Bounded operators

Jan Dereziński

Department of Mathematical Methods in Physics
Warsaw University
Hoża 74, 00-682, Warszawa, Poland

Lecture notes, version of Jan. 2007

January 31, 2007

Contents

These lecture notes are devoted to the most basic properties of bounded operators in Banach and Hilbert spaces. We also introduce the notions of Banach algebras. We avoid using more advanced concepts (locally convex topologies, unbounded operators, applications of the Baire category theorem).

1 Vector spaces

Let $\mathbb{K}$ denote the field $\mathbb{C}$ or $\mathbb{R}$.

If the vector space $\mathcal{X}$ over $\mathbb{K}$ is isomorphic to $\mathbb{K}^n$, we say that $\mathcal{X}$ is of a finite dimension and its dimension is n.

If $A \subset \mathcal{X}$, then Span$A$ denotes the set of finite linear combinations of elements of A. Clearly, Span$A$ is a subspace of $\mathcal{X}$.

1.1 Linear operators

Let $L(\mathcal{X},\mathcal{Y})$ denote the set of linear transformations from $\mathcal{X}$ to $\mathcal{Y}$ and $L(\mathcal{X}) := L(\mathcal{X},\mathcal{X})$. For $A \in L(\mathcal{X},\mathcal{Y})$, Ker$A$ denotes the kernel of A and Ran A the range of A. A is injective iff Ker$A = \{0\}$.

If A is bijective, then A^{-1} is linear.

If $x \in \mathcal{X}$, then $|x\rangle$ denotes the operator

$$\mathbb{K} \ni \lambda \mapsto |x\rangle\lambda := \lambda x \in X.$$

Sometimes we will write $|x)$ instead of $|x\rangle$.

If v is a linear functional on $\mathcal{X}$, then its action on $x \in \mathcal{X}$ will be usually denoted by $\langle v|x\rangle$.

1.2 Factor spaces

Let $\mathcal{X}$ be a vector space and $\mathcal{V} \subset \mathcal{X}$ a subspace. We define

$$x \sim y \iff x - y \in \mathcal{V}.$$

Then $\sim$ is an equivalence relation in $\mathcal{X}$ compatible with the addition of vectors and multiplication of vectors by numbers. The set $\mathcal{X}/\sim$ is denoted by $\mathcal{X}/\mathcal{V}$, it has a natural structure of a vector space. The dimension of $\mathcal{X}/\mathcal{V}$ is called the codimension of $\mathcal{V}$. We have the linear map

$$\mathcal{X} \ni x \mapsto x + \mathcal{V} \in \mathcal{X}/\mathcal{V}$$

whose kernel equals $\mathcal{V}$.

1.3 Direct sums

If $\mathcal{X}$ is a vector space and $\mathcal{X}_1, \mathcal{X}_2$ subspaces, then we write $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2$ when $\mathcal{X}_1 \cap \mathcal{X}_2 = \{0\}$ and $\mathcal{X} = \mathcal{X}_1 + \mathcal{X}_2$. We say that $\mathcal{X}$ is an internal direct sum of X_1 and X_2.

$P \in L(\mathcal{X})$ is called an idempotent if $P^2 = P$. If P is an idempotent, then so is $1 - P$.

If $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2$, then there exists a unique idempotent P such that $\operatorname{Ran} P = \mathcal{X}_1$ and $\operatorname{Ker} P = \mathcal{X}_2$. Conversely, if P is an idempotent, then $\mathcal{X} = \operatorname{Ran} P \oplus \operatorname{Ker} P$.

If $\mathcal{X}_1$ and $\mathcal{X}_2$ are vector spaces, then $\mathcal{X}_1 \times \mathcal{X}_2$ has a natural structure of a vector space, and

$$\mathcal{X}_1 \times \mathcal{X}_2 = \mathcal{X}_1 \times \{0\} \oplus \{0\} \times \mathcal{X}_2$$

in the sense of the internal direct sum. We often write $\mathcal{X}_1 \oplus \mathcal{X}_2$ instead of $\mathcal{X}_1 \times \mathcal{X}_2$ and we call this the external direct sum.

If $\mathcal{X}_i$, $i \in I$ is a family of vector spaces, then we write $\underset{i\in I}{\oplus} \mathcal{X}_i$ for the Cartesian product $\underset{i\in I}{\times} \mathcal{X}_i$ equipped with the vector space structure and $\overset{\mathrm{al}}{\underset{i\in I}{\oplus}} \mathcal{X}_i$ for its subspace consisting of finite linear combinations of elements of $\mathcal{X}_i$.

1.4 2×2 matrices

Let $P^0 \in L(\mathcal{X})$ be an idempotent and $P^1 := 1 - P^0$ the complementary idempotent.

If $H \in L(\mathcal{X})$, then full information about H is contained in the matrix $H^{ij} := P^i H P^j$, with $i, j \in \{0, 1\}$. We can write H as

$$H = \begin{bmatrix} H^{00} & H^{01} \\ H^{10} & H^{11} \end{bmatrix}. \tag{1.1}$$

Theorem 1.1 *Suppose that H^{11} is bijective.*

(1) *The following identity, sometimes called the Schur-Frobenius decomposition, is valid:*

$$H = \begin{bmatrix} 1 & H^{01}(H^{11})^{-1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} H^{00} - H^{01}(H^{11})^{-1}H^{10} & 0 \\ 0 & H^{11} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ (H^{11})^{-1}H^{10} & 1 \end{bmatrix}$$

(2) *Set*

$$G_0 = H^{00} - H^{01}(H^{11})^{-1}H^{10}.$$

Then H^{-1} is bijective iff G_0^{-1} is and

$$\begin{aligned} H^{-1} &= \begin{bmatrix} G_0^{-1} & -G_0^{-1}H^{01}(H^{11})^{-1} \\ -(H^{11})^{-1}H^{10}G_0^{-1} & (H^{11})^{-1} + (H^{11})^{-1}H^{10}G_0^{-1}H^{01}(H^{11})^{-1} \end{bmatrix} \\ &= \left(1^{00} - (H^{11})^{-1}H^{10}\right) G_0^{-1} \left(1^{00} - H^{01}(H^{11})^{-1}\right) + (H^{11})^{-1}. \end{aligned}$$

1.5 Convexity

Let $\mathcal{V}$ be a vector space over $\mathcal{K}$, where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$.

$V \subset \mathcal{X}$ is convex iff $x_1, x_2 \in V$, $0 \le t \le 1$ implies $tx_1 + (1-t)x_2 \in V$.

If $\mathcal{V} \subset \mathcal{X}$ is convex and $e \in \mathcal{V}$, then we say that it is an extremal point of $\mathcal{V}$ iff $e_1, e_2 \in \mathcal{V}$ and $e = \frac{1}{2}(e_1 + e_2)$ implies $e = e_1 = e_2$. $\mathrm{E}(\mathcal{V})$ will denote the set of extremal points of $\mathcal{V}$.

$\mathrm{CH}\mathcal{V}$ denotes the convex hull of $\mathcal{V}$, that is

$$\mathrm{CH}\mathcal{V} := \{tx_1 + (1-t)x_2 \; : \; x_1, x_2 \in \mathcal{V}, \quad t \in [0,1]\}.$$

1.6 Absolute convexity

A set $\mathcal{V} \subset \mathcal{X}$ is balanced iff $x \in \mathcal{V}$ and $|\alpha| \leq 1$ implies $\alpha x \in \mathcal{V}$. It is called symmetric iff $x \in \mathcal{V}$ implies $-x \in \mathcal{V}$. If $\mathbb{K} = \mathbb{R}$ then symmetric is synonymous with balanced.

A balanced convex set is called absolutely convex. ACH$\mathcal{V}$ denotes the absolutely convex hull of $\mathcal{V}$, that is

$$\mathrm{ACH}\mathcal{V} := \{\alpha_1 x_1 + \alpha_2 x_2 \ : \ x_1, x_2 \in \mathcal{V}, \ \ |\alpha_1| + |\alpha_2| \leq 1\}.$$

1.7 Cones

$\mathcal{C} \subset \mathcal{X}$ is called a cone iff

(1) $x \in \mathcal{C}$, $\lambda > 0$ implies $\lambda x \in \mathcal{C}$.

(2) $x, y \in \mathcal{C}$ implies $x + y \in \mathcal{C}$.

A cone $\mathcal{C}$ is called pointed iff $\mathcal{C} \cap (-\mathcal{C}) \subset \{0\}$.

It is called generating iff Span$\mathcal{C} = \mathcal{X}$. In the real case it means that $\mathcal{C} - \mathcal{C} = \mathcal{X}$.

If $\mathcal{A} \subset \mathcal{X}$, then Cone($\mathcal{A}$) denotes the smallest cone containing $\mathcal{A}$.

We can introduce the relation

$$x \geq y \quad \Leftrightarrow \quad x - y \in \mathcal{C}.$$

We will write $x > y$ iff $x \geq y$ and $x \neq y$.

Theorem 1.2 (1) $\mathcal{C} = \{x \ : \ x \geq 0\}$.

(2) *$x \geq y$ and $\lambda > 0$ imply $\lambda x \geq \lambda y$.*

(3) *$x_1 \geq y_1$ and $x_2 \geq y_2$ imply $x_1 + x_2 \geq y_1 + y_2$.*

(4) *If the cone is pointed then $x \geq y$ and $y \geq x$ imply $y = x$.*

(5) *If $0 \in \mathcal{C}$, then $x \geq x$*

Thus in the case of a pointed cone containing zero, $\geq$ is an order.

1.8 Hahn-Banach Theorem

Let $\mathcal{X}$ be a real vector space. Let $U \subset \mathcal{X}$. For $x \in \mathcal{X}$, define

$$p_U(x) := \inf\{\lambda > 0 \ : \ \lambda^{-1} x \in U\}.$$

Then $p_U : \mathcal{X} \to [0, \infty]$ is called the supporting functional of U.

We say that $U \subset \mathcal{X}$ is absorbing iff for any $x \in \mathcal{X}$ there exists λ_0 such that for $\lambda > \lambda_0$, we have $x \in \lambda U$.

Theorem 1.3 (1) *$p_U(\lambda x) = \lambda p_U(x)$ for $\lambda > 0$.*

(2) *p_U is finite iff U is absorbing.*

(3) *$U = -U$ (U is symmetric) iff $p_U(x) = p_U(-x)$.*

(4) *$U = \alpha U$ for $|\alpha| = 1$ (U is balanced) iff $p_U(\lambda x) = |\lambda| p_U(x)$.*

(5) *U is convex iff $p_U(x + y) \leq p_U(x) + p_U(y)$.*

(6) *Let $W := p_U^{-1}([0, 1])$. Then $U \subset W$, W is convex and contains 0. Moreover, $p_U = p_W$.*

A function $p : \mathcal{X} \to [0, \infty[$ is called a Minkowski functional iff

$$p(x_1 + x_2) \le p(x_1) + p(x_2), \quad x_1, x_2 \in \mathcal{X},$$

$$p(\lambda x) = \lambda p(x), \quad \lambda > 0.$$

Clearly, if p is a Minkowski functional, then $U := \{x \ : \ p(x) \le 1\}$ is a convex absorbing set containing 0. Moreover, p is equal to the supporting functional of U.

