \leq 1/n$ (since Φ is a linear isometry). As noted above, $\lambda \in \sigma(T) \Rightarrow p_n(\lambda) \in \sigma(p_n(T))$ (since p_n is a polynomial). Hence (exercise 18(iii), sheet 5) there exists $x_n \in H$, $\|x_n\| = 1$ and

$$\Big\|(p_n(T) - p_n(\lambda))x_n\Big\| \leq \frac{1}{n}$$

Hence

$$\begin{aligned}\|(f(T) - \mu)x_n\| &= \|(f(T) - f(\lambda))x_n\| = \\ &= \|(f(T) - p_n(T) + p_n(T) - p_n(\lambda) + p_n(\lambda - f(\lambda))x_n\| \leq \\ &\leq \|f(T) - p_n(T)\|\|x_n\| + \|(p_n(T) - p_n(\lambda))x_n\| + \\ &+ |p_n(\lambda) - f(\lambda)|\|x_n\| \leq \\ &\leq 1/n + 1/n + 1/n = 3/n\end{aligned}$$

Hence there exists a sequence $\{x_n\}_{n\in\mathbb{N}} \subset H$, $\|x_n\| = 1$, such that $\|(f(T) - \mu)x_n\| \to \infty$, $n \to \infty$. So $\{x_n\}_{n\in\mathbb{N}}$ is a Weyl sequence for $f(T)$ at μ $(= f(\lambda))$. Hence $\mu \in \sigma(f(T))$, so $f(\sigma(T)) \subset \sigma(f(T))$. So, in total $\sigma(f(T)) = f(\sigma(T))$.

(e) $f(T)$ is normal, since $f(T)f(T)^* = f(T)\overline{f}(T) = (f\overline{f})(T) = (\overline{f}f)(T) = f(T)^*f(T)$. The rest is clear. □

Remark (Holomorphic functional calculus). Let $\Omega \subseteq \mathbb{C}$ open, $f\colon \Omega \to \mathbb{C}$ holomorphic, Γ a rectifiable, closed curve in Ω, then, by Cauchy's Integral Formula,

$$f(z) = \frac{1}{2\pi i} \oint_\Gamma \frac{f(\xi)}{\xi - z}\,\mathrm{d}\xi \qquad \text{for any } z \text{ inside (!) } \Gamma$$

This allows the following construction (due to Dunford-Schwartz) of a holomorphic functional calculus:

For $T \in B(X)$ (X is a $\mathbb{C}$-Banach space), $\sigma(T) \subseteq D \subseteq C$ open, $f\colon D \to \mathbb{C}$ holomorphic, $\Gamma = \{\gamma_i\}_{i=1}^n$ a collection of (Jordan) curves in D, such that $\sigma(T)$ lies inside Γ and each γ_i is positively oriented. Then let (!)

$$f(T) := \frac{1}{2\pi i} \oint_\Gamma f(\xi)(\xi - T)^{-1}\,\mathrm{d}\xi$$

Then $f \mapsto f(T)$ is an involutive linear algebra homomorphism from $H(\sigma(T))$ to $B(X)$ (and the map is continuous in an appropriate sense).

The continuous functional calculus now allows (!) to 'compute' $f(T) \in B(H)$ for T bounded and self-adjoint, and f continuous. However, there is nothing on diagonalization.

Remark. Recall: Assume $T \in K(H)$, $T^* = T$ and let $T = \sum_{k=0}^\infty \mu_k E_k$ be the spectral decomposition of T from the spectral theorem for compact normal operators (1.20 + 1.21). (here, $\mu_0 = 0 \in \sigma(T)$ and E_0 is the orthogonal projections on $N(T)$, $\{\mu_k\}_{k=1}^\infty$ is the sequence of distinct and real eigenvalues of T, $\mu_k \neq 0$, $k \neq 0$. E_k are the corresponding projections on eigenspaces. From 1.20 and 1.21 it follows (!) that the map

$$\begin{aligned} C(\sigma(T)) &\longrightarrow B(H) \\ f &\longmapsto \sum_{k=0}^\infty f(\mu_k)E_k \end{aligned}$$

satisfies (a), (b) and (c) in theorem 2.1 (continuous functional calculus) — hence, this map *is* the continuous functional calculus of T.

That is, diagonalistation of $T \in K(H)$, $T^* = T$ allowed already to define the continuous calculus of T. On the other hand (in this case!), the continuous calculus allows to 're-find' the orthogonal spectral projections E_k. Recall that $\sigma(T) \setminus \{0\}$ consists of isolated points. Hence, define for $j \geq 1$, $f_j\colon \sigma(T) = \{0\} \cup \{\mu_1, \mu_2, \dots\} \to \mathbb{R}$ by $f_j(\mu_j) = 1$ and $f_j(t) = 0$ otherwise. Then (!!) f_j is continuous on $\sigma(T)$ and ($\mu_0 = 0$!)

$$f_j(T) = \sum_{k=1}^\infty f_j(\mu_k)E_k = E_j \qquad \Longrightarrow \qquad T = \sum_{k=1}^\infty \mu_k f_k(T)$$

Problem: If we imagine(!) there is, also for general bounded self-adjoint operators, something like spectral projections, one could hope to 'compute' them by using — as above — a function f with values in $\{0,1\}$ and define an operator $f(T)$. But: f has to be continuous to use the continuous calculus — but in general, '$f \in C(\sigma(T))$' and '$f(\sigma(T)) \in \{0,1\}$' are incompatible. Note: For $f(T)$ to be an orthogonal projection, f needs to take alues in $\{0,1\}$. (if $E^2 = E, E^* = E, E = f(T)$, f has to be real and $f(T) = E = E^2 = f(T)^2 = f^2(T)$, so $f^2 = f$, so f can only take the values 0 and 1 on $\sigma(T)$.) This is one (major!) motivation for the following: Extending the continuous functional calculus to define $f(T)$ for bounded, measurable functions on $\sigma(T)$ (i.e. Borel-measurable).

Definition. Let T be a set. A family of subsets $\Sigma \subseteq 2^T$ is called a σ-algebra iff

(a) $\emptyset \in \Sigma$

(b) $E \in \Sigma \Rightarrow T \setminus E \in \Sigma$

(c) $\{E_i\}_{i\in\mathbb{N}} \subseteq \Sigma \Rightarrow \bigcup_{i\in\mathbb{N}} E_i \in \Sigma$

For any family $\Gamma \subset 2^T$ there exists a smalles σ-algebra $\Sigma = \Sigma(\Gamma)$, containing Γ. The smallest σ-algebra in a topological space that contains the topology is called the *Borel-σ-algebra*. Its elements are called *Borel-sets*.

Definition. A family of subset $\Delta \subseteq 2^T$ is called a *Dynkin-system* iff

(a) $T \in \Delta$

(b) $E, F \in \Delta$, $E \subseteq F \Longrightarrow F \setminus E \in \Delta$.

(c) $E_1, E_2, \dots \in \Delta$, $E_i \cap E_j = \emptyset$ for $i \neq j \Longrightarrow \bigcup_{n=1}^{\infty} E_n \in \Delta$.

For a family $\Gamma \subseteq 2^T$ tehre exists a smallest Dynkin-system $\Delta(T)$ containing T. (Note: Σ σ-algebra $\Rightarrow$ Σ Dynkin-system).

A family $\Gamma \subseteq 2^T$ is called *stable under intersection* iff $A, B \in \Gamma \Longrightarrow A \cap B \in \Gamma$.

Proposition 2.5. *If $\Gamma \subseteq 2^T$ is stable under intersection, $\Sigma(\Gamma) = \Delta(\Gamma)$.*

Definition. A function $f\colon T \to \mathbb{R}$ is called *measurable (Borel-measurable)* iff $f^{-1}([a,b))$ is a Borel-set for all $a, b \in \mathbb{R}$.

A function $f\colon T \to \mathbb{C}$ is *measurable* iff both $\operatorname{Re} f$ and $\operatorname{Im} f$ are measurable.

Definition. A function of the form

$$f = \sum_{i=1}^{n} \alpha_i \chi_{E_i} \qquad \alpha_i \in \mathbb{C} \qquad E_i \in \Sigma$$

is called a *step function*. All step functions are measurable.

Proposition 2.6. *Let $f, g, f_n\colon T \to \mathbb{K}$ be measurable functions, $\alpha \in \mathbb{K}$.*

(a) *Then* $f+g$, fg, f/g $(g \neq 0)$, αf, $|f|$, $\max\{f,g\}$, $\min\{f,g\}$ *(if* $\mathbb{K}=\mathbb{R}$*),* $\sup_n f_n$, $\inf_n f_n$, $\liminf_n f_n$, $\limsup_n f_n$ *(and hence,* $\lim_n f_n$*) are all measurable.*

(b) *There exists a sequence* $\{\varphi_n\}_{n\in\mathbb{N}}$ *of step functions such that* $f(t)=\lim_{n\to\infty}\varphi_n(t)$ *for all* $t \in T$*. If* $f \geq 0$*, then the* φ_n *can be chosen such that* $\varphi_1(t), \leq \varphi_2(t) \leq \cdots \leq f(t)$ *for all* $t \in T$*.*

(c) *If* f *is measurable and bounded, then there exists a sequence of step functions that converges uniformly (on* T*) to* f*.*

Definition. Let T be a set, Σ a σ-algebra, $\mu\colon \Sigma \to \mathbb{R}$ (or $\mathbb{C}$) is called a *signed (or complex) measure* iff for all sequences $\{A_i\}_{i\in\mathbb{N}} \subseteq \Sigma$, $A_i \cap A_j = \emptyset$ $(i \neq j)$, $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$. (so $\mu(A) \geq 0$ is not needed!). $M(T,\Sigma)$ is the space of all complex measures on Σ. It s a vectorspace (!). If T is a topological/metric space, we let Σ be the Borel σ-algebra.

Definition. The *variation* of a measure $\mu \in M(T,\Sigma)$ is the positive measure $|\mu|$ defined by

$$|\mu|(A) = \sup_{\mathcal{Z}} \sum_{E\in\mathcal{Z}} |\mu(E)|$$

where the sup is over all decompositions $\mathcal{Z}$ of A in finitely many disjoint elemnts of Σ. Then (!!) $|\mu|(T) < \infty$. The variation norm (!) of μ is $\|\mu\| = |\mu|(T)$.

Proposition 2.7. $(M(T,\Sigma), \|\cdot\|)$ *is a Banach space.*

Theorem 2.8 (Riesz' representation theorem)**.** *Let* K *be a compact topological space (ex.* $[a,b]$*). Then* $C(K)'$ *(i.e. the dual of* $C(K)$*, with* $\|\cdot\|_\infty$*) is isometrically isomorphic to* $M(K,\Sigma)$*,* $C(K)' \cong M(K,\Sigma)$*,* Σ *being the Borel-sets on* K*, via the map* $T\colon M(K) \to C(K)', \mu \mapsto T\mu$ *given by*

$$(T\mu)(f) = \int_K f \,\mathrm{d}\mu$$

and $\|\mu\| = \|T\mu\|_{C(K)'}$*.*

Proof. See Rudin 'Real & Complex Analysis'.

Fact: $\|S\|^2 = \|S^*S\|$ for all $S \in B(H)$, since $\|Sx\|^2 = \langle Sx, Sx\rangle = \langle x, S^*Sx\rangle = |\langle x, S^*Sx\rangle|$.

Definition. Let $D \subseteq \mathbb{C}$ be compact. Let $\mathcal{M}_\mathrm{b}(D)$ be the set of (Borel) measurable and bounded (complex valued) functions on D. ($f \in \mathcal{M}_\mathrm{b}(D) \iff f\colon D \to \mathbb{C}$, f Borel-measurable and bounded).

We will need the following lemma:

Lemma 2.9.

(a) $(\mathcal{M}_\mathrm{b}(D), \|\cdot\|_\infty)$ *is a Banach space.*

(b) *Assume* $U \subseteq \mathcal{M}_\mathrm{b}(D)$ *satisfies:*

(1) $C(D) \subseteq U$

(2) $\{f_n\}_{n\in\mathbb{N}} \subseteq U$*,* $\sup_n \|f_n\|_\infty < \infty$*,* $f(t) := \lim_{n\to\infty} f_n(t)\ \forall t \in D \implies f \in U$*.*

Then $U = \mathcal{M}_\mathrm{b}(D)$*. So* $\mathcal{M}_\mathrm{b}(D)$ *is the smallest space of functions which*

(1) contains all continuous functions

(2) is closed under pointwise limits of uniformly bounded sequences.

Proof. (a) follows from the fact that the space of all bounded functions on D, with the supremum norm, is a Banach space, and that $\mathcal{M}_\mathrm{b}(D)$ is closed under pointwise limits. To prove (b) let

$$\mathscr{F} = \{S \subseteq \mathcal{M}_\mathrm{b}(D) \colon S \supseteq C(D) \text{ and } S \text{ satisfies (2) in lemma 2.9}\}$$

and let $V = \bigcap \mathscr{F}$. Note that $\sigma(D) \subseteq V \subseteq \mathcal{M}_\mathrm{b}(D)$. Then V is a vector space: for $f_0 \in V$ define $V_{f_0} = \{g \in \mathcal{M}_\mathrm{b}(D) \colon f_0 + g \in V\} \supseteq \{0\}$. For $f_0 \in C(D) \subseteq V$, we have $C(D) \subseteq V_{f_0}$. Now, V_{f_0} satisfies condition (2) for any $f_0 \in V$, for assume $\{f_n\}_{n\in\mathbb{N}} \subseteq V_{f_0}$, with $\|f_n\|_\infty \le C < \infty$, and $f(t) = \lim_{n\to\infty} f_n(t)$ exists for all $t \in D$. Since $f_n \in V_{f_0}$ for all n, we have $f_0 + f_n \in V$ for all n. Setting $h_n := f_n + f_0$ we have $\|h_n\|_\infty \le \|f_n\|_\infty + \|f\|_\infty \le C + \|f_0\| =: \widetilde{C}$ and $h_n(t) = f_n(t) + f_0(t) \xrightarrow{n\to\infty} f(t) + f_0(t)$ for all $t \in D$ and $\{h_n\}_{n\in\mathbb{N}} \subseteq V$. By definition of V, it follows that $f + f_0 \in V$, i.e. $f \in V_{f_0}$. Hence, V_{f_0} satisfies condition (2).

In summary, for $f \in C(D)$, $C(D) \subseteq V_{f_0}$, and, for any $f_0 \in V$, V_{f_0} satisfies condition (2). Hence, $V \subseteq V_{f_0}$ for $f_0 \in C(D)$. In other words, for $f_0 \in C(D)$ and $g \in V$, it follows that $f_0 + g \in V$. Now, take $g_0 \in V$. Then the above shows that $f + g_0 \in V$ for all $f \in C(D)$, hence $C(D) \subseteq V_{g_0}$. Also, V_{g_0} satisfies condition (2), hence $V \subseteq V_{g_0}$ for any $g_0 \in V$, which shows that V is closed under addition. Also, $g \in V$ and $\alpha \in \mathbb{C}$ implies $\alpha g \in V$. Hence, V is a linear subspace of $\mathcal{M}_\mathrm{b}(D)$.

We will show that $V = \mathcal{M}_\mathrm{b}(D)$. For this it suffices to show that all step functions lie in V, since the step functions are dense in $\mathcal{M}_\mathrm{b}(D)$: For any $f \in \mathcal{M}_\mathrm{b}(D)$ there exists a sequence $\{f_n\}_{n\in\mathbb{N}}$ of step functions such that $\|f_n - f\|_\infty \to 0$. In particular, $\|f_n\|_\infty \le C$ for some $C > 0$ and $\lim_{n\to\infty} f_n(t) = f(t)$ for all $t \in D$. Then $f \in D$ since V satisfies condition (2).

We still need to prove that $\chi_E \in V$ for all $E \in \Sigma$ (the Borel σ-algebra on D). Let $\Delta = \{E \in \Sigma \colon \chi_E \in V\}$. Claim: Δ is a Dynkin-system. First $\chi_D \equiv 1 \in \mathcal{P} \subseteq C(D) \subseteq V$. Let $E, F \in \Delta$, $\subseteq F$, then $\chi_{F \smallsetminus E} = \chi_F - \chi_E$ and so, since $\chi_E, \chi_F \in V$, also $\chi_{F\smallsetminus E} = \chi_F - \chi_E \in V$. With $E_1, E_2, \dots \in \Delta$, $E_i \cap E_j = \emptyset$ for $i \neq j$, $E = \bigcup_{k=1}^\infty E_k$, we have $\chi_E = \sum_{k=1}^\infty \chi_{E_k}$ (pointwise convergence). Let $f_n := \sum_{k=1}^n \chi_{E_k}$, then $\{f_n\}_{n\in\mathbb{N}} \subseteq V$ (V vectorspace). $\|f_n\|_\infty \le 1$ (E_n disjoint), and $f_n(t) \to \chi_E(t)$ as $n \to \infty$ for all $t \in D$, so $\chi_E \in V$, hence, $E \in \Delta$. So Δ is a Dynkin-system. Claim: $\mathcal{T} \subseteq \Delta$ (the open sets). Idea: Use 2.5 with $\Gamma = \mathcal{T}$. Then $\Delta(\mathcal{T}) = \Sigma(\mathcal{T}) = \Sigma$ and $\mathcal{T} \subseteq \Delta \subseteq \Sigma = \Delta(\mathcal{T})$. So $\Sigma = \Delta(\mathcal{T})$ being the smallest Dynkin-system containing $\mathcal{T}$, we have $\Delta = \Sigma$.