Theorem 1.4 (Hahn-Banach) *Let $\mathcal{X}$ be a real vector space with a Minkowski functional p. Let $\mathcal{Y}$ be a subspace of $\mathcal{X}$ and v a functional on $\mathcal{Y}$ satisfying*

$$\langle v|x\rangle \le p(x), \quad x \in \mathcal{Y}.$$

Then there exists a linear functional $\tilde{v}$ on $\mathcal{X}$ such that

$$\langle \tilde{v}|x\rangle \le p(x), \quad x \in \mathcal{X}.$$

Lemma 1.5 *The Hahn-Banach Theorem is true if* $\dim \mathcal{X}/\mathcal{Y} = 1$

Proof. Let $x_1 \in \mathcal{X}\backslash\mathcal{Y}$, so that $\mathcal{X} = \mathcal{Y} + \mathbb{R}x_1$. Let $\alpha := \sup\limits_{y\in\mathcal{Y}}(v(y) - p(y - x_1))$. We have for $y_1, y_2 \in \mathcal{Y}$,

$$v(y_1) + v(y_2) = v(y_1 + y_2) \le p(y_1 + y_2) \le p(y_1 - x_1) + p(x_1 + y_2).$$

Therefore,

$$\alpha \le p(x_1 + y_2) - v(y_2), \quad y_2 \in \mathcal{Y}, \tag{1.2}$$

In particular, α is finite. We set

$$\tilde{v}(x + ty) := v(y) + t\alpha, \quad y \in \mathcal{Y}, \ t \in \mathbb{R}.$$

Using (1.2) with $y_2 = \frac{y}{t}$, we obtain

$$\tilde{v}(y + tx_1) \le v(y) + t\left(p(x_1 + \frac{y}{t}) - v(\frac{x}{t})\right) = p(y + tx_1).$$

□

Proof of Theorem 1.4 Consider the family $\mathcal{P}$ of spaces equipped with a linear functional $(\mathcal{Y}_i, v_i)$ such that $\mathcal{Y} \subset \mathcal{Y}_i$, $v_i\Big|_{\mathcal{Y}} = v$ and $v_i(x) \le p(x)$, $x \in \mathcal{Y}_i$. We will write $(\mathcal{Y}_1, v_1) \prec (\mathcal{Y}_2, v_2)$ iff $\mathcal{Y}_1 \subset \mathcal{Y}_2$ and v_1 is a restriction of v_2. Let $(\mathcal{Y}_i, v_i)$, $i \in I$ be a linearly ordered subfamily. Then it has an upper bound $(\mathcal{Y}_0, v_0)$ where $\mathcal{Y}_0$ is the union of $\mathcal{Y}_i$ and v_0 is the obvous extension. $\mathcal{P}$ is nonempty, since $(\mathcal{Y}, v)$ belongs to $\mathcal{P}$. Thus by the Kuratowski-Zorn lemma $\mathcal{P}$ has a maximal element $(\mathcal{Y}_{\max}, v_{\max})$. By the previous lemma, $\mathcal{Y}_{\max} = \mathcal{X}$. □

1.9 Hahn-Banach Theorem for seminorms

Let $\mathcal{X}$ be a real or complex vector space. A function $p : \mathcal{X} \to [0, \infty[$ satisfying

$$p(x_1 + x_2) \le p(x_1) + p(x_2), \quad x_1, x_2 \in \mathcal{X},$$

$$p(\lambda x) = |\lambda| p(x), \quad \lambda \in \mathbb{K},$$

is called a seminorm.

Clearly, if If p is a seminorm, then $U := \{x \ : \ p(x) \le 1\}$ is an absolutely convex absorbing set containing 0. Moreover, p is equal to the supporting functional of U.

Theorem 1.4 implies easily

Theorem 1.6 *Let $\mathcal{X}$ be a real or complex vector space with a seminorm p. Let $\mathcal{Y}$ be a subspace of $\mathcal{X}$ and v a functional on $\mathcal{Y}$ satisfying*

$$|\langle v|x\rangle| \leq p(x), \quad x \in \mathcal{Y}.$$

Then there exists a linear functional $\tilde{v}$ on $\mathcal{X}$ such that

$$|\langle \tilde{v}|x\rangle| \leq p(x), \quad x \in \mathcal{X}.$$

2 Algebras

2.1 Algebras

Let $\mathfrak{A}$ be a vector space over $\mathbb{C}$. $\mathfrak{A}$ is called an associative algebra iff it is equipped with a multiplication satisfying

$$\begin{array}{l} A(BC) = (AB)C, \\ A(B+C) = AB + AC, \quad (B+C)A = BA + CA, \\ (\alpha\beta)(AB) = (\alpha A)(\beta B). \end{array}$$

(We will say simply an algebra instead of an associative algebra) $\mathfrak{A}$ is called a commutative algebra iff $A, B \in \mathfrak{A}$ implies $AB = BA$.

$\mathfrak{B} \subset \mathfrak{A}$ is called a subalgebra if it is a linear subspace and $A, B \in \mathfrak{B} \Rightarrow AB \in \mathfrak{B}$. Clearly, a subalgebra is also an algebra.

Let $\mathcal{V}$ be a vector space. Clearly, the set of linear maps in $\mathcal{V}$, denoted by $L(\mathcal{V})$, is an algebra.

A subalgebra of $L(\mathcal{V})$ is called a a concrete algebra.

An identity of an algebra $\mathfrak{A}$ is an element $\mathbb{I} \in \mathfrak{A}$ such that

$$A = \mathbb{I}A = A\mathbb{I}, \quad A \in \mathfrak{A}.$$

Any algebra has at most one identity. In fact, if $\mathbb{I}_1, \mathbb{I}_2$ are identities, then

$$\mathbb{I}_1 = \mathbb{I}_1\mathbb{I}_2 = \mathbb{I}_2.$$

We say that $\mathfrak{A}$ is unital if it possesses an identity. In what follows, for $\lambda \in \mathbb{C}$ we will simply write λ instead of $\lambda\mathbb{I}$.

$P \in \mathfrak{A}$ is called an idempotent iff $P^2 = P$. $P\mathfrak{A}P$ is a subalgebra called the reduced algebra.

If $\mathfrak{A} \subset L(\mathcal{V})$ is a concrete algebra and $E \in \mathfrak{A}$ is its identity, then E is an idempotent in $L(\mathcal{V})$. We can then restrict $\mathfrak{A}$ to $\operatorname{Ran} E$.

An idempotent P is called finite discrete iff $P\mathfrak{A}P$ is finite dimensional. It is called abelian iff $P\mathfrak{A}P$ is commutative.

Fix an algebra $\mathfrak{A}$. Let $\mathfrak{B} \subset \mathfrak{A}$. $\mathrm{Alg}(\mathfrak{B})$ denotes the smallest subalgebra of $\mathfrak{A}$ containing $\mathfrak{B}$. $\mathrm{Alg}(\mathfrak{B})$ will be called the subalgebra generated by $\mathfrak{B}$.

2.2 Commutant

Fix an algebra $\mathfrak{A}$. Let $\mathfrak{B} \subset \mathfrak{A}$.

The relative commutant of $\mathfrak{B}$ in $\mathfrak{A}$ is defined as

$$\mathfrak{B}' := \{A \in \mathfrak{A} \ : \ AB = BA, \ B \in \mathfrak{B}\}$$

Clearly, $\mathfrak{B}'$ is a subalgebra of $\mathfrak{A}$ and $\mathfrak{B}' = \mathrm{Alg}(\mathfrak{B})'$. If $\mathfrak{A}$ contains 1, then so does $\mathfrak{B}'$.

$\mathfrak{B}''$ is a subalgebra of $\mathfrak{A}$ containing $\mathfrak{B}$.

If there is a risk of confusion (it is not clear which $\mathfrak{A}$ we have in mind), we will write $\mathfrak{B}' \cap \mathfrak{A}$ instead of $\mathfrak{B}'$.

The center of $\mathfrak{A}$ is defined as

$$\mathfrak{A} \cap \mathfrak{A}' = \{A \in \mathfrak{A} \ : \ AB = BA, \ B \in \mathfrak{A}\}.$$

2.3 Direct sums

If $\mathfrak{A}_1$, $\mathfrak{A}_2$ are algebras, then we can define $\mathfrak{A}_1 \oplus \mathfrak{A}_2$.

If $\mathfrak{A}$ is an algebra and $P \in \mathfrak{A} \cap \mathfrak{A}'$ is an idempotent, then clearly $P\mathfrak{A} = P\mathfrak{A}P$ is a subalgebra. $\mathfrak{A}$ is naturally isomorphic to $P\mathfrak{A} \oplus (1-P)\mathfrak{A}$.

2.4 Homomorphisms

Let $\mathfrak{A}$, $\mathfrak{B}$ be algebras. A map $\phi : \mathfrak{A} \to \mathfrak{B}$ is called a homomorphism if it is linear and preserves the multiplication, that means it satisfies

(1) $\phi(\lambda A) = \lambda\phi(A)$;

(2) $\phi(A + B) = \phi(A) + \phi(B)$;

(3) $\phi(AB) = \phi(A)\phi(B)$.

A homomorphism of $\mathfrak{A}$ into $L(\mathcal{V})$ is called a representation of $\mathfrak{A}$ in $\mathcal{V}$.

If $\mathfrak{A}$ is a unital algebra and $\phi : \mathfrak{A} \to \mathfrak{B}$ is a homomorphism, then $\phi(1)$ is an idempotent in $\mathfrak{B}$. ϕ is called unital iff

$$\phi(1) = 1.$$

2.5 Ideals

$\mathfrak{B}$ is a left ideal of an algebra $\mathfrak{A}$ iff it is a linear subspace of $\mathfrak{A}$ and $A \in \mathfrak{A}$, $B \in \mathfrak{B} \Rightarrow AB \in \mathfrak{B}$. Similarly we define the right ideal.

If $A \in \mathfrak{A}$, then $\mathfrak{A}A$ is a left ideal.

$\mathfrak{B}$ is called a two-sided ideal if it is a left and right ideal. In what follows we will write an ideal instead of a two-sided ideal.

We say that an ideal $\mathfrak{I}$ is proper iff $\mathfrak{I} \neq \mathfrak{A}$. We say that an ideal $\mathfrak{I}$ is nontrivial iff $\mathfrak{I} \neq \mathfrak{A}$ and $\mathfrak{I} \neq \{0\}$.

Theorem 2.1 *The kernel of a homomorphism is an ideal. If $\mathfrak{I}$ is an ideal of $\mathfrak{A}$, then $\mathfrak{A}/\mathfrak{I}$ has a natural structure of an algebra. The map*

$$\mathfrak{A} \ni A \mapsto A + \mathfrak{I} \in \mathfrak{A}/\mathfrak{I}$$

is a homomorphism whose kernel equals $\mathfrak{I}$.

2.6 Left regular representation

The so-called left regular representation

$$\mathfrak{A} \ni A \mapsto \lambda(A) \in L(\mathfrak{A})$$

is defined by

$$\lambda(A)B := AB, \quad A, B \in \mathfrak{A}.$$

If $\mathfrak{A}$ is unital, then λ is injective. If $\mathfrak{A}$ is not unital, then λ can be extended to a representation

$$\mathfrak{A} \ni A \mapsto \lambda_1(A) \in L(\mathfrak{A}_{\mathbf{1}})$$

in the obvious way, which is injective.

In any case, we see that every algebra is isomorphic to a concrete algebra.

2.7 Invertible elements

Let $\mathfrak{A}$ be an algebra. $A \in \mathfrak{A}$ is left invertible in $\mathfrak{A}$ iff there exists an element $B \in \mathfrak{A}$, called a left inverse of A, such that $BA = 1$. It is called right invertible iff there exists $C \in \mathfrak{A}$ such that $AC = 1$.

Theorem 2.2 *If $\mathfrak{I} \subset \mathfrak{A}$ is a proper left or right ideal, then no elements of $\mathfrak{I}$ are left or right invertible.*

Theorem 2.3 *Let $A \in \mathfrak{A}$. TFAE:*

(1) *A is left and right invertible.*

(2) *There exists a unique $B \in \mathfrak{A}$ such that $AB = BA = 1$*

Proof. Let B, C be a left and right inverse of A. Then

$$B = B1 = BAC = 1C = C.$$

□

If the above conditions are satisfied, then we say that A is invertible, and the element B, called the inverse of A, is denoted A^{-1}

Theorem 2.4 *1. If A is invertible and B is a left or right inverse of A, then $B = A^{-1}$.*

2. If A, B are invertible, then

$$(AB)^{-1} = B^{-1}A^{-1}, \quad A^{-1} - B^{-1} = A^{-1}(B - A)B^{-1}.$$

2.8 Spectrum, resolvent

Let $\mathfrak{A}$ be a unital algebra and $A \in \mathfrak{A}$. The resolvent set of A, denoted $\mathrm{rs}_{\mathfrak{A}}(A)$, is defined as

$$\mathrm{rs}_{\mathfrak{A}}(A) := \{z \in \mathbb{C} \ : \ z - A \text{ is invertible in } \mathfrak{A}\}. \tag{2.3}$$

The spectrum of A, denoted $\mathrm{sp}_{\mathfrak{A}}(A)$ is defined as

$$\mathrm{sp}_{\mathfrak{A}}(A) := \mathbb{C} \setminus \mathrm{rs}_{\mathfrak{A}}(A).$$

If $\mathfrak{A}$ is non-unital, then in (2.3) we demand that $z - A$ is invertible in $\mathfrak{A}_{\mathbf{1}}$.

Theorem 2.5 *1.* $\mathrm{sp}_{\mathfrak{A}}(AB) \cup \{0\} = \mathrm{sp}_{\mathfrak{A}}(BA) \cup \{0\}$.

2. If $\mathfrak{B}$ is a subalgebra of $\mathfrak{A}$ and $B \in \mathfrak{B}$, then $\mathrm{sp}_{\mathfrak{B}}(B) \supset \mathrm{sp}_{\mathfrak{A}}(B)$.

3. If $\pi : \mathfrak{A} \to \mathfrak{B}$ is a unital homomorphism, then $\mathrm{sp}_{\mathfrak{A}}(A) \supset \mathrm{sp}_{\pi(\mathfrak{A})}(\pi(A))$.

Proof. To prove 1., let $z \notin \mathrm{sp}_{\mathfrak{A}}(AB) \cup \{0\}$. Let $C := (z - AB)^{-1}$. Then $z^{-1}(1 + BCA)$ is the inverse of $z - BA$. □

A is nilpotent of degree n iff $A^n = 0$ and $A^{n-1} \neq 0$. A is quasinilpotent iff $\mathrm{sp}_{\mathfrak{A}}(A) = \{0\}$.

Theorem 2.6 (1) *Every nilpotent A is quasinilpotent.*

(2) *If P is an idempotent not equal to 0 or 1, then* $\mathrm{sp}_{\mathfrak{A}}(P) = \{0, 1\}$.

Proof. (1) Let $z \neq 0$. Then $\sum_{j=0}^{\infty} z^{-j-1}A^j$ is a finite sum and is the inverse of $(z - A)$.

(2) We check that for $z \neq 0, 1$, $(z-1)^{-1}P + z^{-1}(1-P)$ is the inverse of $z - P = (z-1)P + z(1-P)$.

□

Theorem 2.7 *Let $A, B \in \mathfrak{A}$, $z \in \mathrm{rs}(A)$ and $AB = BA$. Then*

$$B(z-A)^{-1} = (z-A)^{-1}B.$$

Proof. We have

$$(z-A)B = B(z-A).$$

Then we multiply both sides by $(z-A)^{-1}$. □

If $A \in \mathfrak{A}$, we define

$$\widetilde{\mathrm{Alg}}_{\mathfrak{A}}(A) := \mathrm{Alg}_{\mathfrak{A}}(\{A, (z-A)^{-1} \ : \ z \in \mathrm{rs}_{\mathfrak{A}}(A)\}).$$

Clearly, by Theorem 2.7, $\widetilde{\mathrm{Alg}}_{\mathfrak{A}}(A)$ is commutative and contains 1.

Theorem 2.8 $\widetilde{\mathrm{Alg}}_{\mathfrak{A}}(A)$ *is the smallest among the subalgebras of* $\mathfrak{A}$ *containg* A, 1 *and such that* $\mathrm{sp}_{\mathfrak{A}}(A) = \mathrm{sp}_{\mathcal{C}}(A)$.

2.9 Functional calculus

Let $K \subset \mathbb{C}$. The rational functions with poles outside K form a commutative algebra that we denote $\mathrm{Rat}(K)$. If $f \in \mathrm{Rat}(K)$, then $f(z) = p(z)q(z)^{-1}$, where $p(z)$, $q(z)$ are polynomials and $q(z) = (z-\lambda_1)^{m_1} \cdots (z-\lambda_n)^{m_n}$ with $z_i \notin K$.

Let $\mathfrak{A}$ be a unital algebra, $A \in \mathfrak{A}$ and $f \in \mathrm{Rat}(\mathrm{sp}(A))$. We define

$$f(A) := p(A)(A-z_1)^{-m_1} \cdots (A-z_n)^{-m_n}. \tag{2.4}$$

Note that the commutativity of $\widetilde{\mathrm{Alg}}(A)$ guarantees that (2.4) does not depend on the order of $\lambda_1, \ldots, \lambda_n$.

Theorem 2.9

$$\mathrm{Rat}(\mathrm{sp}(A)) \ni f \mapsto f(A) \in \widetilde{\mathrm{Alg}}(A) \subset \mathfrak{A} \tag{2.5}$$

is a unital homomorphism. Moreover,

(1) *if*

$$\mathrm{Rat}(\mathrm{sp}(A)) \ni f \mapsto \pi(f) \in \mathfrak{A} \tag{2.6}$$

is a unital homomorphism satisfying $\pi(\mathrm{id}) = A$, *where* $\mathrm{id}(z) = z$, *then* $\pi(f) = f(A)$;

(2) $\mathrm{sp}(f(A)) = f(\mathrm{sp}(A))$;

(3) $g \in \mathrm{Rat}(f(\mathrm{sp}(A))) \Rightarrow g \circ f(A) = g(f(A))$.

Proof. Elementary reasoning shows that (2.5) is a homomorphism.

Let us show (1). To this end it is sufficient to show that if $\lambda \in \mathrm{rs}(A)$, then

$$\pi((\lambda - \mathrm{id})^{-1}) = (\lambda - A)^{-1}. \tag{2.7}$$

We know that $\pi(\lambda - \mathrm{id}) = \lambda - A$. Moreover, $(\lambda - \mathrm{id})^{-1} \in \mathrm{Rat}(\mathrm{sp}(A))$ and $(\lambda - \mathrm{id})^{-1}(\lambda - \mathrm{id}) = 1$. Hence

$$\pi((\lambda - \mathrm{id})^{-1})(\lambda - A) = (\lambda - A)\pi((\lambda - \mathrm{id})^{-1}) = 1.$$

Therefore, (2.7) is true.

Let us prove (2) First we show that if $f \in \mathrm{Rat}(\mathrm{sp}(A))$, then

$$\mathrm{sp}(f(A)) \subset f(\mathrm{sp}(A)). \tag{2.8}$$

If $\mu \notin f(\mathrm{sp}(A)))$, then $z \mapsto f(z) - \mu \neq 0$ on $\mathrm{sp}(A)$. Therefore $z \mapsto (f(z) - \mu)$ belongs to $\mathrm{Rat}(\mathrm{sp}(A))$ Therefore $(f(A) - \mu)^{-1}$ exists in $\mathfrak{A}$. Hence $\mu \in \mathrm{rs}(f(A))$, which means that (2.8) is true.

Let us prove

$$\mathrm{sp}(f(A)) \supset f(\mathrm{sp}(A)). \tag{2.9}$$

Let $\mu \notin \mathrm{sp}(f(A))$. Obviously, $(f(A) - \mu)^{-1}$ exists in $\mathfrak{A}$. Let $\lambda \in \mathbb{C}$, $\mu = f(\lambda)$. Then $z \mapsto g(z) := (f(z) - \mu)(z - \lambda)^{-1}$ belongs to $\mathrm{Rat}(\mathrm{sp}(A))$. Hence $g(A)$ is well defined as an element of $\mathfrak{A}$.

Then we check that $g(A)(f(A)-\mu)^{-1}$ is the inverse of $\lambda - A$. Hence $\lambda \notin \mathrm{sp}(A)$. Therefore $\mu \notin f(\mathrm{sp}(A))$, which proves (2.9). □

3 Banach spaces

3.1 Norms and seminorms

Definition 3.1 *Let $\mathcal{X}$ be a vector space over $\mathbb{K}$. $\mathcal{X} \ni x \mapsto \|x\| \in \mathbb{R}$ is called a seminorm iff*

$$1)\ \|x\| \geq 0$$
$$2)\ \|\lambda x\| = |\lambda|\|x\|,$$
$$3)\ \|x + y\| \leq \|x\| + \|y\|.$$

If in addition

$$4)\ \|x\| = 0 \iff x = 0,$$

then it is called a norm.

If $\mathcal{X}$ is a space with a seminorm, then $\mathcal{N} := \{x \in \mathcal{X} \ : \ \|x\| = 0\}$ is a linear subspace. Then on $\mathcal{X}/\mathcal{N}$ we define

$$\|x + \mathcal{N}\| := \|x\|,$$

which is a norm on $\mathcal{X}/\mathcal{N}$.

If $\|\cdot\|$ is a norm, then

$$\mathrm{d}(x, y) := \|x - y\|$$

defines a metric.

Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on $\mathcal{X}$. They are equivalent iff there exist $c_1, c_2 > 0$ such that

$$c_1\|x\|_1 \leq \|x\|_2 \leq c_2\|x\|_1.$$

The equivalence of norms is an equivalence relation. If $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent, then the corresponding metrics are equivalent.

Theorem 3.2 (1) *All norms on a finite dimensional vector space are equivalent.*

(2) *Finite dimensional vector spaces are complete.*

(3) *Every finite dimensional subspace of a normed space is closed.*

For $r > 0$, $(\mathcal{X})_r$ denotes the closed ball in $\mathcal{X}$ of radius r, that is $(\mathcal{X})_r := \{x \in \mathcal{X} \ : \ \|x\| \leq r\}$.

If $\mathcal{V} \subset \mathcal{X}$, then $\mathcal{V}^{\mathrm{cl}}$ will denote the closure of $\mathcal{V}$, $\mathcal{V}^o$ its interior.

3.2 Banach spaces

Definition 3.3 *$\mathcal{X}$ is a Banach space if it has a norm and is complete.*

Definition 3.4 *Let x_i, $i \in I$, be a family of vectors in a normed space. Then*

$$\sum_{i\in I} x_i = x \iff \bigwedge_{\epsilon>0} \bigvee_{I_0\in 2^I_{\rm fin}} \bigwedge_{I_0\subset I_1\in 2^I_{\rm fin}} \left\| \sum_{i\in I_1} x_i - x \right\| < \epsilon.$$

We say then that $\sum_{i\in I} x_i$ is convergent to x.

Clearly,

$$\left\| \sum_{i\in I} x_i \right\| \leq \sum_{i\in I} \|x_i\|.$$

If $c_n \in \mathbb{R}$ and $\sum_{i\in I} c_i$ is convergent, then only a countable number of terms $c_n \neq 0$.

Theorem 3.5 *1) Let $\mathcal{X}$ be a Banach space, $x_i \in \mathcal{X}$ and*

$$\sum_{i\in I} \|x_i\| < \infty.$$

Then there exists

$$\sum_{i\in I} x_i.$$

2) Conversely, if $\mathcal{X}$ is a normed space such that

$$\sum_{n=1}^{\infty} \|x_n\| < \infty$$

implies the convergence of

$$\sum_{n=1}^{\infty} x_n,$$

then $\mathcal{X}$ is a Banach space.

Proof. 1) Since only a countable number of terms x_n is different from zero, the nonzero terms can be treated as a usual sequence indexed by integers. Let

$$y_N := \sum_{n=1}^{N} x_n.$$

For $n \leq m$

$$\|y_n - y_m\| = \left\| \sum_{i=n+1}^{m} x_i \right\| \leq \sum_{i=n+1}^{m} \|x_i\| \to_{n,m\to\infty} 0.$$

Hence (y_N) is Cauchy and therefore convergent.

2) Let (x_n) be a Cauchy sequence in $\mathcal{X}$. By induction we can find a subsequence (x_{n_j}) of the sequence (x_n) such that

$$\|x_{n_{j+1}} - x_{n_j}\| < 2^{-n}.$$

By assumption,

$$\sum_{j=1}^{\infty}(x_{n_{j+1}} - x_{n_j})$$

is convergent. The mth partial sum equals $x_{n_{m+1}} - x_{n_1}$. Hence x_{n_j} is convergent to some $x \in \mathcal{X}$. Since (x_n) was Cauchy, it also has to be convergent to x. □

Theorem 3.6 *Let $\mathcal{X}_0$ be a normed space. Then there exists a unique up to an isometry Banach space $\mathcal{X}$, such that $\mathcal{X}_0 \subset \mathcal{X}$ and $\mathcal{X}_0$ is dense in $\mathcal{X}$. $\mathcal{X}$ is called the completion of $\mathcal{X}_0$ and is denoted $\mathcal{X}_0^{\mathrm{cpl}}$.*

3.3 Bounded operators in a Banach space

Let $\mathcal{X}$ and $\mathcal{Y}$ be normed spaces. An operator $A : \mathcal{X} \to \mathcal{Y}$ is called bounded iff there exists a number C such that

$$\|Ax\| \leq c\|x\|, \quad x \in \mathcal{X}. \tag{3.1}$$

$\|A\|$ is defined as the least c possible in (3.1), or

$$\|A\| := \sup_{x\neq 0} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|\leq 1} \|Ax\| < \infty.$$

The set of such operators is denoted $B(\mathcal{X}, \mathcal{Y})$. We write $B(\mathcal{X}) := B(\mathcal{X}, \mathcal{X})$.

Theorem 3.7 *The following conditions are equivalent:*
1. A is bounded;
2. A is uniformly continuous;
3. A is continuous;
4. A is continuous in one point.

Proof. $1 \Rightarrow 2 \Rightarrow 3 \Rightarrow 4$ is obvious. Clearly, 4. holds $\iff$ A is continuous at 0. Let us show that it implies the boundedness of A.