To prove $\mathcal{T} \subseteq \Delta$: For any open E (relatively open in D), there exists a sequence of continuous functions $\{f_n\}_{n\in\mathbb{N}}$, $0 \le f_n \le 1$ such that $f_n(t) \to \chi_E(t)$ for all $t \in D$. (Then: $\|fn\|_\infty \le 1$ and $f_n \in V$ for all n (since $C(D) \subseteq V$), and V satisfies (2) in Lemma 2.9, so $\chi_E \in V$, hence $E \in \Delta$.) To construct the sequence $\{f_n\}$: Use Tietze-Urysohn in the version: For K, H closed sets, $K \cap H = \emptyset$, $\exists$ g continuous such that $g|_H = 0$ and $g|_K = 1$, $0 \leq g \leq 1$. E is open, so E^c is closed. Use $F_n = \{y\colon \operatorname{dist}(y, E^c) \geq 1/n\}$, which is closed. Take f_n as g in Tietze-Urysohn for E^c, F_n. $\square$

We now extend the continuous functional calculus: Let $x, y \in H$, $f \in C(\sigma(T))$, $T \in B(H)$, $T = T^*$. Let $\ell_{x,y}(f) = \langle x, f(T), y\rangle \in \mathbb{C}$. Then $\ell_{x,y}\colon C(\sigma(T)) \to \mathbb{C}$ is clearly linear (since Φ is) and $|\ell_{x,y}(f)| \leq \|f(T)\|\|x\|\|y\| = \|x\|\|y\|\|f\|_\infty$, so $\ell_{x,y} \in C(\sigma(T))'$ (a bounded linear functional on $C(\sigma(T))$). By Riesz' representation theorem (theorem 2.8), there exists a complex measure $\mu_{x,y}$ such that

$$\langle x, f(T)y\rangle = \ell_{x,y}(f) = \int_{\sigma(T)} f \,\mathrm{d}\mu_{x,y} \tag{$**$}$$

Note: the map $H \times H \to M(\sigma(T)), (x, y) \mapsto \mu_{x,y}$ is sesquilinear (i.e. bilinear except $\mu_{\lambda x,y} = \overline{\lambda}\mu_{x,y}$). Also $\|\mu_{x,y}\| = \|\ell_{x,y}\| \leq \|x\|\|y\|$ (see above). So $(x, y) \mapsto \mu_{x,y}$ is bounded. (here, $\|\mu_{x,y}\|$ is the total variation of the complex measure $\mu_{x,y}$. Note: the right side of $(**)$ makes sense not only for $f \in C(\sigma(T))$, but for all $f \in \mathcal{M}_\mathrm{b}(\sigma(T))$, which helps to define $f(T)$ for $f \in \mathcal{M}_\mathrm{b}(\sigma(T))$.

Theorem 2.10 (Measurable functional calculus)**.** *Let $T \in B(H)$, $T = T^*$. Then there exists a unique map $\widehat{\Phi}\colon \mathcal{M}_\mathrm{b}(\sigma(T)) \to B(H)$ such that*

(a) $\widehat{\Phi}(t) = T$, $\widehat{\Phi}(1) = \mathrm{id}_H$

(b) $\widehat{\Phi}$ *is an involutive algebra homomorphism.*

(c) $\widehat{\Phi}$ *is continuous (bounded).*

(d) If $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathcal{M}_\mathrm{b}(\sigma(T))$, $\sup_n \|f_n\|_\infty$ *such that* $f_n(t) \to f(t)$, $n \to \infty$ $\forall t \in \sigma(T)$, *then* $\left\langle x, \widehat{\Phi}(f_n)y\right\rangle \to \left\langle x, \widehat{\Phi}(f)y\right\rangle$ $\forall x, y \in H$.

(d') Under the same assumptions as in (d), $\widehat{\Phi}(f_n)x \to \widehat{\Phi}(f)x$ $\forall x \in H$*.c*

Proof. Uniqueness: By the theorem on the continuous functional calculus, the conditions (a), (b), (c) determine uniquely $\widehat{\Phi}(f)$ for $f \in C(\sigma(T))$. Then uniqueness for $f \in \mathcal{M}_\mathrm{b}(\sigma(T))$ follows from lemma 2.5 (see condition (d)).

Existence: Let $f \in \mathcal{M}_\mathrm{b}(\sigma(T))$, $x, y \in H$. Let $\mu_{x,y}$ be the measure constructed above and look at the map

$$H \times H \ni (x, y) \longmapsto \int_{\sigma(T)} f \,\mathrm{d}\mu_{x,y} \in \mathbb{C}$$

This is sesquilinear (since $(x, y) \mapsto \mu_{x,y}$ is sesquilinear) and bounded since

$$\left|\int_{\sigma(T)} f \,\mathrm{d}\mu_{x,y}\right| \leq \|f\|_\infty\|\mu_{x,y}\| \leq \|f\|_\infty\|x\|\|y\|.$$

Hence, by Lax-Milgram, there exists a unique operator $\widehat{\Phi}(f) \in B(H)$, such that

$$\int_{\sigma(T)} f \,\mathrm{d}\mu_{x,y} = \left\langle x, \widehat{\Phi}(f)y\right\rangle,$$

and $\|\widehat{\Phi}(f)\| \leq \|f\|_\infty$. It remains to verify, that the map $\widehat{\Phi}\colon \mathcal{M}_\mathrm{b}(\sigma(T)) \to B(H)$ satisfies conditions (a) through (d'):

(a) Note that for $f \in C(\sigma(T))$, $\widehat{\Phi}(f) = \Phi(f) = f(T)$ where Φ is the continuous functional calculus. Hence, $\widehat{\Phi}(t) = \Phi(t) = T$ and $\widehat{\Phi}(1) = \Phi(1) = \mathrm{id}_H$.

(c) This holds, because $\|\widehat{\Phi}(f)\| \leq \|f\|_\infty$.

(d) This follows from Lebesgue's theorem on dominated convergence, since constant functions are integrable on $\sigma(T)$: There exists some $C > 0$ such that $|f_n(t)| \leq C$ for all $n \in \mathbb{N}$ and all $t \in \sigma(T)$, and $f_n(t) \to f(t)$ for all $t \in \sigma(T)$. Hence,

$$\left\langle x, \widehat{\Phi}(f_n)y \right\rangle = \int_{\sigma(T)} f_n \,\mathrm{d}\mu_{x,y} \xrightarrow{n\to\infty} \int_{\sigma(t)} f \,\mathrm{d}\mu_{x,y} = \left\langle x, \widehat{\Phi}(f)y \right\rangle, \quad \forall x, y \in H.$$

(b) The map $\widehat{\Phi}$ is linear by construction. To prove that $\widehat{\Phi}(fg) = \widehat{\Phi}(f)\widehat{\Phi}(g)$ for all $f, g \in \mathcal{M}_\mathrm{b}(\sigma(T))$, note that for continuous f and g this follows from the multiplicativity of the continuous functional calculus. Now, let $g \in C(\sigma(T))$ and let

$$U = \{f \in \mathcal{M}_\mathrm{b}(\sigma(T))\colon \widehat{\Phi}(fg) = \widehat{\Phi}(f)\widehat{\Phi}(g)\} \subseteq \mathcal{M}_\mathrm{b}(\sigma(T)).$$

Then $C(\sigma(T)) \subseteq U$. We will use lemma 2.9 to conclude that $U = \mathcal{M}_\mathrm{b}(\sigma(T))$: Let $\{f_n\}_{n\in\mathbb{N}} \subseteq U$ such that there exists some $C > 0$ with $\|f_n\|_\infty \leq C < \infty$, and $f(t) := \lim_{n\to\infty} f_n(t)$ exists for all $t \in \sigma(T)$. By (d) $\left\langle x, \widehat{\Phi}(fg)y \right\rangle = \lim_{n\to\infty} \left\langle x, \widehat{\Phi}(f_n g)y \right\rangle = \lim_{n\to\infty} \left\langle x, \widehat{\Phi}(f_n)\widehat{\Phi}(g)y \right\rangle = \left\langle x, \widehat{\Phi}(f)\widehat{\Phi}(g)y \right\rangle$. Hence, $\widehat{\Phi}(fg) = \widehat{\Phi}(f)\widehat{\Phi}(g)$, so $f \in U$. Hence, $U = \mathcal{M}_\mathrm{b}(\sigma(T))$ by lemma 2.9, and $\widehat{\Phi}(fg) = \widehat{\Phi}(f)\widehat{\Phi}(g)$ for all $f \in \mathcal{M}_\mathrm{b}(\sigma(T))$ and $g \in C(\sigma(T))$. Now, for $f \in \mathcal{M}_\mathrm{b}(\sigma(T))$, let

$$V = \{g \in \mathcal{M}_\mathrm{b}(\sigma(T))\colon \widehat{\Phi}(f)\widehat{\Phi}(g) = \widehat{\Phi}(fg)\} \subseteq \mathcal{M}_\mathrm{b}(\sigma(T)).$$

Then the previous argument shows $C(\sigma(T)) \subseteq V$, and repeating it we get $V = \mathcal{M}_\mathrm{b}(\sigma(T))$. Hence, $\widehat{\Phi}(fg) = \widehat{\Phi}(f)\widehat{\Phi}(g)$ for all $f, g \in \mathcal{M}_\mathrm{b}(\sigma(T))$. Similarly, it follows that $\widehat{\Phi}(\overline{f}) = \widehat{\Phi}(f)^*$ for all $f \in \mathcal{M}_\mathrm{b}(\sigma(T))$.

(d') Note that (using (d) on $\{\overline{f_n} f_n\}_{n\in\mathbb{N}}$): $\|\widehat{\Phi}(f_n)x\|^2 = \|f_n(T)x\|^2 = \langle x, f_n(T)^* f_n(T)x\rangle = \left\langle x, \overline{f_n}(T) f_n(T)x \right\rangle = \left\langle x, (\overline{f_n} f)(T)x \right\rangle \to \left\langle x, (\overline{f} f)(T)x \right\rangle = \|\widehat{\Phi}(f)x\|^2$. Now, in a Hilbert space weak convergence and convergence of the norms implies strong convergence, Hence, $\widehat{\Phi}(f_n)x \xrightarrow{n\to\infty} \widehat{\Phi}(f)x$. □

Theorem 2.11 (Lax-Milgram)**.** *Let H be a $\mathbb{C}$-Hilbert space, $B\colon H \times H \to \mathbb{C}$ sesquilinear.*

(a) The following are equivalent:

(i) B is continuous.

(ii) B is partially continuous, i.e. $x \mapsto B(x, y)$ and $y \mapsto B(x, y)$ are continuous.

(iii) $\exists M \geq 0\colon |B(x, y)| \leq M\|x\|\|y\|$

(b) If B is continuous, there exists a unique $S \in B(H)$ such that

$$B(x, y) = \langle x, Sy \rangle \qquad \forall x, y \in H$$

Lemma 2.12. *Let $T \in B(H)$, $T = T^*$.*

(a) For $A \subseteq \sigma(T)$ a Borel set, $E_A := \chi_A(T)$ ($\chi_A \in \mathcal{M}_{\mathrm{b}}(\sigma(T))$) is an orthogonal projection.

(b) $\chi_\emptyset(T) = 0$, $\chi_{\sigma(T)}(T) = \mathrm{id}_H$.

(c) Let $A_1, A_2, \dots \subseteq \sigma(T)$ be disjoint Borel sets and let $A = \bigcup_{n=1}^\infty A_n$, $x \in H$. Then $\sum_{n=1}^\infty \chi_{A_n}(T)x = \chi_A(T)x$.

(d) For Borel sets $A, B \subseteq \sigma(T)$, $\chi_A(T)\chi_B(T) = \chi_{A\cap B}(T)$.

Note: In general, one does not have $\sum_{n=1}^\infty \chi_{A_n}(T) = \chi_A(T)$ (norm convergence), since, for this to hold, we need $\|\chi_{A_n}\| \to 0, n \to \infty$.

Proof.

(a) $\chi_A^2 = \chi_A$ and $\overline{\chi_A} = \chi_A$, so (since $\widehat{\Phi}$ is multiplicative and involutive), $\chi_A(T)\chi_A(T) = \chi_A^2(T) = \chi_A(T)$, $\chi_A(T)^* = \chi_A(T)$, so $\chi_A(T)$ is an orthogonal projection.

(b) $\chi_\emptyset(T) = \widehat{\Phi}(0) = 0$, $\chi_{\sigma(T)} = \widehat{\Phi}(1) = \mathrm{id}_H$, by measurable functional calculus.

(c) Let $f_k = \sum_{n=1}^k \chi_{A_n}$, then $\|f_k\|_\infty \le 1$ (since $f_k(t) \in \{0, 1\}$). Also $f_k(t) \to \sum_{n=1}^\infty \chi_{A_n}(t)$ and $\sum_{n=1}^\infty \chi_{A_n} = \chi_A(t) =: f(t)$ for all $t \in \sigma(T)$. Then, by (d') in theorem 2.10, $f_k(T)x \mapsto f(T)x$, i.e. $\sum_{n=1}^\infty \chi_{A_n}(T)x = \chi_A(T)x$.

(d) $\chi_A\chi_B = \chi_{A\cap B}$, so $\chi_A(T)\chi_B(T) = \widehat{\Phi}(\chi_A)\widehat{\phi}(\chi_B) = \widehat{\Phi}(\chi_A\chi_B) = \widehat{(}\chi_{A\cap B}) = \chi_{A\cap B}(T)$.

□

Lemma 2.12 says that the map $E \colon \Sigma \to B(H), A \mapsto \chi_{A\cap\sigma(T)}(T)$ (where Σ is the Borel σ-algebra of $\mathbb{R}$) is a spectral measure in the sense of the following definition:

Definition 2.13. Let Σ be the Borel σ-algebra on $\mathbb{R}$. A map $E \colon \Sigma \to B(H), A \mapsto E_A$ is called a *spectral measure* (*projection valued measure*) iff all E_A are orthogonal projections ($E_A^2 = E_A = E_A^*$) and

(a) $E_\emptyset = 0$, $E_\mathbb{R} = \mathrm{id}_H$

(b) For pairwise disjoint $A_1, A_2, \dots \in \Sigma$,

$$\sum_{n=1}^\infty (E_{A_n}x) = E_A x \qquad A = \bigcup_{n=1}^\infty A_n \qquad \forall x \in H$$

A spectral measure E is said to have *compact support* iff there exists a compact set $K \subseteq \mathbb{R}$ such that $E_K = \mathrm{id}_H$.

Note: one easily (!) sees that $E_A E_B = E_B E_A = E_{A\cap B}$ always hold for a spectral measure. We shall now discuss how to integrate a measurable bounded function wrt. a spectral measure. The idea will be to show that the measurable functional calculus is given exactly in this way.

Step 1 Let $f = \chi_A$, $A \in \Sigma$ be a characteristic function. Let

$$\int f \,\mathrm{d}E := E_A \quad \in B(H)$$

Step 2 Let f be a step function, $f = \sum_{i=1}^n \alpha_i \chi_{A_i}$, $\alpha_i \in \mathbb{C}$, $A_i \in \Sigma$. Then

$$\int f \,\mathrm{d}E := \sum_{i=1}^n \alpha_i E_{A_i} \quad \in B(H)$$

(to prove: this does not depend on the way f has been written as a step function)

Step 3 Let f be a measurable bounded function. Then there exists a sequence $\{f_n\}_{n\in\mathbb{N}}$ of step functions which converges uniformly to f. (see 'recall on Lebesgue theory'). ($f_n \to f$, uniformly). Idea: Let

$$\int f \,\mathrm{d}E = \lim_{n\to\infty} \int f_n \,\mathrm{d}E \quad \in B(H)$$

The existence of the limit and the independence of the choice of sequence $\{f_n\}$ follows from the following lemma, since if $f_n \to f$ uniformly, then $\{f_n\}$ is (uniformly) Cauchy, so: $\{\int f_n \,\mathrm{d}E\}_{n\in\mathbb{N}} \subseteq B(H)$ is also Cauchy (by 2.14). Hence the limit exists. Similarly, this limit is independent of the sequence $\{f_n\}$.