Suppose A is not bounded. Then there exists a sequence (x_n) such that $\|x_n\| = 1$ and

$$\|Ax_n\| \geq n.$$

Then

$$\lim_{n\to\infty} \frac{x_n}{\sqrt{n}} = 0, \quad \lim_{n\to\infty} \left\| A\frac{x_n}{\sqrt{n}} \right\| = \infty.$$

Thus A is not continuous at 0. □

Example 3.8 *A linear operator from $\mathbb{C}^m$ to $\mathbb{C}^n$ can be defined by a matrix $[a_{ij}]$.*

(1) *If $\mathbb{C}^m$ is equipped with the norm $\|\cdot\|_1$ and $\mathbb{C}^n$ with the norm $\|\cdot\|_\infty$, then $\|A\| = \max\{|a_{ij}|\}$.*

(2) *If $\mathbb{C}^m$ is equipped with the norm $\|\cdot\|_\infty$ and $\mathbb{C}^n$ with the norm $\|\cdot\|_1$, then $\|A\| \leq \sum_{i,j} |a_{ij}|$.*

(3) *If $\mathbb{C}^m$ is equipped with the norm $\|\cdot\|_1$ and $\mathbb{C}^n$ with the norm $\|\cdot\|_1$, then $\|A\| = \max_j\{\sum_i |a_{ij}|\}$.*

(4) *If $\mathbb{C}^m$ is equipped with the norm $\|\cdot\|_\infty$ and $\mathbb{C}^n$ with the norm $\|\cdot\|_\infty$, then $\|A\| = \max_i\{\sum_j |a_{ij}|\}$.*

Proposition 3.9 *All linear operators on a finite dimensional space are bounded.*

Theorem 3.10 *If $\mathcal{Y}$ is a Banach space, then $B(\mathcal{X},\mathcal{Y})$ is a Banach space. Besides, if $A \in B(\mathcal{X},\mathcal{Y})$ and $B \in B(\mathcal{Y},\mathcal{Z})$, then*

$$\|BA\| \leq \|B\|\|A\|.$$

Proof. Clearly, $B(\mathcal{X},\mathcal{Y})$ is a normed space. Let us show that it is complete. Let (A_n) be a Cauchy sequence in $B(\mathcal{X},\mathcal{Y})$. Then $(A_n x)$ is a Cauchy sequence in $\mathcal{Y}$. Define

$$Ax := \lim_{n\to\infty} A_n x.$$

Obviously, A is lenear.

Fix n. Clearly,

$$(A - A_n)x = \lim_{m\to\infty}(A_m - A_n)x.$$

Hence

$$\begin{aligned} &\|(A - A_n)x\| \\ &= \lim_{m\to\infty} \|(A_m - A_n)x\| \leq \|x\| \lim_{m\to\infty} \|(A_m - A_n)\|. \end{aligned}$$

Thus,

$$\|A - A_n\| \leq \lim_{m\to\infty} \|A_m - A_n\|.$$

Therefore, by the Cauchy condition,

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

Thus the sequence A_n is convergent to A. □

Theorem 3.11 *Let $\mathcal{X},\mathcal{Y}$ be Banach spaces and $\mathcal{X}_0$ a dense subspace of $\mathcal{X}$. Let $A_0 \in B(\mathcal{X}_0,\mathcal{Y})$. Then there exists a unique $A \in B(\mathcal{X},\mathcal{Y})$ such that $A\Big|_{\mathcal{X}_0} = A_0$. Moreover, $\|A\| = \|A_0\|$.*

Theorem 3.12 *Let $\mathcal{X}$, $\mathcal{Y}$ be normed spaces. Let $A : \mathcal{X} \to \mathcal{Y}$ be bounded, $\mathcal{X}_0$ dense in $\mathcal{X}$ and* $\operatorname{Ran} A$ *dense in $\mathcal{Y}$. Then $A\mathcal{X}_0$ is dense in $\mathcal{Y}$.*

Proof. Let $y \in \mathcal{Y}$ and $\epsilon > 0$. There exists $y_1 \in \operatorname{Ran} A$ such that $\|y - y_1\| < \epsilon/2$. Let $x_1 \in \mathcal{X}$ such that $Ax_1 = y_1$. Then there exists $x_0 \in \mathcal{X}_0$ such that $\|x - x_0\| < \|A\|^{-1}\epsilon/2$. Hence

$$\|y - Ax_0\| \leq \|y - y_0\| + \|A(x_1 - x_0)\| < \epsilon.$$

□

3.4 Quotient spaces

Let $\mathcal{V}$ be a subspace of a normed space $\mathcal{X}$. In the space $\mathcal{X}/\mathcal{V}$ we introduce

$$\|x + \mathcal{V}\| := \inf\{\|y\| \ : \ y \in x + \mathcal{V}\}.$$

Theorem 3.13 (1) *$\mathcal{X}/\mathcal{V} \ni x + \mathcal{V} \mapsto \|x + \mathcal{V}\|$ is a seminorm and $\|x + \mathcal{V}\| \leq \|x\|$.*

(2) *If $\mathcal{V}$ is closed, it is a norm,*

$$\mathcal{X} \ni x \mapsto x + \mathcal{V} \in \mathcal{X}/\mathcal{V}$$

has norm 1 *and maps open sets onto open sets*

(3) *If $\mathcal{X}$ is Banach, so is $\mathcal{X}/\mathcal{V}$.*

Proof. To prove the triangle inequality we use the property

$$\inf(a+b) \leq \inf a + \inf b.$$

$$\|x+\mathcal{V}\| = 0 \Rightarrow x+\mathcal{V} = \mathcal{V}$$

follows easily from the closedness of $\mathcal{V}$.

To prove the completeness, assume that $y_n \in \mathcal{X}/\mathcal{V}$ satisfies

$$\sum_{n=1}^{\infty} \|y_n\| < \infty.$$

Then, there exists a sequence $x_n \in \mathcal{X}$ such that $y_n = [x_n]$ and

$$\|x_n\| \leq \|y_n\| + 2^{-n}.$$

Hence

$$\sum_{n=1}^{\infty} \|x_n\| < \infty.$$

But $\mathcal{X}$ is complete, hence $\sum_{n=1}^{\infty} x_n$ converges to some $x \in X$. But

$$\|[x - \sum_{n=1}^{N} x_n]\| \leq \|x - \sum_{n=1}^{N} x_n\| \underset{N\to\infty}{\to} 0,$$

hence

$$\sum_{n=1}^{\infty} y_n$$

converges to $[x]$. □

Theorem 3.14 *Let $A \in B(\mathcal{X},\mathcal{Y})$ and $\mathcal{V}$ a closed subspace in* Ker$A$. *Then there exists a unique operator $\tilde{A} \in B(\mathcal{X}/\mathcal{V},\mathcal{Y})$ such that*

$$\tilde{A}(x+\mathcal{V}) := Ax.$$

It satisfies $\|A\| = \|\tilde{A}\|$.

3.5 Dual of a normed space

Let $\mathcal{X}$ be a normed vector space. Recall that the dual space to $\mathcal{X}$, denoted $\mathcal{X}^{\#}$ is the space of bounded linear functionals. In other words, $\mathcal{X}^{\#} := B(\mathcal{X},\mathbb{K})$. Clearly,

$$\|v\| := \sup_{\|x\|\leq 1} |\langle v|x\rangle| < \infty.$$

It follows from Theorem 3.10, that

Theorem 3.15 *$\mathcal{X}^{\#}$ is a Banach space.*

Theorem 3.16 (Hahn-Banach) *Let $\mathcal{X}_0$ be a subspace of a Banach space $\mathcal{X}$. Let $v_0 \in \mathcal{X}_0^{\#}$. Then there exists $v \in X^{\#}$ such that $\|v\| = \|v_0\|$ and*

$$v\Big|_{\mathcal{X}_0} = v_0.$$

Corollary 3.17 *Let $x_0 \in \mathcal{X}$. Then*

$$\sup_{\|v\|\leq 1,\ v\in\mathcal{X}^\#} |\langle v|x_0\rangle| = \|x_0\|.$$

The set

$$\{v \in \mathcal{X}^\# \ : \ \langle v|x_0\rangle = \|x_0\|\} \tag{3.2}$$

is a non-empty convex subset of $\mathcal{X}^\#$.

Proof. The inequality $\leq$ is obvious. Consider $\mathcal{X}_0 = \mathbb{C}x_0$ and $v_0 \in \mathcal{X}_0^\#$ such that

$$\langle v_0|\lambda x_0\rangle := \lambda\|x_0\|.$$

Then $\|v_0\| = 1$. We extend v_0 to v on $\mathcal{X}$ such that $\|v\| = 1$ This proves the inequality (3.2). □

Definition 3.18 *(3.2) is called the set of normed tangent functionals at x_0.*

Theorem 3.19 *If $\mathcal{V}$ is a subspace of $\mathcal{X}$ of a finite codimension, then $\mathcal{V}$ is closed.*

We define

$$\mathcal{X} \ni x \mapsto Jx \in (\mathcal{X}^\#)^\# \tag{3.3}$$

by

$$\langle Jx|v\rangle := \langle v|x\rangle.$$

J is isometric

$$\begin{aligned}\|Jx\| &= \sup_{v\in\mathcal{X}^\#,\|v\|\leq 1} |\langle Jx|v\rangle| \\ &= \sup_{v\in\mathcal{X}^\#,\|v\|\leq 1} |\langle v|x\rangle| = \|x\|.\end{aligned}$$

We will identify $\mathcal{X}$ with the subset $\mathrm{Ran}\, J$ of $X^{\#\#}$. We say that X is reflexive $\iff$ $X^{\#\#} = \mathcal{X}$.

Proposition 3.20 *X is reflexive $\Leftrightarrow$ $X^\#$ is reflexive.*

3.6 Examples of Banach spaces

Let I be an arbitrary set. Let $v = (v^i)_{i\in I}$ be sequences indexed by I with values in $\mathbb{C}$. Set

$$\|v\|_p := \begin{cases} \left(\sum_{i\in I} |v^i|^p\right)^{1/p}, \ 1 \leq p < \infty \\ \sup_{i\in I} |v^i|, \ p = \infty. \end{cases}$$

Lemma 3.21 *The Hölder inequality*

$$\left|\sum_{i\in I} v^i w^i\right| \leq \|v\|_p\|w\|_q, \ p^{-1} + q^{-1} = 1,$$

and the Minkowski inequality are true:

$$\|v + w\|_p \leq \|v\|_p + \|w\|_p.$$

Definition 3.22

$$L^p(I) := \{(v^i), \ i \in I \ : \ \|v\|_p < \infty\}$$

$$C_\infty(I) := \{(v^i), \ i \in I \ : \ \bigwedge_{\epsilon>0} \bigvee_{I_0\in 2^I_{\mathrm{fin}}} \sup_{i\in I\setminus I_0} |v^i| < \epsilon.\}.$$

Theorem 3.23 *$L^p(I)$ with the norm $\|\cdot\|_p$ and $C_\infty(I)$ with the norm $\|\cdot\|_\infty$ are Banach spaces.*

Proof. The Minkowski inequality shows that these are normed spaces.

Le us show that $L^p(I)$ is complete. Let $v_n = (v_n^i)$ satisfy

$$\sum_{n=1}^{\infty} \|v_n\|_p < \infty.$$

Then v_n^i is also summable

$$\sum_{n=1}^{\infty} v_n^i =: v^i.$$

Put $v := (v^i)$. By the Minkowski inequality,

$$\|v - \sum_{j=1}^{n} v_j\|_p < \sum_{j=n+1}^{\infty} \|v^j\|_p.$$

Hence

$$\lim_{n\to\infty} \sum_{j=1}^{n} v^j = v.$$

We check that $v \in L^p(I)$. Thus by Theorem 3.5, $L^p(I)$ is a Banach space.

To see that $C_\infty(I)$ is complete, we check that it is a closed subspace of $L^\infty(I)$. □

Theorem 3.24

$$L^p(I)^\# = L^q(I), \quad p^{-1} + q^{-1} = 1, \ 1 \leq p < \infty,$$
$$C_\infty(I)^\# = L^1(I).$$

Hence $L^1(I)$, $C_\infty(I)$ and $L^\infty(I)$ are not reflexive for infinite sets I.

Proof. Let $C_c(I)$ denote the set of sequences with a finite number of nonzero terms. Let $\mathcal{X} := L^p(I)$ for $p < \infty$ and $\mathcal{X} := C_\infty(I)$ for $p = \infty$. Clearly, $C_c(I)$ is dense in $\mathcal{X}$. Hence ever functional v given on $C_c(I)$, has a unique extension to a functional on $\mathcal{X}$.