Lemma 2.14. *For any step function f we have*

$$\left\| \int f \,\mathrm{d}E \right\| \leq \|f\|_\infty$$

Proof. Let $x \in H$ with $\|x\| \leq 1$ and $f = \sum_{i=1}^n \alpha_i \chi_{A_i}$, $\alpha_i \in \mathbb{C}$, $A_i \in \Sigma$. Note: Can assume, wlog. that $A_i \cap A_j = \emptyset$ ($i \neq j$). Then

$$\begin{aligned}
\left\| \left(\int f \,\mathrm{d}E \right) x \right\|^2 &= \left\| \sum_{i=1}^n E_{A_i}(x) \right\|^2 = \left\langle \sum_{i=1}^n \alpha_i E_{A_i}(x), \sum_{j=1}^n \alpha_j E_{A_j}(x) \right\rangle = \\
&= \sum_{i=1}^n \sum_{j=1}^n \overline{\alpha_i}\alpha_j \langle E_{A_i}(x), E_{A_j}(x)\rangle = \sum_{i=1}^n |\alpha_i|^2 \langle E_{A_i}(x), E_{A_i}(x)\rangle \leq \\
&\leq \left(\max_{1\leq i\leq n} |\alpha_i|^2 \right) \sum_{i=1}^n \|E_{A_i}(x)\|^2 = \left(\max_{1\leq i\leq n} |\alpha_i|^2 \right) \left\| \sum_{i=1}^n E_{A_i}(x) \right\|^2 = \\
&= \|f\|_\infty^2 \left\| E_{\bigcup_{i=1}^n A_i}(x) \right\|^2 \leq \|f\|_\infty^2.
\end{aligned}$$

□

If E has compact support, for example $E_K = \mathrm{id}_H$ for $K \subseteq \mathbb{R}$ compact, and $f \colon K \to \mathbb{C}$ is bounded and measurable, then we set

$$\int f \,\mathrm{d}E := \int \chi_K f \,\mathrm{d}E = \int_K f \,\mathrm{d}E.$$

This definition is independent of the choice of K, since if $E_K = \mathrm{id}_H$ and $K \cap A = \emptyset$, then $E_A = 0$. In summary we have

Theorem 2.15. *Let E be a spectral measure (on $\mathbb{R}$) and f a bounded, measurable function. Then*

$$\int f \,\mathrm{d}E \quad \in B(H)$$

is well-defined, and the map $\mathcal{M}_{\mathrm{b}}(\mathbb{R}) \to B(H), f \mapsto \int f \,\mathrm{d}E$ is linear and bounded. In fact,

$$\left\| \int f \,\mathrm{d}E \right\| \leq \|f\|_\infty$$

If f is real, then $\int f \,\mathrm{d}E$ is self-adjoint. If $K \subseteq \mathbb{R}$ is compact with $E_K = \mathrm{id}_H$, then it is enough that f be defined and bounded and measurable on K.

Remark. Let E be a spectral measure with compact support. Then $\int f \,\mathrm{d}E$ is well-defined for f being a polynomial. Hence,

$$T := \int_K \lambda \,\mathrm{d}E_\lambda = \int \mathrm{id}_K \;\mathrm{d}E_\lambda$$

is a bounded, self-adjoint operator on H. Hence, given a bounded self-adjoint operator T, the measurable functional calculus gives rise to a spectral measure, by $E \colon \Sigma \to B(H), A \mapsto \chi_{A \cap \sigma(T)}(T) =: E_A$, with compact support, since $\sigma(T) \subseteq \mathbb{R}$ is compact. On the other hand, given a spectral measure with compact support, $T = \int_K \lambda \,\mathrm{d}E_\lambda$ defines a self-adjoint bounded operator. We will see that these are inverse operations.

Theorem 2.16. *Let E be a spectral measure on $\mathbb{R}$ with compact support and let T be the bounded self-adjoint operator given by*

$$T = \int \lambda \,\mathrm{d}E_\lambda.$$

Then the map $\Psi \colon \mathcal{M}_{\mathrm{b}}(\sigma(T)) \to B(H), f \mapsto \int_{\sigma(T)} f \,\mathrm{d}E$ is the (unique) measurable functional calculus associated to T, i.e. $\Psi = \widehat{\Phi}$ in the notation of theorem 2.10. In particular, $E_{\sigma(T)} = \mathrm{id}_H$ and for measurable $A \subseteq \sigma(T)$,

$$\chi_A(T) = \int \chi_A \,\mathrm{d}E.$$

Hence, the given spectral measure and that defined through the operator T coincide.

Proof. We will prove that Ψ satisfies (a) through (d) in theorem 2.10, which will show that $\Psi = \widehat{\Phi}$ by uniqueness. Let E be a spectral measure on $\mathbb{R}$ with compact support, say $E_K = \mathrm{id}_H$, and let $T = \int_K \lambda \, \mathrm{d}E_\lambda$. Let $f\colon \sigma(T) \to \mathbb{C}$ be bounded and measurable and define an extension

$$\tilde{f}(t) = \begin{cases} 0 & t \notin \sigma(T) \\ f(t) & t \in \sigma(T) \end{cases}$$

Then $\tilde{f}$ is measurable and bounded, so by theorem 2.15, $\Psi(f) := \int \tilde{f} \, \mathrm{d}E$ exists, and $\Psi\colon \mathcal{M}_\mathrm{b}(\sigma(T)) \to B(H)$ is linear and continuous. Ψ is also multiplicative: For A, B measurable sets

$$\Psi(\chi_{A\cap B}) = \int \chi_{A\cap B} \, \mathrm{d}E = E_{A\cap B} = E_A E_B = \left(\int \chi_A \, \mathrm{d}E\right)\left(\int \chi_B \, \mathrm{d}E\right) = \Psi(\chi_A)\Psi(\chi_B).$$

Since $\chi_{A\cap B} = \chi_A \chi_B$, it follows that $\Psi(\chi_A\chi_B) = \Psi(\chi_A)\Psi(\chi_B)$. Then, by linearity, $\Psi(fg) = \Psi(f)\Psi(g)$ for all step functions f, g. Finally, by a limiting argument, also $\Psi(fg) = \Psi(f)\Psi(g)$ for $f, g \in \mathcal{M}_\mathrm{b}(\sigma(T))$. Similarly, $\Psi(\overline{f}) = \Psi(f)^*$ since $E_A^* = E_A$. Hence, Ψ satisfies (b) and (c) in theorem 2.10.

To prove (d), consider any sequence $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathcal{M}_\mathrm{b}(\sigma(T))$ such that $\|f_n\|_\infty \le C$ for some C and $f(t) = \lim_{n\to\infty} f_n(t)$ exists for all $t \in \sigma(T)$. Note that for E a spectral measure, and fixed $x, y \in H$, the map $\nu_{x,y}\colon \Sigma \to \mathbb{C}, A \mapsto \langle x, E_A y\rangle$ is a complex measure and that

$$\langle x, \Psi(f)y\rangle = \int f \, \mathrm{d}\nu_{x,y} = \int f(\lambda) \, \mathrm{d}\langle x, E_\lambda y\rangle,$$

since this is trivial for $f = \chi_A$ and then follows for step functions by linearity and for arbitrary $f \in \mathcal{M}_\mathrm{b}(\sigma(T))$ by approximation. Then (d) follows for Ψ from Lebesgue's theorem on dominated convergence:

$$\langle x, \Psi(f_n)y\rangle = \int f_n \, \mathrm{d}\nu_{x,y} \xrightarrow{n\to\infty} \int f \, \mathrm{d}\nu_{x,y} = \langle x, \Psi(f)y\rangle$$

since constant functions are integrable by compact support of E.

It remains to prove (a), i.e. that $\Psi(1) = \mathrm{id}_H$ and $\Psi(t) = T$ for $t\colon \sigma(T) \to \mathbb{C}, t \mapsto t$. The first is equivalent to proving $E_{\sigma(T)} = \mathrm{id}_H$. The second follows from the first, since

$$\Psi(t) = \int_{\sigma(T)} \lambda \, \mathrm{d}E_\lambda = T$$

from theorem 2.15. So it remains to prove that $E_{\sigma(T)} = \mathrm{id}_H$. For this we prove $E_{\rho(T)} = 0$ and use that $\mathbb{R} = \sigma(T) \cup \rho(T)$ and $\sigma(T) \cap \rho(T) = \emptyset$. E has, by assumption, compact support K. Choose an interval $(a, b]$ such that $K \subseteq (a, b]$. Then $E_{(a,b]} = E_K = \mathrm{id}_H$. Let $\mu \in \rho(T)$. Then there exists a neighbourhood U of μ such that $E_U = 0$: Note that $T - \mu I$, and hence $\mu I - T$, is invertible. It follows by 0.14 that there exists $\delta > 0$ such that $\|S - (\mu - T)\| \le \delta$ implies that S is invertible, and $\|S^{-1}\| \le C := \|(\mu - T)^{-1}\| + 1$. We can assume, that $\delta = \frac{b-a}{N}$ for some $N \in \mathbb{N}$ and that $\delta < \frac{1}{C}$. Let $a_k = a + k\delta$ for $k = 0, 1, \ldots, N$ and consider the step function $f = \sum_{k=1}^N a_k \chi_{(a_{k-1}, a_k]}$. By 2.15, with $E_k = E_{(a_{k-1}, a_k]} = \int \chi_{(a_{k-1}, a_k]} \, \mathrm{d}E$, we have

$$\left\| T - \sum_{k=1}^N a_k E_k \right\| = \left\| \int \lambda \, \mathrm{d}E - \int f \, \mathrm{d}E \right\| \le \|\lambda - f\|_\infty = \max_{k=0,\ldots,N} |a_k - a_{k-1}| \le \delta$$

Note: $(a_{k-1}, a_k] \cap (a_{j-1}, a_j] = \emptyset$, $k \neq j$, so $(a,b] = \bigcup_{k=1}^N (a_{k-1}, a_k]$ (disjoint union!), so $\mathrm{id} = E_{(a,b]} = \sum_{k=1}^N E_{(a_{k-1},a_k]} = \sum_{k=1}^N E_k$ ($K \subseteq (a,b]$, $E_K = \mathrm{id}$). So $\mu\,\mathrm{id} = \sum_{k=1}^N \mu E_k$, hence

$$\left\| (\mu - T) - \sum_{k=1}^N (\mu - a_k) E_k \right\| = \left\| T - \sum_{k=1}^N a_k E_k \right\| \leq \delta$$

hence, the operator $\sum_{k=1}^N (\mu - a_k) E_k$ is invertible and $\|(\sum_{k=1}^N (\mu - a_k)E_k)^{-1}\| \leq C$. But (since $E_k E_j = E_j E_k = \delta_{kj} E_k$, since the $(a_{i-1}, a_i]$ are disjoint)

$$\left\| \left(\sum_{k=1}^N (\mu - a_k) E_k \right)^{-1} \right\| = \max \left\{ |\mu - a_k|^{-1} \colon E_k \neq 0 \right\}$$

So $|\mu - a_k|^{-1} < C$ implies $E_k = 0$. Now choose N so large that, for some a_{k_0}, $|\mu - a_{k_0}| < 1/C$. Then $E_{k_0} = 0$. Hence $E_U = 0$ for some neighbourhood U. Let now $\tilde{K} \subseteq \rho(T)$ be compact, and $\{U_\mu \colon \mu \in \tilde{K}\}$ is an open cover, $\exists \mu_1, \ldots, \mu_n \in \tilde{K}$ such that $\tilde{K} \subseteq \bigcup_{i=1}^n U_{\mu_i}$, then $E_{\tilde{K}} = 0$. There exist $V_1, \ldots, V_M$ disjoint such that $\bigcup_{i=1}^n U_{\mu_i} = \bigcup_{i=1}^M V_i$, $V_i \cap V_j = \emptyset$ for $i \neq j$ and $\forall i \exists j \colon V_i \subseteq U_{\mu_j}$. So $E_{\bigcup U_{\mu_i}} = E_{\bigcup V_j} = \sum_{j=1}^M E_{V_j} = 0$. Recall: $\forall \tilde{K} \subseteq \mathbb{R} \setminus \sigma(T)$, $\tilde{K}$ compact: $E_{\tilde{K}} = 0$. It follows that, for all $x \in H$, the positive measure $\Sigma(K \setminus \sigma(T)) \to \mathbb{R}_+, A \mapsto \langle x, E_A x \rangle$ is regular (see definition below). Since, for all $\tilde{K} \subseteq \mathbb{R} \setminus \sigma(T)$, $\langle x, E_{\tilde{K}} x \rangle = \langle x, 0 \rangle = 0$, it follows that this measure is 0 on $K \setminus \sigma(T)$. Hence, $\langle x, E_{\mathbb{R} \setminus \sigma(T)} x \rangle = 0$ for all $x \in H$, so $E_{\mathbb{R} \setminus \sigma(T)} = 0$. Hence, $E_{\sigma(T)} = \mathrm{id}$. $\square$

Definition. A measure on the Borel σ-algebra is called *Borel-measure*. A positive Borel-measure μ is called *regular* iff

(a) $\mu(C) < \infty$ for all compact C.

(b) $\forall A \in \Sigma$ (Borel-sets)

$$\mu(A) = \sup\{\mu(C) \colon C \subseteq A, C \text{ compact}\} = \inf\{\mu(O) \colon O \supseteq A, O \text{ open}\}$$

A signed/complex measure is regular iff its variation measure $|\mu|$ is regular.

Theorem. *Let T be (a) a compact metric space or (b) a complete, seperable metric space or (c) an open subset of $\mathbb{R}^n$. Then every finite Borel measure ($\mu(T) < \infty$) μ on T is regular. Also, the Lebesgue measure on $\mathbb{R}^n$ is regular.*

Recall: $A \subseteq \sigma(T)$, measurable (Borel)

$$\chi_A(T) = \Psi(\chi_A) = \int \chi_A \,\mathrm{d}E = E_A$$

so the spectral measure associated to T ($A \mapsto \chi_{A \cap \sigma(T)}(T)$). and the spectral measure used to define T ($= \int \lambda \,\mathrm{d}E$, $A \mapsto E_A$) coincide.

Theorem 2.17 (Spectral theorem for self-adjoint bounded operators)**.** *Let $T \in B(H)$, $T = T^*$. Then there exists a unique spectral measure E, with compact support in $\mathbb{R}$, such that*

$$T = \int_{\sigma(T)} \lambda \,\mathrm{d}E_\lambda$$

The map

$$f \mapsto \int_{\sigma(T)} f(\lambda) \,\mathrm{d}E_\lambda = f(T)$$

defines the measurable functional calculus $f(T)$ given by

$$\langle x, f(T)y\rangle = \int_{\sigma(T)} f(\lambda) \,\mathrm{d}\langle x, E_\lambda y\rangle \qquad x, y \in H$$

Proof. Let $T \in B(H)$, $T = T^*$ and let E be its associated spectral measure ($E\colon A \mapsto \chi_{A\cap\sigma(T)}(T)$). Define $S := \int_{\sigma(T)} \lambda \,\mathrm{d}E_\lambda$. We need to prove $S = T$. (the rest of the statements in 2.17 are recapitulations). Firstly, E has compact support: $\sigma(T) \subseteq \mathbb{C}$ compact, and $\chi_{\sigma(T)}(T) = E_{\sigma(T)} = \mathrm{id}$ (lemma 2.12 b). Let, for $g \in \mathcal{M}_\mathrm{b}(\sigma(T))$, $g(T)$ denote the measurable functional calculus of T (from 2.10) on g and $\Psi(g)$ the measurable functional calculus of S (from 2.16) on g. Let t be the function $t\colon \sigma(T) \to \mathbb{R}, t \mapsto t$, and choose, for $\varepsilon > 0$, a step function f on $\sigma(T)$ such that $\|t - f\|_\infty \leq \varepsilon$. According to 2.10 $\|T - f(T)\| = \|t(T) - f(T)\| \leq \|t - f\|_\infty \leq \varepsilon$ and according to 2.15 (used on the spectral measure associated to T, and $S := \int_{\sigma(T)} \lambda \,\mathrm{d}E_\lambda$), $\|S - \Psi(f)\| = \|\int_{\sigma(T)} (t - f) \,\mathrm{d}E\| \leq \|t - f\|_\infty \leq \varepsilon$. Finally, with $f = \sum \alpha_i \chi_{A_i}$ ($\alpha_i \in \mathbb{C}$, $A_i \in \Sigma$), and since, by definition, $E_{A_i} = \chi_{A_i}(T)$ (the spectral measure E is the one associated to T: $A \mapsto \chi_{A\cap\sigma(T)}(T)$

$$f(T) - \Psi(f) = \sum_{i=1}^{n} \alpha_i \chi_{A_i}(T) - \sum_{i=1}^{n} \alpha_i E_{A_i} = 0$$

Hence, for all $\varepsilon > 0$, $\|S - T\| \leq \|S - \Psi(f)\| + \|\Psi(f) - f(T)\| + \|f(T) - T\| \leq \varepsilon + 0 + \varepsilon = 2\varepsilon$, so $\|S - T\| = 0$, hence, $T = S = \sum_{\sigma(T)} \lambda \,\mathrm{d}E_\lambda$. $\square$

Remark.

(1) This is the generalization of $T = \sum \mu_i E_i$ for compact operators (in particular, matrices) — hence, T is built of the orthogonal projections E_A (weighted with λ) but in a continuous way.

(2) The symbol $\mathrm{d}\langle x, E_\lambda y\rangle$ means integration wrt. the Borel measure $A \mapsto \langle x, E_A y\rangle$ ($x, y \in H$ fixed). (In the proof of 2.10 it was called $\mu_{x,y}$).

(3) T as before ($T \in B(H)$, $T = T^*$), $S \in B(H)$ any operator. Then $A \mapsto \langle x, SE_A y\rangle = \langle S^* x, E_A y\rangle$ ($x, y \in H$) is also a complex measure. (check!)

Corollary 2.18. *Let $T \in B(H)$ be self-adjoint with spectral measure E. Let $S \in B(H)$. Then $ST - TS = [S, T] = 0$ if and only if $[S, E_A] = 0$ for all $A \in \Sigma$.*

Proof. One has $[S,T]=0$ if and only if $[S,T^n]=0$ for all $n\geq 0$ if and only if $\langle x, ST^n y\rangle = \langle x, T^n S y\rangle$ for all $x,y\in H$ and $n\geq 0$. Note that

$$\langle x, ST^n y\rangle = \langle S^* x, T^n y\rangle = \int_{\sigma(T)} \lambda^n \,\mathrm{d}\langle S^* x, E_\lambda y\rangle = \int_{\sigma(T)} \lambda^n \,\mathrm{d}\langle x, SE_\lambda y\rangle$$

and

$$\langle x, T^n S y\rangle = \int_{\sigma(T)} \lambda^n \,\mathrm{d}\langle x, E_\lambda S y\rangle.$$

Hence, $[S,T]=0$ if and only if

$$\int_{\sigma(T)} \lambda^n \,\mathrm{d}\langle x, SE_\lambda y\rangle = \langle x, T^n S y\rangle = \int_{\sigma(T)} \lambda^n \,\mathrm{d}\langle x, E_\lambda S y\rangle$$

for all $x,y\in H$ and $n\geq 0$. So, thinking of the measures $\mathrm{d}\langle x, SE_\lambda y\rangle$ and $\mathrm{d}\langle x, E_\lambda S y\rangle$ as linear functionals on $C(\sigma(T))$, these coincide on the polynomial functions for all $x,y\in H$ if and only if $[S,T]=0$. Since the polynomial functions are dense in $C(\sigma(T))$, the measures $\mathrm{d}\langle x, SE_\lambda y\rangle$ and $\mathrm{d}\langle x, E_\lambda S y\rangle$ for all $x,y\in H$ define the same linear functional on $C(\sigma(T))$ if and only if $[S,T]=0$. Hence, by Riesz' representation theorem, the two measures coincide for all $x,y\in H$ if and only if $[S,T]=0$. Now,

$$\langle x, SE_A y\rangle = \int_{\sigma(T)} \chi_A \,\mathrm{d}\langle x, SE_\lambda y\rangle = \int_{\sigma(T)} \chi_A \,\mathrm{d}\langle x, E_\lambda S y\rangle = \langle x, E_A S y\rangle$$

for all $x,y\in H$ if and only if $[S,T]=0$. Hence, $[S,E_A]=0$ if and only if $[S,T]=0$. □

Example 2.19.

(a) Assume $H\cong\mathbb{C}^n$ to be finite dimensional and T a self-adjoint matrix, and assume T has m distinct eigenvalues $\mu_1,\ldots,\mu_m$. Then $T=\sum_{i=1}^m \mu_i E_{\{\mu_i\}}$ where the $E_{\{\mu_i\}}$ are the orthogonal projections on the eigenspace corresponding to μ_i. The spectral measure of T is $E_A=\sum_{\mu_i\in A} E_{\{\mu_i\}}$.

(b) Let $T\in K(H)$ be self-adjoint. Then

$$T=\sum_{i=0}^\infty \mu_i E_{\{\mu_i\}}$$

in norm with the μ_i and $E_{\{\mu_i\}}$ as in (a). As above, the spectral measure is, for $A\in\Sigma$,

$$E_A=\sum_{\mu_i\in A} E_{\{\mu_i\}},$$

but only pointwise!

(c) Let $H=L^2[0,1]$ and $(Tx)(t)=tx(t)$ for $x\in L^2[0,1]$. Then, T is self-adjoint, $\sigma(T)=\sigma_c(T)=[0,1]$. For $A\in\Sigma$, set

$$E_A x=\chi_{A\cap[0,1]}x,\quad x\in H=L^2[0,1].$$

Then E is the spectral measure of T, and

$$T=\int_{[0,1]}\lambda\,\mathrm{d}E_\lambda.$$

Theorem 2.20. *Let $T \in B(H)$ be self-adjoint with spectral measure E. Then*

(a) $\lambda \in \rho(T)$ iff there exists and open neighbourhood $U \subseteq \mathbb{R}$ of λ such that $E_U = 0$.

(b) λ is an eigenvalue of T iff $E_{\{\lambda\}} \neq 0$. In this case, $E_{\{\lambda\}}$ is the orthogonal projection on the eigenspace associated to λ.

(c) The isolated point λ of $\sigma(T)$ are eigenvalues.

Proof.

(a) By construction of E, $E_{\rho(T)} = 0$. Assume $U \subseteq \mathbb{R}$ is an open neighbourhood of λ such that $E_U = 0$. Define

$$f(t) = \begin{cases} \frac{1}{\lambda - t} & t \notin U \\ 0 & t \in U \end{cases}$$

Then f is measurable and bounded on $\sigma(T)$, i.e. $f \in \mathcal{M}_{\mathrm{b}}(\sigma(T))$. So is $g(t) = \lambda - t$. Then, by the functional calculus, $f(T)(\lambda - T) = f(T)g(T) = (fg)(T) = \chi_{U^c}(T) = E_{U^c}(T) = \mathrm{id}_H$ since $E_U = 0$. Similarly, $(\lambda - T)f(T) = \mathrm{id}_H$. Hence, $\lambda \in \rho(T)$.

(b) It is enough to prove $\mathrm{R}(E_{\{\lambda\}}) = \mathrm{N}(\lambda - T)$. So let $x \in \mathrm{R}(E_{\{\lambda\}})$, i.e. $E_{\{\lambda\}}x = x$. Therefore,

$$\langle y, (\lambda - T)x \rangle = \left\langle y, (\lambda - T)E_{\{\lambda\}}x \right\rangle = \int_{\sigma(T)} (\lambda - t)\chi_{\{\lambda\}}(t) \,\mathrm{d}\langle x, E_t y \rangle = 0$$

since $(\lambda - t)\chi_{\{\lambda\}}(t) = 0$ for all t for all $y \in H$. Hence, $x \in \mathrm{N}(\lambda - T)$. On the other hand, for $x \in \mathrm{N}(\lambda - T)$, i.e. $Tx = \lambda x$, by 2.4(c) $f(T)x = f(\lambda)x$ for $f \in C(\sigma(T))$ and therefore by 2.9(b), this also holds for any $f \in \mathcal{M}_{\mathrm{b}}(\sigma(T))$. In particular for $f = \chi_{\{\lambda\}}$ one has $E_\lambda x = x$, hence $x \in \mathrm{R}(\lambda - T)$.

(c) Let $U \subseteq \mathbb{R}$ be open such that $U \cap \sigma(T) = \{\lambda\}$. Then $U \setminus \{\lambda\} \subseteq \mathbb{R} \setminus \sigma(T) = \rho(T)$, so $E_{U \setminus \{\lambda\}} = 0$. If $E_{\{\lambda\}} = 0$, then $E_U = E_{U \setminus \{\lambda\}} + E_{\{\lambda\}} = 0$, i.e. $\lambda \in \rho(T)$. But by assumption $\lambda \in \sigma(T)$, hence $E_{\{\lambda\}} \neq 0$ and by (b) λ is an eigenvalue. ☐

Corollary 2.21. *Let $T \in B(H)$ be self-adjoint with sepctral measure E. Then $\sigma(T)$ is the smallest compact set such that $E_{\sigma(T)} = \mathrm{id}_H$.*

Proof. Take some compact $K \subseteq \mathbb{R}$ with $E_K = \mathrm{id}_H$ and $\lambda \notin K$. Then $E_{K^c} = 0$ and $\lambda \in K^c$ which is open. So by 2.20(a) $\lambda \in \rho(T)$. Hence, $\sigma(T) \subseteq K$. ☐

One way (among many others) of thinking of the spectral theorem for self-adjoint matrices is: Every self-adjoint linear map $T \colon \mathbb{C}^n \to \mathbb{C}^n$ is unitarily equivalent to a diagonal matrix, i.e. there exists a unitary map $U \colon \mathbb{C}^n \to \mathbb{C}^n$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ such that $(UTU^*x)_i = \lambda_i x_i$.

Theorem 2.22. *Let $T \in B(H)$ be self-adjoint with spectral measure E and assume T has a cyclic vector, i.e. there exists $x_0 \in H$ such that $\overline{\operatorname{span}\{T^i x_0 \colon i \in \mathbb{N}\}} = H$. Let μ be the finite positive measure $\mathrm{d}\langle x_0, E_\lambda x_0 \rangle$. Then there exists a unitary operator $U \colon H \to L^2(\mathbb{R}, \mu)$ such that*

$$(UTU^{-1}\varphi)(t) = t\varphi(t), \quad \text{for all } \varphi \in L^2(\mathbb{R}, \mu).$$

Proof. Let $\varphi \in C(\sigma(T))$, then

$$\int_{\sigma(T)} |\varphi(t)|^2 \,\mathrm{d}\mu(t) = \int_{\sigma(T)} \overline{\varphi(t)}\varphi(t)\,\mathrm{d}\langle x_0, E_t x_0\rangle = \langle x_0, \varphi(T)^*\varphi(T)x_0\rangle = \|\varphi(T)x_0\|^2$$

by the functional calculus, i.e. $\widehat{V}\colon C(\sigma(T)) \to H, \varphi \mapsto \varphi(T)x_0$ is a linear isometry with the $L^2(\mathbb{R},\mu)$-norm on $C(\sigma(T))$. Since $C(\sigma(T))$ is dense in $L^2(\mathbb{R},\mu)$ (proposition below), $\widehat{V}$ has a unieq extension to a map $V\colon L^2(\mathbb{R},\mu) \to H$, which is also an isometry. Note: The image of V is closed in H (since V is an isometry). Now for $\varphi(t) = t^n$, $\varphi \in C(\sigma(T))$ and $T^n x_0 = \varphi(T)x_0$, so $T^n x_0 \in \widehat{V}(C(\sigma(T)) \subseteq V(L^2(\mathbb{R},\mu)) \subseteq H$. But since x_0 is cyclic, this implies that V is onto. So V is a unitary operator. For $\varphi \in C(\sigma(T))$,

$$T(V(\varphi)) = T(\varphi(T)x_0) = (T \circ \varphi(T))x_0 = (\widehat{\Phi}(t) \circ \widehat{\Phi}(\varphi))x_0 = \widehat{\Phi}(t\cdot\varphi)x_0 = V(t\cdot\varphi)$$

So $(V^{-1}TV)\varphi = t\cdot\varphi$ for all $\varphi \in C(\sigma(T))$ and hence, by density, also for all $\varphi \in L^2(\mathbb{R},\mu)$. Let $U = V^{-1} = V^*$. ☐

Proposition. *If μ is a finite regular Borel measure on a compact metric or topological space M, then $C(M) \subseteq L^p(\mu)$, $1 \le p < \infty$, is dense (in L^p norm).*

Theorem 2.23 (Multiplication operator version of the spectral theorem)**.** *Every bounded self-adjoint operator is unitarily equivalent to a multiplication operator. More precisely, for any self-adjoint $T \in B(H)$, there exists a measure space (Ω,Σ,μ) (if H is separable, this space is σ-finite), a bounded measurable function $f\colon \Omega \to \mathbb{R}$ and a unitary operator $U\colon H \to L^2(\Omega,\Sigma,\mu)$ such that $(UTU^{-1})\varphi = f\varphi$ μ-a.e. for all $\varphi \in L^2(\Omega,\Sigma,\mu)$.*

Proof (only for H separable). By lemma 2.24 and theorem 2.22 there exist unitary operators $U_i\colon H_i \to L^2(\mathbb{R},\mu_i)$ and bounded measurable functions $f_i\colon \sigma(T_i) \to \mathbb{R}$ such that $(U_iT_iU_i^{-1})(\varphi_i) = f_i\cdot\varphi_i$ μ_i-a.e. for all $\varphi_i \in L^2(\mathbb{R},\mu_i)$. Let Ω be the disjoint union of the sets $\sigma(T_i)$ $(i<N)$. Let $\Sigma = \{A \subseteq \Omega\colon \forall i\colon A\cap\sigma(T_i) \text{ is Borel}\}$ and let $\mu\colon \Sigma \to [0,\infty]$ be given by $\mu(A) = \sum_{i<N}\mu_i(A\cap\sigma(T_i))$ (recall that μ_i is the measure $\mathrm{d}\langle x_iE_\lambda x_i\rangle$, $x_i \in H_i$ the cyclic vector of T_i, and E_λ the spectral measure). Then (!) Σ is a σ-algebra, and μ is a σ-finite measure. Let $f(t) = f_i(t)$ if $t \in \sigma(T_i)$ (really, $f(t,i) = f_i(t)$, $(t,i) \in \Omega$). We write $f = \{f_i\}_{i<N}$ and $\varphi = \{\varphi_i\}_{i<N}$, so $\varphi \in L^2(\Omega,\Sigma,\mu)$. Let $U\colon H \to L^2(\Omega,\Sigma,\mu)$ be defined by $U(\{x_i\}_{i<N}) = \{U_ix_i\}_{i<N}$, then (check!) U is unitary and $(UTU^{-1})\varphi = f\cdot\varphi$. ☐

Note: the measure μ is not unique.

Lemma 2.24. *Let H be a separable Hilbert space and $T \in B(H)$ be self-adjoint. Then there exists a decomposition $H = \bigoplus_2 H_i$ such that for all $i<N$, $T(H_i) \subseteq H_i$ and the restricted operator $T_i = T|_{H_i}\colon H_i \to H_i$ has a cyclic vector $x_i \in H_i$.*

Proof. Let $\mathcal{H}$ be the set of finite or countable families $\{H_i\}_{i\in I}$ of pairwise orthogonal closed subspaces of H, such that $T(H_i) \subseteq H_i$ and $H_i = \overline{\operatorname{span}\{T^nx_i\colon n \ge 0\}}$ for some $x_i \in H_i$. Note $\{\{0\}\} \in \mathcal{H}$, so $\mathcal{H} \neq \emptyset$. $\mathcal{H}$ is partially ordered by inclusion. Let $\mathcal{K}$ be chain in $\mathcal{H}$, i.e. a totally ordered subset of $\mathcal{H}$. Write $k \in \mathcal{K}$ as $k = \{H_{ik}\colon i<N_k\}$. Let $h_0 = \bigcup_{k\in\mathcal{K}} k = \{H_{ik}\colon i<N_k, k\in\mathcal{K}\}$. Then $h_0 \in \mathcal{H}$ since H is separable. Hence, by

Zorn's lemma, $\mathcal{H}$ has a maximal element $h \in \mathcal{H}$. Let $U = \overline{\operatorname{span} \bigcup_{k \in h} k} \subseteq H$. If $U \neq H$, then there would exists $x \in U^{\perp} \setminus \{0\}$ (U is closed). Let $V = \overline{\operatorname{span}\{T^n x \colon n \in \mathbb{N}\}}$, then $T(V) \subseteq V$ and x is a cyclic vector for $T|_V$. Also (!!) $V \perp U$. Hence $h \subseteq h \cup \{V\} \in \mathcal{H}$. Since h is maximal, $h = h \cup \{V\}$. Hence, $V \in h$. Now $x \in V$, so $x \in U$. This is a contradiction to $x \in U^{\perp} \setminus \{0\}$. □

Remark. The spectral theorem can be generalised to normal bounded operators T by using that $T = S_1 + iS_2$ for self-adjoint operators S_1 and S_2 with $[S_1, S_2] = 0$:

$$S_1 = \frac{T + T^*}{2} \qquad S_2 = \frac{T - T^*}{2i}$$

Theorem 2.25 (Spectral theorem for normal bounded operators)**.** *For any bounded, normal operator T there exists a unique spectral measure G (with compact support) on the Borel σ-algebra of $\mathbb{C}$ such that*

$$T = \int_{\sigma(T)} z \, \mathrm{d}G_z$$

The formula

$$f(T) = \int_{\sigma(T)} f(z) \, \mathrm{d}G_z$$

defines the uniquely determined measurable functional calculus. Every normal bounded operator is unitarily equivalent to a multiplication operator defined by multiplication by a bounded measurable (complex-valued) function.

3 Unbounded operators, symmetric operators and quadratic forms

In functional analysis 1 we briefly saw that not all operators are bounded. The typical examples are (partial) differential operators. We do have that $T = \frac{\mathrm{d}}{\mathrm{d}x}$ from $X = (C^1[0,1], \|{-}\|_{C^1})$ to $Y = (C^0[0,1], \|{-}\|_{C^0})$ is bounded, where $\|f\|_{C^1} = \|f\|_\infty + \|f'\|_\infty$ and $\|f\|_{C^0} = \|f\|_\infty$, since $\|Tf\|_{C^0} = \|f'\|_\infty \leq \|f\|_\infty + \|f'\|_\infty = \|f\|_{C^1}$. However, for many practical purposes, the spaces X and Y are not the right thing to study $\frac{\mathrm{d}}{\mathrm{d}x}$, for example, what about self-adjointness? Furthermore, we would like $\frac{\mathrm{d}}{\mathrm{d}x}$ to be defined on a Hilbert space. Note that $(C^1[0,1], \|{-}\|_2)$ is not a Hilbert space. For example $(L^2[0,1], \|{-}\|_2)$ is, but then we have lost the ability to differentiate in a natural way. One can repair this with the theory of distributions, weak derivatives and Sobolev spaces. However even if one can solve these problems, there is no way to end up with a bounded operator $T \colon L^2[0,1] \to L^2[0,1]$ extending the derivative: Let $e_n(x) = e^{inx}$, $x \in [0,1]$. Then $e_n \in C^\infty[0,1] \subseteq C^1[0,1] \subseteq L^2[0,1]$ and $\|e_n\|_2 = 1$, but

$$\|Te_n\|_2^2 = \int_0^1 \left| \frac{\mathrm{d}}{\mathrm{d}x} e_n(x) \right|^2 \mathrm{d}x = \int_0^1 n^2 \left| e^{inx} \right|^2 \mathrm{d}x = n^2.$$

Hence, there cannot exist any constant $C > 0$ such that $\|Te_n\|_2 \leq C\|e_n\|_2$ for all $n \in \mathbb{N}$. This is motivation for the need to study unbounded operators. We will need to study operators defined on subspaces of Hilbert spaces, for example $C^1[0,1] \subseteq L^2[0,1]$. The theory of unbounded operators is less "complete" than that of bounded operators, but it can still be applied to a vast number of interesting problems. From now on, let H be a complex Hilbert space.