The vectors e_i span $C_c(I)$. Hence the functional v is determined by $v_i := \langle v|e_i\rangle$ and its action on $x \in \mathcal{X}$ is given by the formula

$$\langle v|x\rangle = \sum_{i\in I} v_i x_i.$$

By the Hölder inequality

$$\Big|\sum_i v_i x_i\Big| \leq \Big(\sum_i |v_i|^q\Big)^{\frac{1}{q}} \Big(\sum_i |x_i|^p\Big)^{\frac{1}{p}},$$

we have $\mathcal{X}^\# \supset L^q(I)$. Assume that $v \in \mathcal{X}^\# \backslash L^q(I)$. We can assume that $I = \mathbb{N}$ and $p > 1$. Define a sequence of vectors $x^n \in \mathcal{X}$

$$x_i^n := \begin{cases} \overline{v}_i |v_i|^{q-2} \Big(\sum_{i=1}^n |v_i|^q\Big)^{-\frac{1}{p}}, & i \leq n \\ 0, & i > n. \end{cases}$$

Then

$$\langle v|x^n\rangle = \Big(\sum_{i=1}^{n} |v_i|^q\Big)^{1-\frac{1}{p}} \underset{n\to\infty}{\rightarrow} \infty,$$

$$\|x^n\|_p = 1.$$

□

We will now describe when $v \in L^q(I)$ is a normed tangent functional for $x \in L^p(I)$: if $1 < p < \infty$

$$v_i = \overline{x}_i |x_i|^{p-2} \Big(\sum_i |x_i|^p \Big)^{-\frac{1}{p}},$$

if $p = 1$

$$v_i = \overline{x}_i |x_i|^{-1}, \quad x_i \neq 0, \qquad |v_i| \leq 1, x_i = 0.$$

For $L^\infty(I)$, normed tangent functionals can form a many element set.

Example 3.25 *Let (X, μ) be a space with a measure.*

(1) *Let $1 \leq p \leq \infty$. Then $L^p(X, \mu)$ equipped with the norm*

$$\|v\| := \Big(\int |v(x)|^p \mathrm{d}\mu(x) \Big)^{1/p}$$

is a Banach space.

(2) *If $1 \leq p < \infty$, $p^{-1} + q^{-1} = 1$ and the measure is semifinite, then $L^q(X, \mu)$ can be identified with the dual of $L^p(X, \mu)$ by the duality*

$$L^q(X, \mu) \times L^p(X, \mu) \ni (w, y) \mapsto \langle w|y \rangle : \int w(x) v(x) \mathrm{d}\mu(x).$$

3.7 Direct sum of Banach spaces

If $\mathcal{X}, \mathcal{Y}$ are Banach spaces and p is an arbitrary norm in $\mathbb{R}^2$, then $\mathcal{X} \oplus \mathcal{Y}$ becomes a Banach space if we equip it with the norm

$$\|(x, y)\|_p = p(\|x\|, \|y\|).$$

All these norms in $X \oplus Y$ are equivalent and generate the product topology.

If $\mathcal{X}_i$, $i \in I$ is a family of Banach spaces, then we can define $L^p(\mathcal{X}_i,\ i \in I)$ and $C_\infty(\mathcal{X}_i, i \in I)$. They are Banach spaces and we have the obvious analog of Theorem 3.24.

3.8 Integration, differentiation

For continuous $]a, b[\ni t \mapsto v(t) \in \mathcal{X}$ we can define the Riemann integral. It has all the usual properties, for instance,

$$\Big\| \int_a^b v(t) \mathrm{d}t \Big\| \leq \int_a^b \|v(t)\| \mathrm{d}t,$$

if $A \in B(\mathcal{X}, \mathcal{Y})$, then

$$A \int_a^b v(t) \mathrm{d}t = \int_a^b Av(t) \mathrm{d}t.$$

Let $]a, b[\ni t \mapsto v(t) \in \mathcal{X}$. The (norm) derivative of $v(t)$ is defined as

$$\frac{\mathrm{d}}{\mathrm{d}t} v(t_0) := \lim_{h \to 0} \frac{v(t_0 + h) - v(t_0)}{h}.$$

It has all the usual properties, for instance,

$$\frac{\mathrm{d}}{\mathrm{d}t}Av(t_0) := A\frac{\mathrm{d}}{\mathrm{d}t}v(t_0),$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_a^t v(s)\mathrm{d}s = v(t).$$

3.9 Analyticity

We assume that $\mathbb{K} = \mathbb{C}$. Let Ω be an open subset of $\mathbb{C}$. We say that $\Omega \ni z \mapsto v(z) \in \mathcal{X}$ is analytic iff for any $z_0 \in \Omega$ there exists

$$\frac{\mathrm{d}}{\mathrm{d}z}v(z_0) := \lim_{h\to 0}\frac{v(z_0+h)-v(z_0)}{h}.$$

Theorem 3.26 (1) *Let $x_0, x_1, \ldots \in \mathcal{X}$ and $r^{-1} := \limsup_{n\to\infty}\|x_n\|^{\frac{1}{n}}$. Then*

$$v(z) := \sum_{n=0}^{\infty} x_n z^n,\ \ z \in \mathbb{C}$$

is absolutely uniformly convergent for $|z| < r_1 < r$ and divergent for $|z| > r$. In $B(0,r)$ it is analytic

(2) *$\Omega \ni z \mapsto v(z) \in \mathcal{X}$ is analytic iff around any $z_0 \in \Omega$ we can develop it into a power series. Its radius of convergence equals*

$$\left(\limsup_{n\to\infty}\left\|\frac{v^{(n)}(z_0)}{n!}\right\|^{\frac{1}{n}}\right)^{-1}.$$

(3) *If f is analytic on Ω, continuous on Ω^{cl} and $z_0 \in \Omega$, then*

$$f(z_0) = \frac{1}{2\pi\mathrm{i}}\int_{\partial\Omega} f(z)\mathrm{d}z.$$

3.10 Invertible elements

Let $A \in B(\mathcal{X},\mathcal{Y})$. We say it is invertible iff it is bijective and $A^{-1} \in B(\mathcal{Y},\mathcal{X})$.

Theorem 3.27 *Let $A \in B(\mathcal{X},\mathcal{Y})$. Suppose that for some $c > 0$*

$$\|Ax\| \geq c\|x\|.$$

Then $\operatorname{Ran} A$ *is closed. If* $\operatorname{Ran} A = \mathcal{Y}$*, then A is invertible and*

$$\|A^{-1}\| \leq c^{-1}$$

Proof. Let $y_n \in \operatorname{Ran} A$ and $y_n \to y$. Let $Ax_n = y_n$. Then x_n is a Cauchy sequence, hence convergent. Let $\lim_{n\to\infty} x_n := x$. But A is bounded, hence $Ax = y$. Therefore, $\operatorname{Ran} A$ is closed. □

Corollary 3.28 *Let $A \in B(\mathcal{X},\mathcal{Y})$. Suppose that for some $c > 0$*

$$\|Ax\| \geq c\|x\|,$$

and $\operatorname{Ran} A$ *is dense in $\mathcal{Y}$. Then A is invertible.*

Theorem 3.29 *Let $A \in B(\mathcal{X}, \mathcal{Y})$ be invertible and $B \in B(\mathcal{X}, \mathcal{Y})$ such that*

$$\|BA^{-1}\| < 1.$$

Then $A + B$ is invertible and

$$(A+B)^{-1} = \sum_{j=0}^{\infty} (-1)^j A^{-1} (BA^{-1})^j.$$

Moreover,

$$\|(A+B)^{-1}\| \leq \|A^{-1}\|(1 - \|BA^{-1}\|)^{-1},$$

$$\|A^{-1} - (A+B)^{-1}\| \leq \|A^{-1}BA^{-1}\|(1 - \|BA^{-1}\|)^{-1}.$$

In particular, invertible elements form an open subset of $B(\mathcal{X}, \mathcal{Y})$ on which the inverse is a continuous function.

4 Banach algebras

4.1 Banach algebras

An algebra $\mathfrak{A}$ over $\mathbb{C}$ is called a normed algebra, if it is equipped with a norm $\mathfrak{A} \ni A \mapsto \|A\| \in \mathbb{R}$ such that

$$\|AB\| \leq \|A\| \|B\|.$$

It is called a Banach algebra if it is complete in the norm $\| \cdot \|$.

If $\mathfrak{A}$ is a Banach algebra, then every norm closed subalgebra of $\mathfrak{A}$ is a Banach algebra. If $\mathfrak{B} \subset \mathfrak{A}$, then the smallest Banach algebra containing $\mathfrak{B}$ is denoted by $\mathrm{Ban}(\mathfrak{B})$.

Let $\mathcal{V}$ be a Banach space. Recall that $B(\mathcal{V})$ denotes the set of bounded operators on $\mathcal{V}$. Clearly, $B(\mathcal{V})$ is a Banach algebra. Every norm closed subalgebra of $B(\mathcal{V})$ is a Banach algebra. Such Banach algebras are called concrete Banach algebras.

4.2 Ideals and quotient algebras

Recall that if $\mathcal{V}$ is a Banach space and $\mathcal{W}$ its closed subspace, then $\mathcal{V}/\mathcal{W}$ has the structure of a Banach space and $\mathcal{V} \ni v \mapsto v + \mathcal{W} \in \mathcal{V}/\mathcal{W}$ is a surjective open map of norm one.

Theorem 4.1 *Let $\mathfrak{I}$ be a two-sided closed ideal of $\mathfrak{A}$. Then $\mathfrak{A}/\mathfrak{I}$ is a Banach algebra and*

$$\mathfrak{A} \ni A \mapsto A + \mathfrak{I} \in \mathfrak{A}/\mathfrak{I} \tag{4.4}$$

is a norm-decreasing homorphism of Banach algebras with the kernel equal to $\mathfrak{I}$.

Proof. By Theorem 2.1 we know that (4.4) is a homomorphism of algebras with the kernel equal to $\mathfrak{I}$. We also know that $\mathfrak{A}/\mathfrak{I}$ is a Banach space and (4.4) is norm decreasing. Hence it is enough to show that

$$\|(A + \mathfrak{I})(B + \mathfrak{I})\| \leq \|A + \mathfrak{I}\| \|B + \mathfrak{I}\|.$$

□

4.3 Spectrum

Note that Theorem 3.29 remains true if we replace $B(\mathcal{X},\mathcal{Y})$ and $B(\mathcal{Y},\mathcal{X})$ with a unital Banach algebra $\mathfrak{A}$

We assume that $\mathbb{K}=\mathbb{C}$. Let $A\in\mathfrak{A}$. We define the resolvent set of A as

$$\mathrm{rs}A := \{z\in\mathbb{C}\ :\ z-A \ \text{ is invertible }\}.$$

We define the spectrum of A as $\mathrm{sp}A := \mathbb{C}\backslash\mathrm{rs}A$.

Theorem 4.2 *Let $A\in\mathfrak{A}$. Then*

(1) *If $\|(\lambda-A)^{-1}\| = c$, then $\{z:\ |z-\lambda|<c^{-1}\}\subset\mathrm{rs}A$.*

(2) $\|(z-A)^{-1}\| \geq (\mathrm{dist}(z,\mathrm{sp}A))^{-1}$.

(3) $\{|z|>\|A\|\}$ *is contained in* $\mathrm{rs}A$.

(4) $\mathrm{sp}A$ *is a compact subset of* $\mathbb{C}$.

(5) $(z-A)^{-1}$ *is analytic on* $\mathrm{rs}A$.

(6) $(z-A)^{-1}$ *cannot be analytically extended beyond* $\mathrm{rs}A$.

(7) $\mathrm{sp}A\neq\emptyset$

Proof. (1) For $|z-\lambda|<c^{-1}$, we have $\|(z-\lambda)(\lambda-A)^{-1}\| = |z-\lambda|c<1$ Hence we can apply Theorem 3.29. This implies (2)

(3) We check that $\sum_{n=0}^{\infty} z^{-n-1}A^n$ is convergent for $|z|>\|A\|$ and equals $(z-A)^{-1}$.

(4) follows from (1) and (3).

(5) We check that the resolvent is differentiable in the complex sense:

$$h^{-1}\big((z+h-A)^{-1}-(z-A)^{-1}\big) = -(z+h-A)^{-1}(z-A)^{-1}\to -(z-A)^{-2}.$$

(6) follows from (2).

(7) $(z-A)^{-1}$ is an analytic function tending to zero at infinity. Hence it cannot be analytic everywhere, unless it is zero, which is impossible. □

For $A\in B(\mathcal{X})$, the point spectrum is defined as the set of eigenvalues of A

$$\mathrm{sp}_{\mathrm{p}}A = \{z\in\mathbb{C}\ :\ \ \text{there exists } x\in\mathcal{X}\ \text{ such that }\ Ax=zx\}.$$

Clearly, $(\mathrm{sp}_{\mathrm{p}}A)^{\mathrm{cl}}\subset\mathrm{sp}A$.

4.4 Spectral radius

Spectral radius of $A\in B(\mathcal{X})$ is defined as

$$\mathrm{sr}A := \sup_{\lambda\in\mathrm{sp}A}|\lambda|.$$

Lemma 4.3 *Let a sequence of reals (c_n) satisfy*

$$c_n+c_m\geq c_{n+m}.$$

Then

$$\lim_{n\to\infty}\frac{c_n}{n} = \inf\frac{c_n}{n}.$$

Proof. Fix $m \in \mathbb{N}$. Let $n = mq + r$, $r < m$. We have

$$c_n \le qc_m + c_r.$$

So

$$\frac{c_n}{n} \le \frac{qc_m}{n} + \frac{c_r}{n}.$$

Hence

$$\limsup_{n\to\infty} \frac{c_n}{n} \le \frac{c_m}{m}.$$

Thus,

$$\limsup_{n\to\infty} \frac{c_n}{n} \le \inf \frac{c_m}{m}.$$

□

Theorem 4.4 *Let $A \in B(\mathcal{X})$. Then*

$$\lim_{n\to\infty} \|A^n\|^{\frac{1}{n}}$$

exists and equals $\mathrm{sr}A$*. Besides,* $\mathrm{sr}A \le \|A\|$.

Proof. Let

$$c_n := \log \|A^n\|.$$

Then

$$c_n + c_m \ge c_{n+m}$$

Hence there exists

$$\lim_{n\to\infty} \frac{c_n}{n}.$$

Consequently, there exists

$$r := \lim_{n\to\infty} \|A^n\|^{1/n}.$$

By the Cauchy criterion, the series

$$\sum_{n=0}^{\infty} A^n z^{-1-n}. \tag{4.5}$$

is absolutely convergent for $|z| > r$, and divergent for $|z| < r$. We easily check that (4.5) equals $(z - A)^{-1}$.
□

4.5 Idempotents

Theorem 4.5 *Let $P, Q \in B(\mathcal{V})$ be idempotents such that* $\mathrm{sr}(P-Q)^2 < 1$*. Then there exists an invertible $U \in B(\mathcal{X})$ such that $P = UQU^{-1}$.*

Proof. Set

$$\tilde{U} := QP + (1-Q)(1-P), \quad \tilde{V} := PQ + (1-P)(1-Q).$$

We have

$$Q\tilde{U} = \tilde{U}P, \quad P\tilde{V} = \tilde{V}Q.$$

We also have

$$\tilde{V}\tilde{U} = \tilde{U}\tilde{V} = 1 - R,$$

$$R = (P-Q)^2 = P + Q - PQ - QP.$$

We check that P and Q commute with R (note in particular that $PR = P - PQP$, etc.).

Set $c := \mathrm{sr}R < 1$. Then on $\mathrm{sp}(1-R) \subset B(1,c)$, the function $z \mapsto z^{\frac{1}{2}}$ is well defined. Hence we can introduce the function

$$(1-R)^{-1/2}$$

(which can be defined by a convergent power series). We set

$$U := \tilde{U}(1-R)^{-1/2} = (1-R)^{-1/2}\tilde{U}, \quad V := \tilde{V}(1-R)^{-1/2} = (1-R)^{-1/2}\tilde{V}.$$

So $UV = VU = 1$, or $V = U^{-1}$ and

$$Q = UPU^{-1}.$$

Proposition 4.6 *Let $t \mapsto P(t)$ be a differentiable function with values in idempotents. Then*

$$P\dot{P}P = 0.$$

Proof.

$$\frac{\mathrm{d}}{\mathrm{d}t}P = \frac{\mathrm{d}}{\mathrm{d}t}P^2 = \dot{P}P + P\dot{P}.$$

Hence $P\dot{P}P = 2P\dot{P}P$. □

4.6 Functional calculus

Let $K \subset \mathbb{C}$ be compact. By $\mathrm{Hol}(K)$ let us denote the set of analytic functions on a neighborhood of K. It is a commutative algebra.

More precisely, let $\widetilde{\mathrm{Hol}}(K)$ be the set of pairs $(f, \mathcal{D})$, where $\mathcal{D}$ is an open subset of $\mathbb{C}$ containing K. We introduce the relation $(f_1, \mathcal{D}_1) \sim (f_2, \mathcal{D}_2)$ iff $f_1 = f_2$ on $\mathcal{D}_1 \cap \mathcal{D}_2$. We set $\mathrm{Hol}(K) := \widetilde{\mathrm{Hol}}(K)/\sim$.

Definition 4.7 *Let $A \in \mathfrak{A}$ and $f \in \mathrm{Hol}(\mathrm{sp}A)$. Let γ be a contour in the domain of f that encircles* $\mathrm{sp}A$ *counterclockwise. We define*

$$f(A) := \frac{1}{2\pi\mathrm{i}}\int_\gamma (z-A)^{-1} f(z)\mathrm{d}z \tag{4.6}$$

Clearly, the definition is independent of the choice of the contour.

Theorem 4.8

$$\mathrm{Hol}(\mathrm{sp}A) \ni f \mapsto f(A) \in \mathfrak{A} \tag{4.7}$$

is a linear map satisfying

(1) $fg(A) = f(A)g(A)$;

(2) $1(A) = 1$;

(3) $\mathrm{id}(A) = A$, *for* $\mathrm{id}(z) = z$;

(4) *If $f(z) := \sum_{n=0}^\infty f_n z^n$ is an analytic function defined by a series absolutely convergent in a disk of radius greater than* $\mathrm{sr}A$, *then*

$$f(A) = \sum_{n=0}^\infty f_n A^n;$$

(5) *(Spectral mapping theorem).* $\mathrm{sp}f(A) = f(\mathrm{sp}A)$

(6) $g \in \mathrm{Hol}(f(\mathrm{sp}A)) \Rightarrow g \circ f(A) = g(f(A))$,

(7) $\|f(A)\| \leq c_{\gamma,A} \sup_{z\in\gamma} |f(z)|$.

Proof. From the formula

$$(z-A)^{-1} = \sum_{n=0}^{\infty} z^{-n-1}A^n, \quad |z| > \mathrm{sr}(A),$$

we get that $1(A) = 1$ and $\mathrm{id}(A) = A$.

It is clear that $f \to f(A)$ is linear. Let us show that it is multiplicative. Let $f_1, f_2 \in \mathrm{Hol}(\mathrm{sp}A)$. Choose a contour γ_2 around the contour γ_1, both in the domains of f_1 and f_2.

$$\begin{aligned}
&(2\pi\mathrm{i})^{-2}\int_{\gamma_1} f_1(z_1)(z_1-A)^{-1}\mathrm{d}z_1 \int_{\gamma_2} f_2(z_2)(z_2-A)^{-1}\mathrm{d}z_2\\
&= (2\pi\mathrm{i})^{-2}\int_{\gamma_1}\int_{\gamma_2} f_1(z_1)f_2(z_2)\left((z_1-A)^{-1}-(z_2-A)^{-1}\right)(z_2-z_1)^{-1}\mathrm{d}z_1\mathrm{d}z_2\\
&= (2\pi\mathrm{i})^{-2}\int_{\gamma_1} f_1(z_1)(z_1-A)^{-1}\mathrm{d}z_1 \int_{\gamma_2}(z_2-z_1)^{-1}f_2(z_2)\mathrm{d}z_2\\
&+(2\pi\mathrm{i})^{-2}\int_{\gamma_2} f_2(z_2)(z_2-A)^{-1}\mathrm{d}z_2 \int_{\gamma_1}(z_1-z_2)^{-1}f_1(z_1)\mathrm{d}z_1.
\end{aligned}$$

But

$$\begin{aligned}
&\int_{\gamma_1}(z_1-z_2)^{-1}f_1(z_1)\mathrm{d}z_1 = 0,\\
&\int_{\gamma_2}(z_2-z_1)^{-1}f_2(z_2)\mathrm{d}z_2 = 2\pi\mathrm{i}f_2(z_1).
\end{aligned}$$

Thus

$$f_1(A)f_2(A) = f_1f_2(A). \tag{4.8}$$

Let us prove the spectral mapping theorem. First we will show

$$\mathrm{sp}f(A) \subset f(\mathrm{sp}A). \tag{4.9}$$

If $\mu \notin f(\mathrm{sp}A)$, then the function $z \mapsto f(z)-\mu \neq 0$ on $\mathrm{sp}A$. Therefore, $z \mapsto (f(z)-\mu)^{-1}$ belongs to $\mathrm{Hol}(\mathrm{sp}A)$. Thus $f(A)-\mu$ is invertible and therefore, $\mu \notin \mathrm{sp}f(A)$. This implies (4.9).

Let us now show

$$\mathrm{sp}f(A) \supset f(\mathrm{sp}A). \tag{4.10}$$

Let $\mu \notin \mathrm{sp}f(A)$. This clearly implies that $f(A)-\mu$ is invertible.

If μ does not belong to the image of f, then of course it does not belong to $f(\mathrm{sp}A)$. Let us assume that $\mu = f(\lambda)$. Then the function

$$z \mapsto g(z) := (f(z)-\mu)(z-\lambda)^{-1}$$

belongs to $\mathrm{Hol}(\mathrm{sp}A)$. Hence $g(A)$ is well defined as an element of $B(\mathcal{X})$. We check that $g(A)(f(A)-f(\lambda))^{-1} = (\lambda-A)^{-1}$. Hence $\lambda \notin \mathrm{sp}A$. Thus $\mu \notin f(\mathrm{sp}A)$. Consequently, (4.10) holds.

Let us show now (6). Notice that if $w \notin f(\mathrm{sp}A)$, then the function $z \mapsto (w-f(z))^{-1}$ is analytic on a neighborhood of

$$(w-f(A))^{-1} = \frac{1}{2\pi i}\int_{\gamma}(w-f(z))^{-1}(z-A)^{-1}\mathrm{d}z.$$

We compute

$$\begin{aligned}
&g(f(A))\\
&= \tfrac{1}{2\pi\mathrm{i}}\int_{\tilde{\gamma}} g(w)(w-f(A))^{-1}\mathrm{d}w\\
&= \tfrac{1}{(2\pi\mathrm{i})^2}\int_{\tilde{\gamma}}\int_{\gamma} g(w)(w-f(z))^{-1}(z-A)^{-1}\mathrm{d}w\mathrm{d}z\\
&= \tfrac{1}{(2\pi\mathrm{i})^2}\int_{\gamma}(z-A)^{-1}\mathrm{d}z\int_{\tilde{\gamma}} g(w)(w-f(z))^{-1}\mathrm{d}w\\
&= \tfrac{1}{2\pi\mathrm{i}}\int_{\gamma} g(f(z))(z-A)^{-1}\mathrm{d}z.
\end{aligned}$$

□

4.7 Spectral idempotents

Let Ω be a subset of $B \subset \mathbb{C}$. Ω will be called an isolated subset of B, if $\Omega \cap (B\backslash\Omega)^{\mathrm{cl}} = \emptyset$ and $\Omega^{\mathrm{cl}} \cap (B\backslash\Omega) = \emptyset$ (or Ω is closed and open in the relative topology of B).

Let $\Omega \subset \mathrm{sp}A$ be an isolated subset of $\mathrm{sp}A$. We define the function

$$1_\Omega(z) := \begin{cases} 1 & z \text{ belongs to a neighborhood of } \Omega, \\ 0 & z \text{ belongs to a neighborhood of } \mathrm{sp}A\backslash\Omega. \end{cases}$$

Clearly, $1_\Omega \in \mathrm{Hol}(\mathrm{sp}A)$ and $1_\Omega^2 = 1_\Omega$ Hence $1_\Omega(A)$ is an idempotent.

If γ is a contour around Ω outside of $\mathrm{sp}A\backslash\Omega$, then

$$1_\Omega(A) = \frac{1}{2\pi \mathrm{i}} \int_\gamma (z - A)^{-1} \mathrm{d}z$$

This operator will be called the spectral idempotent of the operator A onto Ω. Let $\mathfrak{A}_\Omega := 1_\Omega(A)\mathfrak{A}1_\Omega(A)$ be the subalgebra of $\mathfrak{A}$ reduced by the projection $1_\Omega(A)$. Then

$$\mathrm{sp}_{\mathfrak{A}_\Omega}(A1_\Omega(A)) = \mathrm{sp}A \cap \Omega.$$

If Ω_1 and Ω_2 are two isolated subsets of $\mathrm{sp}A$, then

$$1_{\Omega_1}(A)1_{\Omega_2}(A) = 1_{\Omega_1 \cap \Omega_2}(A)$$

4.8 Isolated eigenvalues

Assume now that λ is an isolated point of $\mathrm{sp}A$. Set

$$P := 1_\lambda(A), \qquad N := (A - \lambda)P.$$

Definition 4.9 *We say that λ is a semisimple eigenvalue if $N = 0$ If $N^n = 0$ and $N^{n-1} \neq 0$, then we say that λ is nilpotent of degree n. It is easy to see that if $A \in L(\mathcal{X})$, then the degree of nilpotence of λ is less than or equal to* $\dim P$.

Proposition 4.10 *The operator N is quasinilpotent, satisfies $PN = NP = N$ and can be written as*

$$N = f(A), \quad f(z) := (z - \lambda)1_\lambda(z). \tag{4.11}$$

Besides,

$$(z - A)^{-1}P = (z - \lambda)^{-1}P + \sum_{j=1}^{\infty} N^j (z - \lambda)^{-j+1}.$$

and $(z - A)^{-1}(1 - P)$ is analytic in the neighborhood of λ. If N is nilpotent of degree n, then there exist $\delta > 0$ and C such that

$$\|(z - A)^{-1}\| \leq C|z - \lambda|^{-n}, \quad z \in B(\lambda, \delta). \tag{4.12}$$

Proof. Clearly, $AP = A1_\lambda(A)$ and $\lambda P = \lambda 1_\lambda(A)$. This shows (4.11). Then note that $f(z) = 0$ for $z \in \mathrm{sp}A$. Hence $\mathrm{sp}N = \{0\}$.