Definition 3.1. An *operator* (operator in a Hilbert space) $T\colon D(T) \to H$ is a linear map whose *domain* $D(T) \subseteq H$ is a linear subspace of H. In general, $D(T)$ is not necessarily a closed subspace. T is said to be *densely defined* (in H) if $D(T)$ is dense in H.

An operator $S\colon D(S) \to H$ is called an *extension* of the operator $T\colon D(T) \to H$ if $D(T) \subseteq D(S)$ and $S|_{D(T)} = T$. We will write $T \subset S$. Two operators S and T are equal if and only if $T \subset S$ and $S \subset T$.

An operator $T\colon D(T) \to H$ is called *symmetric* if $\langle y, Tx\rangle = \langle Ty, x\rangle$ all for $x, y \in D(T)$.

Remark. Let $H = L^2[0,1]$ and $T = i\frac{\mathrm{d}}{\mathrm{d}x}\colon D(T) \to H$ with $D(T) = \{x \in C^1[0,1]\colon x(0) = x(1) = 0\}$, and $S = i\frac{\mathrm{d}}{\mathrm{d}x}\colon D(S) \to H$ with $D(S) = \{x \in C^1[0,1]\colon x(0) = x(1)\}$. Then both S and T are symmetric operators in H and $T \subset S$.

Proposition 3.2 (Hellinger-Toeplitz)**.** *Let H be a Hilbert space and $T\colon H \to H$ be a linear operator, i.e. $D(T) = H$, such that $\langle y, Tx\rangle = \langle Ty, x\rangle$ for all $x, y \in H$. Then T is bounded. Hence, T is self-adjoint as a bounded operator.*

We will now define the adjoint of a densely defined operator. Let $T\colon D(T) \to H$ be a densely defined operator and set

$$D(T^*) = \{y \in H\colon x \mapsto \langle y, Tx\rangle \text{ is bounded on } D(T)\}.$$

Then $D(T^*) \subseteq H$ is a linear subspace of H. For $y \in D(T^*)$, $x \mapsto \langle y, Tx\rangle$ is bounded on $D(T)$ and $\overline{D(T)} = H$, hence this continuous linear functional has a unique extension to a bounded linear functional on H. By Riesz' representation theorem, there exists a unique $z \in H$ such that this extension is given by $\langle z, -\rangle$. We define $T^*y = z$. It is easy to see, that the map $T^*\colon D(T^*) \to H$ is linear. Furthermore, by definition $\langle T^*y, x\rangle = \langle y, Tx\rangle$ for all $x \in D(T)$ and $y \in D(T^*)$.

Definition 3.3. The operator T^* described above is called the *adjoint* operator of T. If $T = T^*$ then T is called *self-adjoint*.

Remark.

1. If $T \in B(H)$ then this definition coincides with the old one.
2. $T = T^*$ in particular demands that $D(T) = D(T^*)$.
3. If $T = T^*$ then T is symmetric. In general, the opposite is false and importantly the spectral theorem holds for self-adjoint operators, not in general for symmetric operators.

Lemma 3.4. *If $T\colon D(T) \to H$ is densely defined and symmetric, then $T \subset T^*$. In particular, $D(T^*) \subset H$ is dense, so T^{**} is well-defined.*

Remark. If T is not symmetric, then it may happen, that $D(T^*)$ is not dense. In fact, it is possible that $D(T^*) = 0$.

For the study of self-adjointness, the concept of closed operators is important. This topic has already been introduced in functional analysis 1.

Definition 3.5. Let X, Y be normed spaces, $D \subseteq X$ a subspace, $T\colon D \to Y$ a linear map. T is a *closed* operator iff for all $\{x_n\} \subseteq D$, $x_n \to x \in X$ such that $\{Tx_n\}$ converges, $x \in D$ and $Tx_n \to Tx$.

Defining the graph $\Gamma(T) = \mathrm{gr}(T) = \{(x, Tx)\colon x \in D(T)\} \subseteq X \times Y$, T is closed iff $\mathrm{gr}(T) \subseteq X \times Y$ is a closed subspace in $X \oplus_2 Y$ ($X \times Y$ with the norm $\|(x,y)\| = (\|x\|_X + \|y\|)^{1/2}$).

Theorem 3.6 (Closed graph theorem)**.** *Let X, Y be Banach spaces, $T\colon X \to Y$ closed and linear. Then T is bounded.*

Proposition 3.7. *Let $T\colon D(T) \to H$ be a densely defined operator. Then*

(a) T^* is closed.

(b) If T^* is densely defined (for example if T is symmetric), then $T \subset T^{**}$

(c) Assume T^* is densely defined, $S\colon D(S) \to H$ is closed and $T \subset S$. Then $T^{**} \subset S$.

Proof.

(a) Let $\{y_n\} \in D(T^*)$, $y_n \to y \in H$, $T^* y_n \to z \in H$. We need to prove $y \in D(T^*)$ and $T^* y = z$. Now, for all $x \in D(T)$,

$$\langle y, Tx\rangle = \lim_{n\to\infty} \langle y_n, Tx\rangle = \lim_{n\to\infty} \langle T^* y_n, x\rangle = \langle z, x\rangle$$

Hence, $x \mapsto \langle y, Tx\rangle$ is continuous, so $y \in D(T^*)$ and $T^* y = z$ (since $T^* y$ is the unique vector such that $\langle y, Tx\rangle = \langle Ty, x\rangle$).

(b) Assume T^* is densely defined, i.e. $\overline{D(T^*)} = H$. Let $x \in D(T)$, $y \in D(T^*)$, then $\langle y, Tx\rangle = \langle T^* y, x\rangle$. Hence, $y \mapsto \langle x, T^* y\rangle = \overline{\langle y, Tx\rangle}$ is continuous, so $x \in D(T^{**})$ (hence $D(T) \subset D(T^{**})$) and $\langle x, T^* y\rangle = \overline{\langle T^{**} x, y\rangle}$. Since $x \in D(T)$, this implies $\langle Tx, y\rangle = \langle T^{**} x, y\rangle$ for all $y \in D(T^*)$. Since $\overline{D(T^*)} = H$, this implies $Tx = T^{**} x$. So $T \subset T^{**}$.

(c) The claim follows from $\overline{\mathrm{gr}(T)} = \mathrm{gr}(T^{**})$. First, we prove $\overline{\mathrm{gr}(T)} \subseteq \mathrm{gr}(T^{**})$. By (b), $T \subset T^{**}$, i.e. $\mathrm{gr}(T) \subseteq \mathrm{gr}(T^{**})$, so $\overline{\mathrm{gr}(T)} \subseteq \overline{\mathrm{gr}(T^{**})}$ and by (a), T^{**} is closed, so $\overline{\mathrm{gr}(T^{**})} = \mathrm{gr}(T^{**})$. To show $\mathrm{gr}(T^{**}) \subseteq \overline{\mathrm{gr}(T)}$, it is enough to show that $\mathrm{gr}(T)^\perp \subseteq \mathrm{gr}(T^{**})^\perp$ (and use $(V^\perp)^\perp = \overline{V}$), where $\perp$ means the orthogonal complement wrt. the scalar product on $H \times H$:

$$\langle (u,v), (x,y)\rangle = \langle u, x\rangle + \langle v, y\rangle$$

Let $(u,v) \in \mathrm{gr}(T)^\perp$, i.e. $0 = \langle (u,v), (x, Tx)\rangle = \langle u, x\rangle + \langle v, Tx\rangle$. Hence, $\langle v, Tx\rangle = -\langle u, x\rangle$ for all $x \in D(T)$. So $x \mapsto \langle v, Tx\rangle$ is bounded on $D(T)$. Hence $v \in D(T^*)$ and $T^* v = -u$. Let $(z, T^{**} z) \in \mathrm{gr}(T^{**})$ for $z \in D(T^{**})$, then $\langle (u,v), (z, T^{**} z)\rangle = \langle u, z\rangle + \langle v, T^{**} z\rangle = \langle u, z\rangle + \langle T^* v, z\rangle = \langle u, z\rangle + \langle -u, z\rangle = 0$, so $(u,v) \in \mathrm{gr}(T^{**})^\perp$. $\square$

Remark.

1) (a) shows that only closed operators can be self-adjoint ($T = T^* \Rightarrow T$ is closed).

2) (c) means T^{**} is the smallest closed extension of T — called the *closure* $\overline{T}$ of T ($\overline{T} = T^{**}$).

Corollary 3.8. *Let $T\colon D(T) \to H$ densely defined.*

(a) T is symmetric $\iff T \subset T^*$. In this case, $T \subset T^{**} \subset T^* = T^{***}$. Hence, T^{**} is also symmetric.

(b) T is closed and symmetric $\iff T = T^{**} \subset T^*$.

(c) T is self-adjoint $\iff T = T^{**} = T^*$.

Note: So far, we have for $T\colon D(T) \to H$ densely defined and symmetric: $T \subset T^{**} \subset T^* = T^{***}$ and for T self-adjoint: $T = T^{**} = T^* = T^{***}$. 'Between' these:

Definition 3.9. Let $T\colon D(T) \to H$ be densely defined and symmetric. T is called *essentially self-adjoint* iff $\overline{T}$ is self-adjoint (i.e. T is essentially self-adjoint iff $T \subset T^{**} = T^*$, since $T^* = T^{***}$).

Remark. If T is symmetric, but not self-adjoint, then $T \subset T^*$, $T \neq T^*$. If $T \subset S$, then $T^* \supset S^*$. Hence, the domain of T is too small — enlarging it (i.e. finding a symmetric extension) will diminish the domain of the adjoint — the 'goal' is to find an extension S of T such that the domains 'meet', i.e. $S = S^*$. T may or may not have self-adjoint extensions. An essentially self-adjoint operator has exactly one self-adjoint extension.

We continue to search (!) for criterions for self-adjointness first:

Lemma 3.10. *$T\colon D(T) \to H$ is densely defined. Then*

(a) $\mathrm{N}(T^* \mp i) = \mathrm{R}(T \pm i)^\perp$. *Hence* $\mathrm{N}(T^* \mp i) = \{0\} \Leftrightarrow R(T \pm i) \subseteq H$ *is dense.*

(b) Assume T is symmetric and closed. Then $R(T \pm i) \subseteq H$ is closed.

Proof. Note that $(T \pm i)^* = T^* \mp i$ (use definitions, the sum $T + i$ is defined on $D(T)$!) $\supseteq$: Let $y \in \mathrm{R}(T \pm i)^\perp$. Then $\langle y, (T \pm i)z\rangle = 0$ for all $z \in D(T)$. So $z \mapsto \langle y, Tz\rangle$ is continuous (it equals $\mp\langle y, iz\rangle$). Hence, $y \in D(T^*)$ and $0 = \langle (T^* \mp i)y, z\rangle$ for all $z \in D(T)$. So $(T^* \mp i)y = 0$ (since $D(T) \subseteq H$ is dense). Hence, $y \in \mathrm{N}(T^* \mp i)$. For $\subseteq$ reed the above argument backwards. To prove (b), let T be symmetric and closed. Then $\langle x, Tx\rangle \in \mathbb{R}$ for all $x \in D(T)$. So

$$\|(T \pm i)x\|^2 = \|Tx\|^2 + \|x\|^2 \pm 2\,\mathrm{Re}\langle ix, Tx\rangle = \|Tx\|^2 + \|x\|^2 \mp 2\,\mathrm{Re}(i\langle x, Tx\rangle) \geq \|x\|^2$$

Hence, $(T + i)^{-1}\colon \mathrm{R}(T + i) \to D(T)$ exists and is bounded. Let $\{x_n\} \subseteq D(T)$ such that $(T + i)x_n \to y \in \overline{\mathrm{R}(T + i)}$. Then $\{(T + i)x_n\} \subseteq \mathrm{R}(T + i)$ is Cauchy, hence $\{x_n\} \subseteq D(T)$ is Cauchy (it is the image of $\{(T + i)x_n\}$ via the bounded map $(T + i)^{-1}$). Hence, there exists some $x \in H$ such that $x_n \to x$ as $n \to \infty$. Now $Tx_n = (T + i)x_n - ix_n \to y - ix$ as $n \to \infty$, so by closedness of T it follows that $x \in D(T)$ and $Tx = y - ix$. So $y = (T + i)x \in \mathrm{R}(T + i)$. Hence, $\overline{\mathrm{R}(T + i)} = \mathrm{R}(T + i)$, i.e. $\mathrm{R}(T + i)$ is closed. $\square$

Theorem 3.11. *Let $T\colon D(T) \to H$ be densely defined and symmetric. Then the following are equivalent:*

(i) T *is self-adjoint.*

(ii) T *is closed and* $\mathrm{N}(T^* \pm i) = 0$.

(iii) $\mathrm{R}(T \pm i) = H$.

Proof. First assume (i). Then by 3.7(a), T is closed. Let $x \in \mathrm{N}(T^*+i)$, i.e. $(T^*+i)x = 0$. Recall that for symmetric S one has $\|(S+i)x\| \geq \|x\|$ for all $x \in D(S)$. Since T^* is symmetric, $x = 0$. Hence, $\mathrm{N}(T^*+i) = 0$. Similarly, $\mathrm{N}(T^*-i) = 0$. Hence, we get (ii).

Now, assuming (ii), by 3.10(b) (T is closed), $\mathrm{R}(T \pm i) \subseteq H$ is closed. Because we have $\mathrm{N}(T^* \pm i) = 0$, 3.10(a) gives $\mathrm{R}(T \pm i) \subseteq H$ is dense. So $\mathrm{R}(T \pm i) = H$.

Lastly, assume (iii). We have $T \subset T^*$. Hence, it remains to prove that $T^* \subset T$. For this, it is enough to prove that $D(T^*) \subseteq D(T)$ (then, if $x \in D(T^*)$, then $x \in D(T)$ and $Tx = T^*x$ since $T \subset T^*$). Let $y \in D(T^*)$. Then $(T^*-i)y \in H = \mathrm{R}(T-i)$, so there exists some $x \in D(T)$ such that $(T^*-i)y = (T-i)x$. But $T \subset T^*$, and $x \in D(T)$, so $Tx = T^*x$. Hence, $(T^*-i)(y-x) = 0$. But $\mathrm{R}(T \pm i) = H$, hence dense, so, by 3.10(a), $\mathrm{N}(T^* \pm i) = 0$. Hence, $y = x \in D(T)$. So $D(T^*) \subseteq D(T)$, so we get $T = T^*$. ☐

Corollary 3.12. *Let $T\colon D(T) \to H$ be densely defined and symmetric. The following are equivalent:*

(i) T *is essentially self-adjoint.*

(ii) $\mathrm{N}(T^* \pm i) = 0$.

(iii) $\mathrm{R}(T \pm i) \subseteq H$ *is dense.*

Proof. (i) implies $T^{**} = (T^{**})^*$, so by corollary 3.8 $T^{**} = T^{***} = T^*$. By 3.11 (ii) used on T^{**} it follows that $\mathrm{N}(T^* \pm i) = \mathrm{N}(T^{***} \pm i) = 0$.

(ii) implies, by corollary 3.8, that $T^{***} = T^*$, so $\mathrm{N}(T^{***} \pm i) = 0$ and T^{**} is closed. So, by 3.11, used on T^{**} it follows that $T^{**} = (T^{**})^*$, i.e. T is essentially self-adjoint.

The equivalence of (ii) and (iii) is 3.10(a). ☐

Definition 3.13. We call the Hilbert space dimensions $\dim \mathrm{N}(T^* \pm i)$ the *deficiency indices.*

Theorem 3.14. *Let $T\colon D(T) \to H$ be symmetric and densely defined. Then T has a self-adjoint extension if and only if* $\dim \mathrm{N}(T^*+i) = \dim \mathrm{N}(T^*-i)$.