Using the Laurent series expansion we get

$$(z - A)^{-1} = \sum_{n=-\infty}^{\infty} C_n (z - \lambda)^n,$$

where

$$C_n = \frac{1}{2\pi \mathrm{i}} \int_\gamma (z-A)^{-1}(z-\lambda)^{-n-1}\mathrm{d}z.$$

Clearly, $C_{-1} = P$ and $C_{-2} = N$. Besides, by Theorem 4.8 we obtain

$$C_{-1-n}C_{-1-m} = C_{-1-n-m}.$$

□

4.9 Spectral theory in finite dimension

Suppose that $\mathcal{X}$ is finite dimensonal of dimension d and $A \in L(\mathcal{X})$. Then sp$A$ has at most d elements. Let sp$A = \{\lambda_1, \dots, \lambda_n\}$.

We say that A is diagonalizable iff

$$A = \sum_{j=1}^{n} \lambda_j 1_{\lambda_j}(A).$$

It is well known that in a finite dimension for every $A \in L(\mathcal{X})$, there exist unique diagonalizable D and nilpotent N satisfying $DN = ND$ such that $A = D + N$. Let m be the degree of nilpotence of N.

In fact, define two functions on a neighborhood of sp$A$: $d(z)$ is equal to λ_i on a neighborhood of $\lambda_i \in$ sp$A$ and $n(z) = z - \lambda_i$ on a neighborhood of $\lambda_i \in$ sp$A$. Both d and n belong to Hol(sp$A$). Clearly, and $D := d(A)$ and $N := n(A)$ satisfy the above requirements.

Clearly then $N = \sum_{j=1}^{n} N_j$ with $N_j = P_j N P_j$ also nilpotent. Let m_j be the degree of nilpotence of N_j. We have

$$\begin{aligned} f(A) &= \textstyle\sum_{k=0}^{m} f^{(k)}(D)\frac{N^k}{k!} \\ &= \textstyle\sum_{j=1}^{n}\sum_{k=0}^{m_j} f^{(k)}(\lambda_j)\frac{N_j^k}{k!}. \end{aligned}$$

4.10 Functional calculus for several commuting operators

Let $K \subset \mathbb{C}^n$ be compact. By Hol(K) let us denote the set of analytic functions on a neighborhood of K. It is a commutative algebra.

Let $\mathfrak{A}$ be a Banach algebra.

Definition 4.11 *Let $A_1, \dots, A_n \in \mathfrak{A}$ commute with one another. Let $F \in$ Hol(sp$A_1 \times \cdots \times$ spA_n). Let $\gamma_1, \dots, \gamma_n$ be contours such that $\gamma_1 \times \cdots \times \gamma_n$ lies in the domain of F and each γ_j encircles* spA_j *counterclockwise. We define*

$$F(A_1, \dots, A_n) := \frac{1}{(2\pi i)^n} \int_{\gamma_1} \mathrm{d}z_1 \cdots \int_{\gamma_n} \mathrm{d}z_n (z_1 - A_1)^{-1} \cdots (z_n - A_n)^{-1} F(z_1, \dots, z_n). \tag{4.13}$$

Clearly, the definition is independent of the choice of the contour.

Theorem 4.12

$$\mathrm{Hol}(\mathrm{sp}A_1 \times \cdots \times \mathrm{sp}A_n) \ni F \mapsto F(A_1, \dots, A_n) \in \mathfrak{A} \tag{4.14}$$

is a linear map satisfying

(1) $FG(A_1, \dots, A_n) = F(A_1, \dots, A_n)G(A_1, \dots, A_n)$;

(2) $1(A_1, \dots, A_n) = 1$;

(3) $\mathrm{id}_j(A_1, \dots, A_n) = A_j$, *for* $\mathrm{id}_j(z_1, \dots, z_n) := z_j$;

(4) *If* $F(z_1,\dots,z_n) := \sum_{m_1,\dots,m_n=0}^{\infty} F_{m_1,\dots,m_n} z_1^{m_1}\cdots z_n^{m_n}$ *is an analytic function defined by a series absolutely convergent in a neighborhood of* $\mathrm{B}(\mathrm{sr}A_1)\times\cdots\times\mathrm{B}(\mathrm{sr}A_n)$, *then*

$$F(A_1,\dots,A_n) = \sum_{m_1,\dots,m_n=0}^{\infty} F_{m_1,\dots,m_n} A_1^{m_1}\cdots A_n^{m_n};$$

(5) *(Weak version of the spectral mapping theorem).* $\mathrm{sp}F(A_1,\dots,A_n) \subset F(\mathrm{sp}A_1,\dots,\mathrm{sp}A_n)$

(6) $g \in \mathrm{Hol}(F(\mathrm{sp}A_1\times\ldots\times\mathrm{sp}A_n)) \Rightarrow g\circ F(A_1,\dots,A_n) = g(F(A_1,\dots,A_n))$,

(7) $\|F(A_1,\dots,A_n)\| \leq c_{\gamma,A_1,\dots,A_n}\sup_{z\in\gamma}|f(z)|$.

Proof. The proof is essentially the same as that of Theorem 4.8. Let us show for instance the weak version of the spectral mapping theorem. Let $\mu \notin F(\mathrm{sp}A_1,\dots,\mathrm{sp}A_n)$. Then the function $(z_1,\dots,z_n)\mapsto F(z_1,\dots,z_n)-\mu\neq 0$ on $\mathrm{sp}A_1\times\ldots\times\mathrm{sp}A_n$. Therefore, $(z_1,\dots,z_n)\mapsto(F(z_1,\dots,z_n)-\mu)^{-1}$ belongs to $\mathrm{Hol}(\mathrm{sp}A_1\times\ldots\times A_n)$. Thus $F(A_1,\dots,A_n)-\mu$ is inverible and therefore, $\mu\notin\mathrm{sp}F(A_1,\dots,,A_n)$. □

5 Hilbert spaces

5.1 Scalar product spaces

Let $\mathcal{V}$ be a vector space.

$$\mathcal{V}\times\mathcal{V}\ni(v,y)\mapsto(v|y)\in\mathbb{C}$$

is called a scalar product if

$$\begin{aligned}
&(v|y+z) = (v|y)+(v|z), \qquad (v|\lambda y) = \lambda(v|y),\\
&(v+y|z) = (v|z)+(y|z), \qquad (\lambda v|y) = \overline{\lambda}(v|y),\\
&(v|v)\geq 0,\\
&(v|v)=0 \Rightarrow v=0.
\end{aligned}$$

Theorem 5.1 (The hermitian property.)

$$(v|y) = \overline{(y|v)}.$$

Proof. We use the polarization identity:

$$\begin{aligned}
(v|y) &= \tfrac{1}{4}\textstyle\sum_{n=0}^{3}(-\mathrm{i})^n(v+\mathrm{i}^n y|v+\mathrm{i}^n y),\\
(y|v) &= \tfrac{1}{4}\textstyle\sum_{n=0}^{3}\mathrm{i}^n(v+\mathrm{i}^n y|v+\mathrm{i}^n y).
\end{aligned}$$

□

We define

$$\|v\| := \sqrt{(v|v)}$$

Theorem 5.2 (The parallelogram identity.)

$$2(\|v\|^2+\|y\|^2) = \|v+y\|^2+\|v-y\|^2.$$

Theorem 5.3 (The Schwarz inequality.)

$$|(v|y)| \leq \|v\|\|y\|$$

Proof.

$$0 \quad \leq (v+ty|v+ty) = \|v\|^2 + t(v|y) + \overline{t(v|y)} + \|y\|^2|t|^2.$$

We set $t = -\frac{\overline{(v|y)}}{\|y\|^2}$ and we get

$$0 \leq \|v\|^2 - \frac{|(v|y)|^2}{\|y\|^2}.$$

□

Theorem 5.4 (The triangle inequality.)

$$\|v+y\| \leq \|v\| + \|y\|$$

Proof.

$$\|v+y\|^2 = \|v\|^2 + (v|y) + (y|v) + \|y\|^2 \leq (\|v\| + \|y\|)^2.$$

□

Hence $\|\cdot\|$ is a norm.

5.2 The definition and examples of Hilbert spaces

Definition 5.5 *A space with a scalar product is called a Hilbert space if it is complete.*

Example 5.6 *Let I be an arbitrary set of indices. Then $L^2(I)$ denotes the space of families $(v^i)_{i\in I}$ with values in $\mathbb{C}$ indexed by I such that*

$$\sum_{i\in I} |v^i|^2 < \infty$$

equipped with the scalar product

$$(v|w) = \sum_{i\in I} \overline{v^i} w^i.$$

The Schwarz inequality guarantees that the scalar product is well defined.

Example 5.7 *Let (X,μ) be a space with a measure. Then $L^2(X,\mu)$ equipped with the scalar product*

$$(v|w) := \int \overline{v}(x) w(x) \mathrm{d}\mu(x)$$

is a Hilbert space.

Theorem 5.8 *Let $\mathcal{V}_0$ be a space equipped with a scalar product (but not necessarily complete). Let $\mathcal{V}_0^{\rm cpl}$ be its completion (see Theorem 3.6). Then there exists a unique scalar product on $\mathcal{V}_0^{\rm cpl}$, which is compatible with the norm on $\mathcal{V}_0^{\rm cpl}$. $\mathcal{V}_0^{\rm cpl}$ with this scalar product is is a Hilbert space.*

5.3 Complementary subspaces

Suppose that (for the time being) $\mathcal{V}$ is a space with a scalar product (not necessarily complete).

If $A \subset \mathcal{V}$, then $A^\perp$ denotes

$$A^\perp := \{v \in \mathcal{V} \ : \ (v|z) = 0, \quad z \in A\}.$$

Proposition 5.9 (1) *$A^\perp$ is a closed subspace.*

(2) $A \subset B \ \Rightarrow A^\perp \supset B^\perp$

(3) $(A^\perp)^\perp \supset \mathrm{Span}(A)^{\mathrm{cl}}$

Proof. 1. and 2. are obvious. To prove 3. we note that $(A^\perp)^\perp \supset A$. But $(A^\perp)^\perp$ is a closed subspace by 1. Hence it contains the least closed subspace containing A, or $\mathrm{Span}(A)^{\mathrm{cl}}$. □

Suppose that $\mathcal{V}$ is Hilbert space.

Theorem 5.10 *Let $\mathcal{W}$ be a closed subspace of $\mathcal{V}$. Then $\mathcal{W}^\perp$ is a closed subspace and*

$$\mathcal{W} \oplus \mathcal{W}^\perp = \mathcal{V}, \quad (\mathcal{W}^\perp)^\perp = \mathcal{W}.$$

Proof. Let

$$\inf_{w\in\mathcal{W}} \|v - w\| =: d.$$

Then there exists a sequence $y_n \in \mathcal{W}$ such that

$$\lim_{n\to\infty} \|v - y_n\| = d.$$

Then using first the parallelogram identity and then $\frac{1}{2}(y_n + y_m) \in \mathcal{W}$ we get

$$\begin{aligned} \|y_n - y_m\|^2 &= 2\|y_n - v\|^2 + 2\|y_m - v\|^2 - 4\|v - \tfrac{1}{2}(y_n + y_m)\|^2 \\ &\leq 2\|y_n - v\|^2 + 2\|y_m - v\|^2 - 4d^2 \to 0. \end{aligned}$$

Therefore, (y_n) is a Cauchy sequence and hence

$$\lim_{n\to\infty} y_n =: y.$$

Clearly, $y \in \mathcal{W}$ and it is an element closest to v. We set $z := v - y$. We will show that $z \in \mathcal{W}^\perp$. Let $w \in \mathcal{W}$. Then

$$\begin{aligned} \|z\|^2 &= \|v - y\|^2 \leq \|v - (y + tw)\|^2 \\ &= \|z - tw\|^2 = \|z\|^2 - \bar{t}(w|z) - t\overline{(w|z)} + |t|^2\|w\|^2. \end{aligned}$$

We set $t = \frac{(w|z)}{\|w\|^2}$. We get

$$0 \leq -\frac{|(w|z)|^2}{\|w\|^2}.$$

Thus $(w|z) = 0$. This shows that $\mathrm{Span}(\mathcal{W} \cup \mathcal{W}^\perp) = \mathcal{V}$.