Proof. Assume that S is a self-adjoint extension of T. Let $V = (S+i)(S-i)^{-1}$ on $D(V) = H$ (this is possible, because $\mathrm{R}(S-i) = H$ by 3.11) and $U = (T+i)(T-i)^{-1}$ on $D(U) = \mathrm{R}(T-i)$. Since $T \subset S$, we have $U \subset V$, and V is unitary. By $\|(S \pm i)x\| \geq \|x\|$, V is injective, and (again by 3.11), $\mathrm{R}(S+i) = H$, so V is surjective. By construction U maps the subspace $\mathrm{R}(T-i)$ onto $\mathrm{R}(T+i)$. Since V is unitary, and $U \subset V$, V maps the orthogonal complement of $\mathrm{R}(T-i)$ onto the orthogonal complement of $\mathrm{R}(T+i)$. But by 3.10(a) $\mathrm{N}(T^* \mp i) = \mathrm{R}(T \pm i)^\perp$, so V maps $\mathrm{N}(T^* + i)$ unitarily onto $\mathrm{N}(T^* - i)$. Hence, the Hilbert space dimensions of $\mathrm{N}(T^* + i)$ and $\mathrm{N}(T^* - i)$ must agree.

For the other direction, assume $\dim N(T^* + i) = \dim N(T^* - i)$ (Hilbert space dimensions). From the proof of 3.10(b) we saw that $\|(T+i)x\|^2 = \|Tx\|^2 + \|x\|^2 = \|(T-i)x\|^2$. So the map $U\colon R(T - i) \to R(T + i), (T - i)x \mapsto (T + i)x$ is well-defined, linear, isometric and surjective. Note that, by 3.10(a), $R(T \pm i)^\perp = N(T^* \mp i)$. So the orthogonal complements of $R(T - i)$ and $R(T + i)$ have the same (Hilbert space) dimension. Hence, we can extend U to a unitary map $V\colon H \to H$ (take an ONB for $R(T - i)$ and $R(T + i)$, and map one to the other). Claim: $V - 1$ is injective. To show this, assume $(V - 1)y = 0$, i.e. $Vy = y$. Then also $V^*y = V^*Vy = y$ or $(V^* - 1)y = 0$. Let now $x \in D(T)$, then

$$2i\langle y, x\rangle = \langle y, (T + i)x - (T - i)x\rangle = \langle y, (V - 1)(T - i)x\rangle = \langle (V^* - 1)y, (T - i)x\rangle = 0$$

Hence, $\langle y, x\rangle = 0$ for all $x \in D(T)$, so (sice $\overline{D(T)} = H$), $y = 0$, so $V - 1$ is injective. We can then define the operator $S\colon R(V - 1) \to H, Vz - z \to i(Vz + z)$ (in other words, $S = i(V + 1)(V - 1)^{-1}$). The goal is to show this S is a self-adjoint extension of T. First we show $T \subset S$. For $x \in D(T)$ we have (see above) $\frac{1}{2i}(V - 1)(T - i)x = x$, i.e. $x \in R(V-1) = D(S)$, and $Sx = i(V-1)(V-1)^{-1}(\frac{1}{2i}(V-1)(T-i)x) = \frac{1}{2}(V+1)(T-i)x = \frac{1}{2}V(T - i)x - \frac{1}{2}(T - i)x = \frac{1}{2}(T + i)x - \frac{1}{2}(T - i)x = Tx$. Secondly, S is symmetric: Let $x \in D(S)$, i.e. $x = Vy - y$ for some $y \in H$, then $\langle x, Sx\rangle = \langle Vy - y, i(V + 1)y\rangle = i(\langle Vy, Vy\rangle + \langle Vy, y\rangle - \langle y, Vy\rangle = \langle y, y\rangle) = i(\langle Vy, y\rangle - \langle y, Vy\rangle) = -2\operatorname{Im}\langle Vy, y\rangle$. Hence, $\langle x, Sx\rangle \in \mathbb{R}$ for all $x \in D(S)$, so S is symmetric. To prove that S is self-adjoint we use theorem 3.11. We need to prove $R(S \pm i) = H$. Note that $S - i = 2i(V - 1)^{-1}$ and $S + i = 2i(V - 1)^{-1}$ on $D(S)$ and so, for all $z \in H$,

$$z = (S - i)(V - 1)\frac{z}{2i} = (S + i)(V - 1)V^*\frac{z}{2i} = R(S \pm i) \qquad \square$$

Definition 3.15. Let $T\colon D(T) \to H$ be densely defined.

(a) $\rho(T) = \{\lambda \in \mathbb{C}\colon T - \lambda I\colon D(T) \to H$ is a bijection and $(T - \lambda I)^{-1} \in B(H)\}$ is called the *resolvent set* of T.

(b) $R\colon \rho(T) \to B(H), \lambda \mapsto R_\lambda = (T - \lambda I)^{-1}$ is the *resolvent map*.

(c) $\sigma(T) = \mathbb{C} \setminus \rho(T)$ is the *spectrum* of T.

(d) If $x \in H \setminus \{0\}$, $\lambda \in \mathbb{C}$ are such that $x \in D(T)$ and $Tx = \lambda x$, then λ is an *eigenvalue* and x an *eigenvector* of T.

Remark.

1. $(T-\lambda I)^{-1}\colon H \to D(T) \subseteq H$; $(T-\lambda I)^{-1}\colon H \to H$ need not be bijective. $(T-\lambda I)^{-1}$ is automatically bounded by the open mapping theorem if T is closed.

2. If T is not closed, then $\sigma(T) = \mathbb{C}$. This is the reason, why we only study the spectral theory of closed operators.

3. Clearly, $\lambda \in \sigma(T)$ if λ is an eigenvalue.

Proposition 3.16. *Let $T\colon D(T) \to H$ be densely defined. Then*

(a) $\rho(T) \subseteq \mathbb{C}$ is open, hence $\sigma(T) \subseteq \mathbb{C}$ is closed.

(b) The resolvent map is analytic and

$$R_\mu - R_\lambda = (\mu - \lambda) R_\mu R_\lambda.$$

Proof. The proof proceeds just as for bounded operators (modulo domains). □

Remark.

1. $\sigma(T)$ need not be compact.
2. $\sigma(T) = \emptyset$ is possible.

4 Spectral theory for unbounded self-adjoint operators

For the spectrum of a self-adjoint (unbounded) operator we first prove:

Proposition 4.1. *Let $T\colon D(T) \to H$ be self-adjoint. Then $\sigma(T) \subseteq \mathbb{R}$.*

Proof. Let $z = \lambda + i\mu \in \mathbb{C} \setminus \mathbb{R}$ (i.e. $\mu \neq 0$). Let $S = T/\mu - \lambda/\mu$ on $D(S) := D(T)$. Then $S = S^*$, in particular (since S is symmetric)

$$\|(T - z)x\|^2 = \|\mu(S - i)x\|^2 = \mu^2 \|(S - i)x\|^2 = \mu^2(\|Sx\|^2 + \|x\|^2) \geq \mu^2 \|x\|^2$$

Hence, $(T - z)^{-1}\colon R(T - z) = R(S - i) \to D(T)$ exists and is bounded (put $y = (T - z)x$, then $\|(T - z)^{-1} y\| = \|x\| \leq \frac{1}{\mu^2} \|(T - z)x\| = \frac{1}{\mu^2}\|y\|^2$). Since $S = S^*$, we have $R(S \pm i) = H$ (theorem 3.11), i.e. $(T - z)^{-1}\colon H \to H$ is a bounded operator $(T - z)^{-1} \in B(H)$, hence $z \in \rho(T)$. □

Theorem 4.2 (Spectral theorem, multipliation operator version)**.** *Let $T\colon D(T) \to H$ be self-adjoint. Then there exists a measure space (Ω, Σ, μ) (if H is separable, it is σ-finite), a measurable function $f\colon \Omega \to \mathbb{R}$ (not necessarily bounded) and a unitary operator $U\colon H \to L^2(\Omega, \mu)$ such that*

(a) $x \in D(T) \iff f \cdot Ux \in L^2(\Omega, \mu)$.

(b) $UTU^*\phi = f\phi =: M_f \phi$ for $\phi \in D(M_f) = \{\phi \in L^2(\Omega, \mu)\colon f\phi \in L^2(\Omega, \mu)\}$.

Proof. We know from the previous theorem that $\sigma(T) \subseteq \mathbb{R}$, so $\pm i \in \rho(T)$. Hence, $R = (T + i)^{-1} = R_{-i} \in B(H)$ is well-defined. Our aim is to prove R is normal and then use the multiplication operator version of the spectral theorem for bounded, normal (!!) operators (discussed briefly in theorem 2.25 and problem 37). Let $z_1, z_2 \in H$. Since T is self-adjoint, theorem 3.11 gives that $T + i$ and $T - i$ are surjective. Hence, there exist $x, y \in D(T)$ such that $z_1 = (T + i)x$, $z_2 = (T - i)y$. Then

$$\langle z_2, R_{-i} z_1 \rangle = \langle (T - i)y, x \rangle = \langle (T^* - i)y, x \rangle = \langle y, (T + i)x \rangle = \langle (T - i)^{-1} z_2, z_1 \rangle$$

so $R_{-i}^* = (T-i)^{-1} = R_i$. Hence, by 3.16(b), $R_{-i}R_{-i}^* = R_{-i}R_i = \frac{1}{-2i}(R_{-i}-R_i) = \frac{1}{2i}(R_i - R_{-i}) = R_iR_{-i} = R_{-i}^*R_{-i}$. Therefore, by 2.25, we have $URU^* = M_g$ (multiplication operator) for a bounded, measurable function $g\colon \Omega \to \mathbb{C}$ and a unitary operator $U\colon H \to L^2(\Omega, \mu)$. We need to construct a multiplication operator representation of T out of this. Note that $(\tau + i)^{-1} = \gamma \Leftrightarrow \tau = \gamma^{-1} - i$ for all $\tau, \gamma \in \mathbb{C}$, $\tau \neq -i$, $\gamma \neq 0$. We shall use this on T and R: Let $f(\omega) = g(\omega)^{-1} - i$ for almost every $\omega \in \Omega$. Note: R_{-i} is injective (but maybe not surjective, $R_{-i}\colon H \to D(T)$), hence M_g is injective. Hence, $\{\omega \in \Omega\colon g(\omega) = 0\}$ is of measure 0 (otherwise, one can construct a function $\varphi \neq 0$ such that $M_g\varphi = 0$).

Now let $x \in D(T)$. Since $R_{-i}\colon H \to H$ has range equal to $D(T)$, we have $x = R_{-i}y$ for some $y \in H$, hence $Ux = UR_{-i}y = g \cdot Uy$. Also, $f \cdot Ux = fg \cdot Uy$. Now $(fg)(\omega) = 1 - ig(\omega)$ for almost every $\omega \in \Omega$, and $g \in L^\infty(\Omega, \mu)$, so fg is essentially bounded. Also, $Uy \in L^2(\Omega, \mu)$, hence $fg \cdot Uy \in L^2(\Omega, \mu)$, i.e. $f \cdot Ux \in L^2(\Omega, \mu)$. On the other hand, assume (for $x \in H$) that $f \cdot Ux \in L^2(\Omega, \mu)$. Since $Ux \in L^2(\Omega, \mu)$, $i \cdot Ux \in L^2(\Omega, \mu)$, hence $(f + i) \cdot Ux \in L^2(\Omega, \mu)$. Since $U\colon H \to L^2(\Omega, \mu)$ is a bijection, there exists some $y \in H$ such that $Uy = (f + i) \cdot Ux$. Hence, $g \cdot Uy = g(f + i) \cdot Ux = Ux$. So $x = U^*UX = U^*(g \cdot Uy) = (U^*M_gU)y = R_{-i}y \in D(T)$.

By (a), $D(M_f) = U(D(T))$, and (as above), for $x \in D(T)$ there exists some $y \in H$ such that $x = R_{-i}y$, so $(T + i)x = y$, i.e. $Tx = y - ix$. Hence, $UTx = Uy - i \cdot Ux = g^{-1} \cdot Ux - i \cdot Ux = f \cdot Ux = M_fUx$, i.e. $UTU^*\varphi = f \cdot \varphi$. Since T is symmetric, also M_f is symmetric, so f is real. □

Theorem 4.3 (Spectral decomposition of self-adjoint operators)**.** *Let $T\colon D(T) \to H$ be self-adjoint. Then there exists a unique spectral measure $A \mapsto E_A$ such that*

$$\langle y, Tx\rangle = \int_{\mathbb{R}} \lambda \, \mathrm{d}\langle y, E_\lambda x\rangle$$

for all $x \in D(T)$ and $y \in H$. If $h\colon \mathbb{R} \to \mathbb{C}$ is measurable and

$$D_h = \Big\{x \in H\colon \int_{\mathbb{R}} |h(\lambda)|^2 \, \mathrm{d}\langle x, E_\lambda x\rangle < \infty\Big\}$$

then

$$\langle y, h(T)x\rangle = \int_{\mathbb{R}} h(\lambda) \, \mathrm{d}\langle y, E_\lambda x\rangle$$

for $x \in D_h$ and $y \in H$ defines a normal operator $h(T)\colon D_h \to H$, which is self-adjoint iff h is real-valued. Note that $\mathrm{d}\langle x, E_\lambda x\rangle$ is given as $A \mapsto \langle x, E_A x\rangle = \|E_A x\|^2 \leq \|x\|^2 < \infty$, so if h is also bounded, then $D_h = H$.

Proof. By 4.2, there exists a measure space (Ω, Σ, μ), a measurable function $f\colon \Omega \to \mathbb{R}$ and a unitary operator $U\colon H \to L^2(\Omega, \mu)$ such that $UTU^*\varphi = f \cdot \varphi = M_f\varphi$ μ-almost everywhere for $\varphi \in D(M_f) = \{\varphi \in L^2(\Omega, \mu)\colon f \cdot \varphi \in L^2(\Omega, \mu)\}$. Let $h \in \mathcal{M}_\mathrm{b}(\mathbb{R})$. Then define the operator $h(M_f) := M_{h \circ f}$, i.e. $(h(M_f)\varphi)(\omega) = h(f(\omega))\varphi(\omega)$ μ-almost everywhere. Note that $h \circ f\colon \Omega \to \mathbb{C}$ is bounded, hence $h(M_f) \in B(L^2(\Omega, \mu))$. Also, $h(M_f)$ is normal.

The map $\mathcal{M}_{\mathrm{b}}(\mathbb{R}) \to B(L^2(\Omega,\mu)), h \mapsto h(M_f)$ is linear, bounded and multiplicative: Let $h, g \in \mathcal{M}_{\mathrm{b}}(\mathbb{R})$ and $\alpha \in \mathbb{C}$. Then $((\alpha h + g)(M_f)\varphi)(\omega) = \alpha h(f(\omega))\varphi(\omega) + g(f(\omega))\varphi(\omega) = (\alpha h(M_f)\varphi + g(M_f)\varphi)(\omega)$ and similarly we have $(h(M_f)g(M_f))(\varphi)(\omega) = h(f(\omega))g(f(\omega))\varphi(\omega) = (hg)(f(\omega))\varphi(\omega) = (hg)(M_f)(\varphi)(\omega)$. Additionally, one has, for any $\varphi \in L^2(\Omega,\mu)$,

$$\|h(M_f)\varphi\|^2_{L^2(\Omega,\mu)} = \int_\Omega |h(f(\omega))\varphi(\omega)|^2 \,\mathrm{d}\mu(\omega) \le \int_\Omega \|h\|^2_\infty |\varphi(\omega)|^2 \,\mathrm{d}\mu(\omega) \le \|h\|^2_\infty \|\varphi\|^2_{L^2(\Omega,\mu)},$$

hence $\|h(M_f)\|_{B(L^2(\Omega,\mu))} \le \|h\|_\infty$.

Hence, in particular, for any Borel set $A \subseteq \mathbb{R}$ the characteristic function χ_A is measurable and bounded. So we define $F_A = \chi_A(M_f) = M_{\chi_A \circ f} = M_{\chi_{f^{-1}(A)}}$. Then $A \mapsto F_A$ defines a spectral measure (however, in general, F does not have compact support): Clearly $F_A^* = F_A = F_A^2$ and we have $F_\emptyset = M_{\chi_\emptyset \circ f} = 0$ and $F_\mathbb{R} = M_{\chi_\mathbb{R} \circ f} = \mathrm{id}$. Let $A_1, A_2, \ldots$ be disjoint Borel sets in $\mathbb{R}$, and let $\varphi \in L^2(\Omega,\mu)$. Then

$$\sum_{n=1}^\infty (F_{A_n}\varphi)(\omega) = \sum_{n=1}^\infty (M_{\chi_{A_n} \circ f}\varphi)(\omega) = \sum_{n=1}^\infty (\chi_{A_n}(f(\omega))\varphi(\omega)$$

Since $A_n \cap A_m = \emptyset$ for $n \neq m$, there exists at most one $n_0 \in \mathbb{N}$ such that $f(\omega) \in A_{n_0}$. In any case

$$\sum_{n=1}^\infty \chi_{A_n}(f(\omega)) = \chi_{\bigcup_{n=1}^\infty A_n}(f(\omega))$$

for all ω. Hence

$$\sum_{n=1}^\infty F_{A_n}\varphi = F_{\bigcup_{n=1}^\infty A_n}\varphi$$

for all $\varphi \in L^2(\Omega,\mu)$. It follows that

$$h(M_f) = \int_\mathbb{R} h(\lambda) \,\mathrm{d}F_\lambda$$

for all $h \in \mathcal{M}_{\mathrm{b}}(\mathbb{R})$ and that this uniquely determines F, i.e. there is only one spectral measure such that this equation holds. Now, let $E_A = U^* F_A U$. Then $A \mapsto E_A$ defines a spectral measure, since U is unitary. We can now, for $h \in \mathcal{M}_{\mathrm{b}}(\mathbb{R})$ define

$$h(T) := \int_\mathbb{R} h(\lambda) \,\mathrm{d}E_\lambda.$$

The map $\mathcal{M}_{\mathrm{b}}(\mathbb{R}) \to B(H), h \mapsto h(T)$ is linear and bounded, in fact one has the bound $\|h(T)\|_{B(H)} \le \|h\|_\infty$. Hence, this defines $h(T)$ for $h \in \mathcal{M}_{\mathrm{b}}(\mathbb{R})$. Note that, for all $h \in \mathcal{M}_{\mathrm{b}}(\mathbb{R})$, $h(T) = U^* h(M_f) U$, since this is true for $h = \chi_A$:

$$h(T) = \int_\mathbb{R} \chi_A \,\mathrm{d}E = E_A = U^* F_A U = U^* \chi_A(M_f) U = U^* h(M_f) U,$$

and the general case follows by linearity and continuity. The map $h \mapsto h(T)$ has the properties of a measurable functional calculus.

Let now $h\colon \mathbb{R} \to \mathbb{C}$ be measurable (but not necessarily bounded). For $x \in H$, $\mathrm{d}\langle x, E_\lambda x\rangle$ is a positive (finite) measure: $A \mapsto \langle x, E_A x\rangle = \langle x, E_A^2 x\rangle = \langle E_A^* x, E_A x\rangle = \langle E_A x, E_A x\rangle = \|E_A x\|^2$. Let

$$D_h := \left\{ x \in H\colon \int_{\mathbb{R}} |h(\lambda)|^2 \,\mathrm{d}\langle x, E_\lambda x\rangle < \infty \right\}.$$

If h is bounded, then $D_h = H$. We claim

(1) $D_h \subseteq H$ is dense.

(2) The integral $\int h(\lambda)\,\mathrm{d}\langle y, E_\lambda x\rangle$ exists for all $x \in D_h$, $y \in H$ (here, $\mathrm{d}\langle y, E_\lambda x\rangle$ is a complex measure).

(3) Therefore (!) there exists $h(T)x \in H$ such that

$$\langle y, h(T)x\rangle = \int_{\mathbb{R}} \mathrm{d}\langle y, E_\lambda x\rangle \qquad \forall y \in H$$

This defines the map $D_h \ni x \mapsto h(T)x \in H$.

To prove this, we transport things to the concrete setting on $L^2(\Omega, \mu)$ via the unitary operator $U\colon H \to L^2(\Omega, \mu)$. Let $x, y \in H$ and write $\varphi = Ux$, $\psi = Uy$. Then

$$\begin{aligned}\langle y, E_A x\rangle &= \langle U^*\psi, E_A U^*\varphi\rangle = \langle \psi, U E_A U^* \varphi\rangle = \langle \psi, F_A \varphi\rangle = \langle \psi, \chi_A(M_f)\varphi\rangle = \\ &= \langle \psi, M_{\chi_{f^{-1}(A)}}\varphi\rangle = \int_\Omega \overline{\psi} M_{\chi_{f^{-1}(A)}} \varphi \,\mathrm{d}\mu = \int_\Omega \overline{\psi} \chi_{f^{-1}(A)} \varphi \,\mathrm{d}\mu = \int_{f^{-1}(A)} \overline{\psi}\varphi \,\mathrm{d}\mu\end{aligned}$$

Let $\nu(B) := \int_B \overline{\psi}\varphi \,\mathrm{d}\mu$ ($B \in \Sigma$). Then

$$\langle y, E_A x\rangle = \int_{f^{-1}(A)} \overline{\psi}\varphi \,\mathrm{d}\mu = \nu(f^{-1}(A)) \qquad A \text{ Borel}$$

That is, the measure $\mathrm{d}\langle y, E_\lambda x\rangle$ is the pushforward $f_*\nu$ of ν via f. Hence, by the transformation theorem:

$$\int_{\mathbb{R}} g(\lambda)\,\mathrm{d}\langle y, E_\lambda x\rangle = \int_\Omega (g \circ f)\,\mathrm{d}\nu = \int_\Omega (g \circ f)\overline{\psi}\varphi \,\mathrm{d}\mu \tag{$*$}$$

for $g\colon \mathbb{R} \to \mathbb{C}$ integrable. Using this we get that $x \in D_h$ iff $\varphi = Ux$ satisfies

$$\int_\Omega (|h|^2 \circ f)\overline{\psi}\varphi \,\mathrm{d}\mu = \int_\Omega |h \circ f|^2 |\varphi|^2 \,\mathrm{d}\mu < \infty$$

By exercise 47(i) the set (in $L^2(\Omega, \mu)$) of such φ's is dense in $L^2(\Omega, \mu)$, hence (since U is unitary) $D_h \subseteq H$ is also dense. Note: For $h\colon \mathbb{R} \to \mathbb{C}$, $x \in D_h$, $y \in H$, we have

$$\left| \int_{\mathbb{R}} h(\lambda)\,\mathrm{d}\langle y, E_\lambda x\rangle \right| \le \int_{\mathbb{R}} |h(\lambda)|\,\mathrm{d}\,|\langle y, E_\lambda x\rangle|$$

where $d\,|\langle y, E_\lambda x\rangle|$ is the variation (measure) of the complex measure $\mathrm{d}\langle y, E_\lambda x\rangle$. (i.e., the 'smallest' positive finite measure $|\rho|$ for a measure ρ such that $|\rho(A)| \le |\rho|(A)$ for all measurable A. Since $|\langle y, E_A x\rangle|$ is a positive measure, it is the variation of $\mathrm{d}\langle y, E_A x\rangle$). Using the analogue of $(*)$ on $\mathrm{d}\,|\langle y, E_\lambda x\rangle|$ we get

$$\int_{\mathbb{R}} |h(\lambda)|\,\mathrm{d}|\langle y, E_\lambda x\rangle| = \int_\Omega |h\circ f||\overline{\psi}\varphi|\,\mathrm{d}\mu \le \left(\int_\Omega |h\circ f|^2|\varphi|^2\right)^{1/2}\left(\int_\Omega |\psi|^2\,\mathrm{d}\mu\right)^{1/2} =$$
$$= \left(\int_\Omega |h\circ f|^2|\varphi|^2\,\mathrm{d}\mu\right)^{1/2}\|\psi\|_{L^2} = \left(\int_\Omega |h\circ f|^2|\varphi|^2\,\mathrm{d}\mu\right)^{1/2}\|y\|_H$$

Hence, since $x \in D_h \Leftrightarrow (h\circ f)\varphi \in L^2(\Omega,\mu)$, we have that, for $x \in D_h$ and $y \in H$,

$$\left|\int_{\mathbb{R}} h(\lambda)\,\mathrm{d}\langle y, E_\lambda x\rangle\right| \le C\|y\|_H \qquad C = \left(\int_\Omega |h\circ f|^2|\varphi|^2\,\mathrm{d}\mu\right)^{1/2} < \infty$$

so that $y \mapsto \int_\Omega h(\lambda)\,\mathrm{d}\langle y, E_\lambda x\rangle$ is a (anti-)linear, bounded functional on H. Hence, by Riesz-Fischer, this functional is given by the scalar product with a (unique) element $z \in H$. Denote this element by $h(T)x := z$. Then, for all $x \in D_h$ and $y \in H$,

$$\langle y, h(T)x\rangle = \int_{\mathbb{R}} h(\lambda)\,\mathrm{d}\langle y, E_\lambda x\rangle \qquad \forall x \in D_h\ \forall y \in H \tag{$\square$}$$

This defines a map $h(T)\colon D_h \to H, x \mapsto h(T)x$ (which is well-defined by the uniqueness of the z above). This is clearly (!) linear and we write

$$h(T) = \int_{\mathbb{R}} h(\lambda)\,\mathrm{d}E_\lambda$$

but this only holds in the sense of ($\square$) above. Let $h\colon \mathbb{R} \to \mathbb{R}$ be the function $h(t) = t$. Let, as before, $x \in D_h$, $y \in H$, and write $\varphi = Ux$, $\psi = Uy$. Note $(h\circ f)(\omega) = f(\omega)$ for all $\omega \in \Omega$. Hence,

$$\int_{\mathbb{R}} h(\lambda)\,\mathrm{d}\langle y, E_\lambda x\rangle = \int_{\mathbb{R}} \lambda\,\mathrm{d}\langle y, E_\lambda x\rangle = \int_\Omega (h\circ f)\overline{\psi}\varphi\,\mathrm{d}\mu = \int_\Omega f\overline{\psi}\varphi\,\mathrm{d}\mu$$

$$= \int_\Omega (M_f\varphi)\,\mathrm{d}\mu = \langle \psi, M_f\varphi\rangle_{L^2} = \langle Uy, M_f Ux\rangle_{L^2} = \langle y, U^* M_f Ux\rangle_H = \langle y, Tx\rangle_H$$

So $D_h = D(T)$ and

$$\langle y, Tx\rangle_H = \int_{\mathbb{R}} \lambda\,\mathrm{d}\langle y, E_\lambda x\rangle \qquad x \in D(T)\ y \in H \tag{$\square\square$}$$

i.e. $T = \int_{\mathbb{R}} \lambda\,\mathrm{d}E_\lambda$ in the weak sense of ($\square\square$). Similarly, of course, $\langle y, h(T)x\rangle = \langle y, U^* M_{h\circ f} Ux\rangle$, $x \in D_h = D(h(T))$, $y \in H$.

That $h(T)$ is self-adjoint if $h\colon \mathbb{R} \to \mathbb{R}$ (h is real), follows from this last formula and exercise 47(ii). $\square$

Corollary 4.4. *Let $T\colon D(T) \to H$, $T^* = T$. Then there is a unique map $\widehat{\Phi}\colon \mathcal{M}_{\mathrm{b}}(\mathbb{R}) \to B(H)$ such that*

(i) $\widehat{\Phi}$ is an involutive algebra-homomorphism.

(ii) $\widehat{\Phi}$ *is continuous. In fact,* $\|h(T)\|_{B(H)} \leq \|h\|_\infty$.

(iii) If $\{h_n\}_{n\in\mathbb{N}} \subseteq \mathcal{M}_{\mathrm{b}}(\mathbb{R})$, $\sup_{n\in\mathbb{N}} \|h_n\|_\infty \leq C < \infty$ *and* $h_n(t) \to h(t)$ *as* $n \to \infty$ *for all* $t \in \mathbb{R}$*, then* $h_n(T)x \to h(T)x$ *as* $n \to \infty$ *for all* $x \in H$.

Corollary 4.5. *Let* $T\colon D(T) \to H$ *be self-adjoint.*

(a) If $x \in D_h$ *(*$h\colon \mathbb{R} \to \mathbb{C}$ *is measurable), then* $\|h(T)x\|_H^2 = \int |h(\lambda)|^2 \,\mathrm{d}E_\lambda$.

(b) If $\{h_n\}_{n\in\mathbb{N}} \subseteq \mathcal{M}_{\mathrm{b}}(\mathbb{R})$ *with*

(i) $h_n(t) \to t$ *as* $n \to \infty$ *for all* $t \in \mathbb{R}$ *(a.e.)*

(ii) $|h_n(t)| \leq |t|$ *for all* $t \in \mathbb{R}$, $n \in \mathbb{N}$

Then, for any $x \in D(T)$, $h_n(T) \to Tx$ *as* $n \to \infty$.

Proof.

(a) Assume first that $h\colon \mathbb{R} \to \mathbb{C}$ is measurable and bounded. Then

$$\begin{aligned}\|h(T)x\|^2 &= \langle h(T)x, h(T)x\rangle = \langle x, h(T)^*h(T)x\rangle = \\ &= \left\langle x, |h|^2(T)x \right\rangle = \int_{\mathbb{R}} |h(\lambda)|^2 \,\mathrm{d}\langle x, E_\lambda x\rangle\end{aligned}$$

for all $x \in H$. In general, i.e. for some measurable but not necessarily bounded $h\colon \mathbb{R} \to \mathbb{C}$, we proved that for all $x \in D_h$ and $y \in H$,

$$\left|\int_{\mathbb{R}} h(\lambda) \,\mathrm{d}\langle y, E_\lambda x\rangle\right| \leq C\|y\|_H,$$

with

$$C = \left(\int_{\mathbb{R}} |h \circ f|^2 |\varphi|^2 \,\mathrm{d}\mu\right)^{1/2}$$

where $U\colon L^2(\mu)$ is such that $UTU^* = M_f$ and $\varphi = Ux$, so that

$$y \mapsto \int_{\mathbb{R}} h(\lambda) \,\mathrm{d}\langle y, E_\lambda x\rangle$$

defines a bounded, antilinear functional on H, which hence is equal to the scalar product with some vector $h(T)x \in H$ and $\|h(T)x\| \leq C$. So we have

$$\|h(T)x\|^2 \leq \int_{\mathbb{R}} |h \circ f|^2 |\varphi|^2 \,\mathrm{d}\mu = \int_{\mathbb{R}} |h(\lambda)|^2 \,\mathrm{d}\langle x, E_\lambda x\rangle.$$

Set now (for h unbounded), for $n \in \mathbb{N}$

$$h_n(t) = \begin{cases} h(t) & |h(t)| \leq n \\ 0 & |h(t)| > n \end{cases}$$

Note that $h_n(t) \to h(t)$ as $n \to \infty$ for all $t \in \mathbb{R}$. Then $h_n \in \mathcal{M}_{\mathrm{b}}(\mathbb{R})$ and $D_h = D_{h-h_n}$ and, for $x \in D_h$,

$$\|(h(T) - h_n(T))x\|^2 = \|(h - h_n)(T)x\|^2 \leq \int_{\mathbb{R}} |(h - h_n)(\lambda)|^2 \,\mathrm{d}\langle x, E_\lambda x\rangle \xrightarrow{n\to\infty} 0$$

by dominated convergence, since

$$|h_n(\lambda) - h(\lambda)|^2 \le (|h_n(\lambda)| + |h(\lambda)|)^2 \le 3|h_n(\lambda)|^2 + 3|h(\lambda)|^2 \le 6|h(\lambda)|^2$$

and $h \in L^2(\mathrm{d}\langle x, E_\lambda x\rangle)$. Hence, $h_n(T)x \to h(T)x$ as $n \to \infty$ for all $x \in D_h$. Also,

$$\int_{\mathbb{R}} |h_n(\lambda)|^2 \,\mathrm{d}\langle x, E_\lambda x\rangle \xrightarrow{n\to\infty} \int_{\mathbb{R}} |h(\lambda)|^2 \,\mathrm{d}\langle x, E_\lambda x\rangle.$$

Since

$$\|h_n(T)x\|^2 = \int_{\mathbb{R}} |h_n(\lambda)|^2 \,\mathrm{d}\langle x, E_\lambda x\rangle$$

for all $x \in D_h$, this implies

$$\|h(T)x\|^2 = \int_{\mathbb{R}} |h(\lambda)|^2 \,\mathrm{d}\langle x, E_\lambda x\rangle.$$

The proof of (b) uses the same method. □

There are other approaches to the spectral theorem. For example (à la Teschl) one studies spectral measures (i.e. integrate with respect to them) and proves that this implies a measurable functional calculus. But one gets back in a different way (i.e. given a self-adjoint operator, how to construct the spectral measure). Namely, the resolvent $R_T(z)$ should be

$$R_T(z) = \int_{\mathbb{R}} (\lambda - z)^{-1} \,\mathrm{d}E_\lambda.$$

i.e. for $x, y \in H$

$$\langle y, R_T(z)x\rangle = \int_{\mathbb{R}} \frac{1}{\lambda - z} \,\mathrm{d}\langle y, E_\lambda x\rangle.$$

So one can start by

$$F_x(z) = \langle x, R_T(z)x\rangle = \int_{\mathbb{R}} \frac{1}{\lambda - z} \,\mathrm{d}\langle x, E_\lambda x\rangle.$$

For a Borel measure μ, the map

$$F_\mu(z) := \int_{\mathbb{R}} \frac{1}{\lambda - z} \,\mathrm{d}\mu(\lambda)$$

is called the *Borel-transform* of μ. It turns out that $F_x \colon \mathbb{H} \to \mathbb{H}$ is holomorphic. Such functions have been studied to great extend, they are called Herglatz or Nevanlinna functions. One can reconstruct the measure μ from $F_\mu(z)$ by the Stieltjes inversion formula:

$$\mu(\lambda) = \lim_{\delta\searrow 0} \lim_{\varepsilon\searrow 0} \frac{1}{\pi} \int_{-\infty}^{\lambda+\delta} \mathrm{Im}(F_\mu(t + i\varepsilon)) \,\mathrm{d}t.$$

So if $F_x(z) = \langle x, R_t T(z)x\rangle$ is a Herglatz function satisfying $|F_x(z)| \le M/\mathrm{Im}(z)$ for all $z \in \mathbb{H}$ and some $M > 0$, then it is Borel-transform of a unique Borel measure $\mu_{x,x}$, with $\mu_{x,x}(\mathbb{R}) \le M$. So, one lets $F_x(z) = \langle x, R_T(z)x\rangle$ and proves that this is holomorphic, maps $\mathbb{H}$ into $\mathbb{H}$ and satisfies $|F_x(z)| \le \|x\|^2/\mathrm{Im}(z)$. That $F_x \colon \mathbb{H} \to \mathbb{H}$ follows from the first resolvent formula. Then one gets the unique measure $\mu_{x,x}$ by the Stieltjes inversion formula and by polarization one can then get the complex measures $\mu_{x,y}$. One then defines operators (which will be the E_A) by

$$s_A(y,x) = \int_{\mathbb{R}} \chi_A(\lambda)\,\mathrm{d}\mu_{y,x}.$$

By Lax-Milgram there exists a unique operator E_A such that $s_A(y,x) = \langle y, E_A x\rangle$ for all $x, y \in H$. Then one shows that $A \mapsto E_A$ is a projection valued measure and that the corresponding self-adjoint operator

$$\int_{\mathbb{R}} \lambda\,\mathrm{d}E_\lambda$$

is equal to the original operator. This procedure needs a lot of complex analysis.

There is a concept called "Resolution of the identity". One sets $E_\lambda = E_{(-\infty,\lambda]}$. Then $E(\lambda)$ is an orthogonal projection, $E(\lambda_1) \leq E(\lambda_2)$ for $\lambda_1 \leq \lambda_2$, $\lim_{\lambda_n \searrow \lambda} E(\lambda_n)x = E(\lambda)x$ and $\lim_{\lambda\to-\infty} E(\lambda)x = 0$ and $\lim_{\lambda\to\infty} E(\lambda)x = x$. A map $\lambda \mapsto E_\lambda$ satisfying these properties is called a *resolution of the identity.* There is a bijective correspondence between resolutions of the identity and projection-valued measures. One then defines for $f\colon \mathbb{R} \to \mathbb{C}$,

$$\int_{\mathbb{R}} f(\lambda)\,\mathrm{d}\langle x, E(\lambda)x\rangle$$

as the Riemann-Stieltjes integral, i.e. the limit of Riemann sums of the form

$$\sum f(\xi_i)[\langle x, E(t_{i+1})x\rangle - \langle x, E(t_i)x\rangle].$$

5 Banach algebras

As an outlook, we will now discuss the Banach algebra approach to spectral theory.

Definition 5.1. A *complex algebra* is a complex vectorspace A with a multiplication such that for all $x, y, z \in A$ and $\alpha \in C$

(i) $x(yz) = (xy)z$

(ii) $(x+y)z = xz + yz$, $x(y+z) = xy + xz$

(iii) $\alpha(xy) = x(\alpha y)$

Definition 5.2. A *Banach algebra* is a complex algebra with a norm $\|-\|$ and a distinguished unit element e such that it is a Banach space and satisfies

(iv) $\|xy\| \leq \|x\|\|y\|$ for all $x, y \in A$

(v) $xe = ex = x$ for all $x \in A$.

(vi) $\|e\| = 1$

Remark.

(1) It is not assumed that A is commutative with respect to multiplication.

(2) If A has no unit then there is a natural way of 'adding a unit', making, in a canonical way, a new algebra with unit. For a complex Banach algebra A, let $A_1 = A \times \mathbb{C}$ with the following operations

  (i) $(a, \alpha) + (b, \beta) = (a + b, \alpha + \beta)$

  (ii) $\beta(a, \alpha) = (\beta a, \beta\alpha)$

  (iii) $(a, \alpha) \cdot (b, \beta) = (ab + \alpha b + \beta a, \alpha\beta)$

  Define $\|(a, \alpha)\| := \|a\| + |\alpha|$. Then A_1 with this norm and the algebraic operations defined in (i), (ii), (iii) is a Banach algebra with unit $(0, 1)$ and $a \mapsto (a, 0)$ is an isometric embedding of A into A_1.

$\|xy\| \leq \|x\|\|y\|$ implies that multiplication is continuous, i.e. for all $a \in A$, $x \mapsto ax$ is continous.

Example.

(a) Let K be a nonempty compact Hausdorff space, $C(K)$ the set of all complex-valued continuous functions on K. $C(K)$ forms a Banach algebra with pointwise addition and multiplication and the usual sup norm $\|f\|_{C(K)} = \sup_{k \in K} |f(k)| = \max_{k \in K} |f(k)|$. This algebra is commutative.

(b) If K is a finite set, say $K = \{1, \dots, n\}$, then $C(K) \cong \mathbb{C}^n$, with coordinatewise multiplication.

(c) Let X be a Banach space. Then $B(X)$ is a Banach algebra with the operator norm and the identity as unit element. If $\dim X = n < \infty$, then $B(X) \simeq M_n(\mathbb{C})$. If $\dim X > 1$, then $B(X)$ is not commutative.

(d) Let $K \subseteq \mathbb{C}$ be nonempty and compact, and $A \subseteq C(K)$ be the subset of holomorphic functions $K^\circ \to \mathbb{C}$. Then A is a Banach algebra (in the norm of $C(K)$). If $K = \overline{\mathbb{D}} \subseteq \mathbb{C}$ ($\mathbb{D} \subseteq \mathbb{C}$ being the open unit disc), then A is called the disc algebra. (In fact, one can take $K \subseteq \mathbb{C}^n$).

(e) $L^1(\mathbb{R}^n)$ with convolution as multiplication:

$$(f * g)(x) = \int_{\mathbb{R}^n} f(y) g(x - y) \,\mathrm{d}y \qquad f, g \in L^1(\mathbb{R}^n) \Rightarrow f * g \in L^1(\mathbb{R}^n)$$

  is almost a Banach algebra. One can make it a commutative Banach algebra by adding a unit as above, or concretely enlarge $L^1(\mathbb{R}^n)$ to the algebra of all complex Borel measures μ on $\mathbb{R}^n$ of the form $\mathrm{d}\mu = f \,\mathrm{d}\lambda^n + \alpha \,\mathrm{d}\delta$, where $f \in \mathbb{L}^1(\mathbb{R}^n)$, $\alpha \in \mathbb{C}$, λ^n is the n-dimensional Lebesgue measure and δ is the Dirac measure.

(f) Let $M(\mathbb{R}^n)$ be the algebra of all complex Borel measures, with convolution (of measures) as multiplication, and with the total variation as norm. This is a commutative Banach algebra with unit.

(g) Let (X, Ω, μ) be a σ-finite measure space, and $A = L^\infty(X, \Omega, \mu)$, then A is a commutative (abelian) Banach algebra with unit (using pointwise defined operations).

(h) Let X be a Banach space, and let $A = K(X)$ be the space of compact operators on X. Then $K(X)$ is a Banach algebra without a unit — it has no unit iff $\dim X = \infty$. A is and ideal of $B(X)$ — in particular, it is a subalgebra of $B(X)$.

(i) The construction in (e) ($L^1(\mathbb{R}^n)$ with convolution) can be generalized, as soon as one has a measure space which is a group. Let G be a σ-compact locally compact topological group and m a (right) Haar-measure on G. For $f, g \in L^1(G, \mathrm{d}m)$, let

$$(f * g)(x) = \int_G f(xy^{-1})g(y)\,\mathrm{d}m(y)$$

then $f * g \in L^1(G, \mathrm{d}m)$. $L^1(G)$ is abelian iff G is abelian.

Definition 5.3. If A is a Banach algebra, and there is a map $(-)^*\colon A \to A$ satisfying

(i) $(x^*)^* = x$

(ii) $(xy)^* = y^*x^*$

(iii) $(x + y)^* = x^* + y^*$

(iv) $(\lambda x)^* = \overline{\lambda} x^*$

then A is called a *$*$-Banach algebra.* If the norm on A satisfies the $*$-identity $\|x^*x\| = \|x^*\|\|x\|$, then A is called a *C^*-algebra.*

Definition 5.4.

(i) An element $x \in A$ is invertible iff there exists $x^{-1} \in A$ such that $x^{-1}x = xx^{-1} = e$.

(ii) Let $G(A) \subseteq A$ be the set of all invertible elements in A.

(iii) For $x \in A$ the *spectrum* $\sigma(x) \subseteq \mathbb{C}$ is the set of all $\lambda \in \mathbb{C}$ such that $\lambda e - x$ is not invertible (in the algebra A). The complement of $\sigma(x)$ is the resolvent set $\rho(x)$.

(iv) The *spectral radius* of x is $r(x) := \sup\{|\lambda|\colon \lambda \in \sigma(x)\}$.

Theorem 5.5.

(0) The inverse x^{-1} is unique.

(1) $G(A) \subseteq A$ is an open subset and $x \mapsto x^{-1}$ is an homeomorphsm of $G(A)$ onto $G(A)$.

(2) $\sigma(x) \subseteq \mathbb{C}$ is compact and nonempty for all $x \in A$.

(3) $r(x)$ satisfies

$$r(x) = \lim_{n\to\infty} \|x^n\|^{1/n} = \inf_{n\geq 1} \|x^n\|^{1/n}$$

Proposition 5.6. *Let A be a Banach algebra, $x \in A$, $\|x\| < 1$. Then* *(a)* $e - x$ *is invertible.*

(b) $(e-x)^{-1} = \sum_{n=0}^{\infty} x^n$ *(converging in norm).*

Theorem 5.7 (Gelfand-Mazur)**.** *Let A be a Banach algebra such that every nonzero $x \in A$ is invertible. Then $A \cong \mathbb{C}$.*

Remark. If $B \subseteq A$ is a subalgebra and $x \in B$, then x may have no inverse as element in B, but have an inverse as element in A. Hence, $\sigma(y)$ depends on with respect to which algebra it is computed. But, since $r(x) = \lim_{n\to\infty} \|x^n\|^{1/n}$, $r(x)$ does not.

As for $B(X)$, if $x \in A$, where A is some Banach algebra, then $p(x) \in A$ makes sense for any polynomial function $p\colon \mathbb{C} \to \mathbb{C}$. Also, if $f\colon \mathbb{C} \to \mathbb{C}$ is entire, i.e. holomorphic on all of $\mathbb{C}$, one can write f as a everywhere convergent power series

$$f(t) = \sum_{n=0}^{\infty} \alpha_i t^i;$$

then $f(x) := \sum_{n=0}^{\infty} \alpha_i x^i \in A$ makes sense, i.e. is norm convergent. Here, $x^0 = e \in A$. If $f(t) = (\alpha - t)^{-1}$, $\alpha \in \mathbb{C}$, then $f(x) := (\alpha e - x)^{-1}$ makes sense in A, if $\alpha \notin \sigma(x)$. So, if $f\colon \Omega \to \mathbb{C}$ is holomorphic on an open set $\Omega \subseteq \mathbb{C}$ containing $\sigma(x)$ and we put a closed curve around $\sigma(x)$ inside Ω, then by Cauchy's formula for all t inside Γ, in particular for all $t \in \sigma(x)$,

$$f(t) = \frac{1}{2\pi i} \oint_{\Gamma} (z-t)^{-1} f(z) \,\mathrm{d}z.$$

Defining integrals of Banach space valued functions (Bochner integrals), one gets

Lemma 5.8. *For all $x \in A$ one has*

$$\frac{1}{2\pi i} \oint_{\Gamma} (\alpha - z)^n (ze - x)^{-1} \,\mathrm{d}z = (\alpha e - x)^n \in A.$$

Then one defines:

Definition 5.9. Let A be a Banach algebra with unit, let $a \in A$ and let $\Omega \subseteq \mathbb{C}$ be open with $\sigma(a) \subseteq \Omega$ and let Γ be a contour surrounding $\sigma(a)$ in Ω. Let $f\colon \Omega \to \mathbb{C}$ be holomorphic. Then we define

$$\tilde{f}(a) = \frac{1}{2\pi i} \oint_{\Gamma} f(z)(ze-a)^{-1} \,\mathrm{d}z \in A.$$

This turns out to be independent of Γ and Ω, i.e. if $g\colon \Omega_1 \to \mathbb{C}$, $\sigma(a) \subseteq \Omega_1$, Γ_1 a contour surrounding $\sigma(a)$ in Ω_1, and $g|_{\Omega\cap\Omega_1} = f|_{\Omega\cap\Omega_1}$, then $\tilde{f}(a) = \tilde{g}(a)$.

Let $\mathrm{Hol}(a) = \{f\colon \exists \Omega \subseteq \mathbb{C}, \Omega \supseteq \sigma(a), f\colon \Omega \to \mathbb{C} \text{ is holomorphic}\}$ for any $a \in A$. $\mathrm{Hol}(a)$ is an algebra, but in general not a Banach algebra.

Theorem 5.10 (Riesz functional calculus; Dunford-Schwartz)**.** *Let A be a Banach algebra, $a \in A$. Then the map $\widetilde{\Phi}\colon \mathrm{Hol}(a) \to A, f \mapsto \tilde{f}(a)$ is an algebra homomorphism. If $f = 1$, then $\tilde{f}(a) = e \in A$. If $f = \mathrm{id}$, then $\tilde{f}(a) = a$. If $f(z) = \sum_{i=0}^{\infty} \alpha_i z^i$ has radius of convergence strictly larger than $r(a)$, then $f \in Hol(a)$ and $\tilde{f}(a) = \sum_{i=0}^{\infty} \alpha_i a^i$.*

Let $\{f_n\}_{n\in\mathbb{N}}$ be holomorphic on some open $G \subseteq \mathbb{C}$ with $G \supseteq \sigma(a)$, and assume that $f_n(z) \to f(z)$, $n \to \infty$, uniformly on compact subsets of G. Then $\|\tilde{f}_n(a) - \tilde{f}(a)\| \to 0$, $n \to \infty$. The map $\widetilde{\Phi}\colon \mathrm{Hol}(a) \to A, f \mapsto \tilde{f}(a)$ is uniquely determined by these properties.

Definition 5.11. For $\Omega \subseteq \mathbb{C}$, let $\mathrm{H}(\Omega)$ be the algebra of all holomorphic functions on Ω. Let $A_\Omega = \{x \in A \colon \sigma(x) \subseteq \Omega\} \subseteq A$. Note that if $a \in A_\Omega$ and $f \in \mathrm{H}(\Omega)$, then $\tilde{f}(a) \in A$ is well-defined.

Let $\widetilde{\mathrm{H}}(A_\Omega)$ be the set of A-valued functions $\tilde{g}$ with domain A_Ω which arise from $g \in \mathrm{H}(\Omega)$ by the formula

$$\tilde{g}(x) = \frac{1}{2\pi i} \oint_\Gamma f(z)(ze - x)^{-1} \,\mathrm{d}z$$

with some contour Γ surrounding $\sigma(x)$ in Ω.

Theorem 5.12. $\widetilde{\mathrm{H}}(A_\Omega)$ *is a complex algebra and the map* $\mathrm{H}(\Omega) \to \widetilde{\mathrm{H}}(A_\Omega), f \mapsto \tilde{f}$ *is an algebra isomorphism, which is continuous in the following sense: If* $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathrm{H}(\Omega)$, $f_n \to f$ *uniformly on compact subsets of* Ω, *then* $\tilde{f}(x) = \lim_{n\to\infty} \tilde{f}_n(x)$ *in norm for all* $x \in A_\Omega$. *If* $u(t) = t$ *and* $v(t) = 1$ *in* Ω, *then* $\tilde{u}(x) = x$ *and* $\tilde{v}(x) = e$ *for all* $x \in A_\Omega$.

Theorem 5.13. *Let* $x \in A_\Omega$, $f \in \mathrm{H}(\Omega)$. *Then* $\tilde{f}(x) \in A$ *is invertible iff* $f(t) \neq 0$ *for all* $t \in \sigma(x)$, *and* $\sigma(\tilde{f}(x)) = f(\sigma(x))$.

Recall that for a self-adjoint compact operator $T \colon H \to H$ on a complex Hilbert space H, there exists an orthonormal basis $\{e_j\}$ and $\{\lambda_k\} \subseteq \mathbb{C}$ such that

$$Tx = \sum_k \lambda_k \langle e_k, x\rangle e_k = \sum_k \lambda_k e_k(x) e_k$$

for all $x \in H$. More generally one has the singular value decomposition for a compact operator $T \in K(H_1, H_2)$, i.e. there exist orthonormal systems $\{e_i\} \subseteq H_1$, $\{f_i\} \subseteq H_2$ and $s_1 \geq s_2 \geq \cdots \geq 0$, $s_k \to 0$ as $k \to \infty$, such that

$$Tx = \sum_k s_k \langle e_k, x\rangle f_k = \sum_k s_k e_k(x) f_k$$

for all $x \in H$.

Studying how fast the convergence $\lambda_k \to 0$ is for compact operators on a Hilbert space leads to the study of the *Schatten classes* S_p, i.e. those operators for which $\{\lambda_k\}_k \in \ell_p$. For example $f(-i\nabla)g(x) \in S_p$ iff $f, g \in L^p(\mathbb{R}^3)$.

On Banach spaces, one also studies operators $T \in B(X, Y)$ which can be written as $Tx = \sum \alpha_k x_k'(x) y_k$ for $\alpha_k \in \mathbb{C}$, $x_k' \in X'$ and $y_k \in Y$, so called *nuclear operators*.