$\mathcal{W} \cap \mathcal{W}^\perp = \{0\}$ is obvious. This implies the uniqueness of the pair $y \in \mathcal{W}$, $z \in \mathcal{W}^\perp$. This ends the proof of $\mathcal{V} = \mathcal{W} \oplus \mathcal{W}^\perp$.

Let us show now that $(\mathcal{W}^\perp)^\perp \subset \mathcal{W}$. Let $v \in (\mathcal{W}^\perp)^\perp$. Then $v = y + z$, where $y \in \mathcal{W}$, $z \in \mathcal{W}^\perp$. But $(z|v) = 0$ and $(z|y) = 0$. We have

$$(v|z) = (y|z) + (z|z).$$

Hence $(z|z) = 0$, or $z = 0$, therefore $v \in \mathcal{W}$ □

Corollary 5.11

$$A^{\perp^\perp} = \mathrm{Span}(A)^{\mathrm{cl}}$$

Proof.

$$\mathrm{Span}(A)^{\mathrm{cl}^{\perp^\perp}} \supset A^{\perp^\perp} \supset \mathrm{Span}(A)^{\mathrm{cl}}$$

follows by Proposition 5.9.

$$\mathrm{Span}(A)^{\mathrm{cl}^{\perp^\perp}} = \mathrm{Span}(A)^{\mathrm{cl}}$$

follows by Proposition 5.10. □

5.4 Orthonormal basis

Assume for the time being that $\mathcal{V}$ is a space with a scalar product.

Definition 5.12 *$A \subset \mathcal{V} \backslash \{0\}$ is an orthogonal system iff $e_1, e_2 \in A$, $e_1 \neq e_2$ implies $(e_1|e_2) = 0$. $A \subset \mathcal{V}$ is na orthonormal system if it is orthogonal and if $e \in A$, then $\|e\| = 1$.*

Theorem 5.13 *Let $(e_1, \dots, e_N)$ be an orthonormal system. We then have* **the Pythagoras Theorem**

$$\|v\|^2 = \sum_{n=1}^{N} |(v|e_n)|^2 + \|v - \sum_{n=1}^{N} (e_n|v) e_n\|^2$$

and **the Bessel inequality***:*

$$\|v\|^2 \geq \sum_{n=1}^{N} |(v|e_n)|^2.$$

Assume now that $\mathcal{V}$ is a Hilbert space.

Definition 5.14 *A maximal orthonormal system is called an orthonormal basis.*

Theorem 5.15 *Let $\{e_i\}_{i \in I}$ be an orthonormal system. It is an orthonormal basis iff one of the following conditions holds:*

(1) $\{e_i \ : \ i \in I\}^{\perp} = \{0\}$.

(2) $(\mathrm{Span}\{e_i \ : \ i \in I\})^{\mathrm{cl}} = \mathcal{V}$

Theorem 5.16 *Every orthonormal system can be completed to an orthonormal basis.*

Proof. Let $\mathcal{B}$ denote the family of all orthonormal systems ordered by inclusion. Let $\{A_i \ : \ i \in I\} \subset \mathcal{B}$ be a subset linearly ordered. Then

$$\cup_{i \in I} A_i$$

is also an orthonormal system. It is also an upper bound of the set $\{A_i \ : \ i \in I\}$. Hence we can apply the Kuratowski-Zorn lemma. □

The definition of an orthogonal basis is similar. From an orthogonal basis $(w_i)_{i \in I}$ we can construct an orthonormal basis $\{\|w_i\|^{-\frac{1}{2}} w_i\}_{i \in I}$.

Theorem 5.17 *Let $(e_i)_{i \in I}$ be an orthonormal basis. Then*

(1)

$$v = \sum_{i \in I} (e_i|v) e_i, \tag{5.15}$$

and

$$\|v\|^2 = \sum_{i \in I} |(v|e_i)|^2.$$

(2) *If*

$$v = \sum_{i \in I} \lambda_i e_i,$$

then $\lambda_i = (e_i|v)$.

Proof. By the Bessel inequality, a finite number of coefficients is greater than $\epsilon > 0$. Hence a countable number of coefficients is non-zero. Let us enumerate the non-zero coefficients $(e_{i_n}|v),\ n = 1, 2, \ldots$ By the Bessel inequality, we get

$$\sum_{i=1}^{\infty} |(e_i|v)|^2 \leq \|v\|^2.$$

Set

$$v_N := \sum_{n=1}^{N} (e_i|v) e_i,$$

Then for $N < M$

$$\|v_M - v_M\|^2 = \sum_{i=N+1}^{M} |(e_i|v)|^2.$$

Hence by the completeness of $\mathcal{V}$ we get the convergence of v_N and thus the convergence of the series. Besides, the vector

$$v - \sum_{i \in I} e_i (e_i|v)$$

is orthogonal to the basis. Hence it is zero. This proves 1. □

Theorem 5.18 *Let B_1 and B_2 be orthonormal bases in $\mathcal{V}$. Then they have the same cardinality.*

Proof. First we prove this for finite B_1 or B_2.

For any $y \in B_1$ there exists a countable number of $x \in B_2$ such that $(x|y) \neq 0$. For every $x \in B_2$ we will find $y \in B_1$ such that $(x|y) \neq 0$. Hence there exists a function $f : B_2 \to B_1$ such that the preimage of every set is countable. Hence

$$|B_2| \leq |B_1 \times \mathbb{N}| = \max(|B_1|, \aleph_0).$$

Similarly we check that

$$|B_1| \leq \max(|B_2|, \aleph_0).$$

□

Definition 5.19 *The cardinality of this basis is called the dimension of the space.*

Definition 5.20 *We say that a linear operator $U : \mathcal{V}_1 \to \mathcal{V}_2$ is unitary iff it is a bijection and*

$$(Uw|Uv) = (w|v), \quad v, w \in \mathcal{V}_1.$$

We say that the Hilbert spaces $\mathcal{V}_1$ and $\mathcal{V}_2$ are isomorphic iff there exists a unitary operator from $\mathcal{V}_1$ to H_2.

Theorem 5.21 *Two Hilbert spaces are isomorphic iff they have the same dimension.*

Proof. Let $\{x_i : i \in I\}$ be an orthonormal basis in $\mathcal{V}$. It suffices to show that $\mathcal{V}$ is isomorphic to $L^2(I)$. We define the unitary operator

$$(Uv)_i := (x_i|v).$$

□

5.5 The space dual to a Hilbert space

Theorem 5.22 (The Riesz Lemma) *The formula*

$$\langle C_{\mathcal{V}}(v)|x\rangle := (v|x)$$

defines an antilinear isometry from $\mathcal{V}$ onto $\mathcal{V}^{\#}$.

Proof. Isometricity:

$$\|C_{\mathcal{V}}(v)\| = \sup_{\|x\|\leq 1} |(v|x)| \leq \|v\|.$$

It suffices to take $x = \frac{v}{\|v\|}$ to get the equality.

Surjectivity: Let $w \in \mathcal{V}^{\#}$ and $\mathcal{W} := \mathrm{Ker} w$. If $\mathcal{W} = \mathcal{V}$, then $w = C(0)$. If not, then let $x_0 \in \mathcal{W}^{\perp}$, $\|x_0\| = 1$. Set

$$v := x_0\overline{\langle w|x_0\rangle}.$$

We will prove that $w = C(v)$.

An arbitrary y can be represented as

$$y = \left(y - \frac{\langle w|y\rangle}{\langle w|x_0\rangle}x_0\right) + \frac{\langle w|y\rangle}{\langle w|x_0\rangle}x_0$$

The first term belongs to $\mathcal{W}$. Hence

$$\begin{aligned}(v|y) &= \left(x_0\overline{\langle w|x_0\rangle}|\tfrac{\langle w|y\rangle}{\langle w|x_0\rangle}x_0\right)\\ &= \tfrac{\langle w|y\rangle\langle w|x_0\rangle}{\langle w|x_0\rangle} = \langle w|y\rangle.\end{aligned}$$

□

The space dual to a Hilbert space has a natural structure of a Hilbert space:

$$(C_{\mathcal{V}}v|C_{\mathcal{V}}x) := (x|v), \quad v, x \in \mathcal{V}.$$

Theorem 5.23 *A Hilbert space is reflexive: $(\mathcal{V}^{\#})^{\#} = \mathcal{V}$ and $C_{\mathcal{V}^{\#}}C_{\mathcal{V}} = 1$.*

Proof. Let $y \in \mathcal{V}$, $v \in \mathcal{V}^{\#}$. Then

$$\langle C_{\mathcal{V}^{\#}}C_{\mathcal{V}}y|v\rangle = (C_{\mathcal{V}}y|v) = (C_{\mathcal{V}}^{-1}v|y) = \langle v|y\rangle = \langle Jy|v\rangle,$$

where J was defined in (3.3) Or, $C_{\mathcal{V}^{\#}}C_{\mathcal{V}} = J$. But $C_{\mathcal{V}^{\#}}$ and $C_{\mathcal{V}}$ are bijective, therefore, J is also a bijection (which we will identify with the identity). □

5.6 Quadratic forms

Let $\mathcal{V}, \mathcal{W}$ be complex vector spaces.

Definition 5.24 $\mathfrak{a}$ *is called a sesquilinear form on $\mathcal{W} \times \mathcal{V}$ iff it is a map*

$$\mathcal{W} \times \mathcal{V} \ni (w, v) \mapsto \mathfrak{a}(w, v) \in \mathbb{C}$$

antilinear wrt the first argument and linear wrt the second argument.

If $\lambda \in \mathbb{C}$, then λ can be treated as a sesquilinear form $\lambda(w,v) := \lambda(w|v)$. If $\mathfrak{a}$ is a form, then we define $\lambda\mathfrak{a}$ by $(\lambda\mathfrak{a})(w,v) := \lambda\mathfrak{a}(w,v)$. and $\mathfrak{a}^*$ by $\mathfrak{a}^*(v,w) := \overline{\mathfrak{a}(w,v)}$. If $\mathfrak{a}_1$ and $\mathfrak{a}_2$ are forms, then we define $\mathfrak{a}_1 + \mathfrak{a}_2$ by $(\mathfrak{a}_1 + \mathfrak{a}_2)(w,v) := \mathfrak{a}_1(w,v) + \mathfrak{a}_2(w,v)$.

Suppose that $\mathcal{V} = \mathcal{W}$. We will write $\mathfrak{a}(v) := \mathfrak{a}(v,v)$. We will call it a quadratic form. The knowledge of $\mathfrak{a}(v)$ determines $\mathfrak{a}(w,v)$:

$$\mathfrak{a}(w,v) = \frac{1}{4}\left(\mathfrak{a}(w+v) + \mathrm{i}\mathfrak{a}(w-\mathrm{i}v) - \mathfrak{a}(w-v) - \mathrm{i}\mathfrak{a}(w+\mathrm{i}v)\right). \tag{5.16}$$

Suppose now that $\mathcal{V}, \mathcal{W}$ are Hilbert spaces. A form is bounded iff

$$|\mathfrak{a}(w,v)| \leq C\|w\|\|v\|.$$

Proposition 5.25 (1) *Let $\mathfrak{a}$ be a bounded sesquilinear form on $\mathcal{W} \times \mathcal{V}$. Then there exists a unique operator $A \in B(\mathcal{V}, \mathcal{W})$ such that*

$$\mathfrak{a}(w,v) = (w|Av).$$

(2) *If $A \in B(\mathcal{V}, \mathcal{W})$, then $(w|Av)$ is a bounded sesquilinear form on $\mathcal{W} \times \mathcal{V}$.*

Proof. (2) is obvious. To show (1) note that $w \mapsto \mathfrak{a}(w|v)$ is an antilinear functional on $\mathcal{W}$. Hence there exists $\eta \in \mathcal{W}$ such that $\mathfrak{a}(w,v) = (w|\eta)$. We put $Av := \eta$.

Theorem 5.26 *Suppose that $\mathcal{D}, \mathcal{Q}$ are dense linear subspaces of $\mathcal{V}, \mathcal{W}$ and $\mathfrak{a}$ is a bounded sesquilinear form on $\mathcal{D} \times \mathcal{Q}$. Then there exists a unique extension of $\mathfrak{a}$ to a bounded form on $\mathcal{V} \times \mathcal{W}$.*

5.7 Adjoint operators

Definition 5.27 *Let $A \in B(\mathcal{V}, \mathcal{W})$. Then the operator A^* given (uniquely) by the formula*

$$(A^*w|v) := (w|Av)$$

is called the (hermitian) conjugate of A.

Note that the definition is correct, because $\mathfrak{a}(w,v) := (w|Av)$ is a bounded sesquilinear form, and hence so is $\mathfrak{a}^*$; and A^* is the operator associated with $\mathfrak{a}^*$.

Theorem 5.28 *The hermitian conjugation has the following properties*

1) $\|A^*\| = \|A\|$
2) $(\lambda A)^* = \overline{\lambda} A^*$
3) $(A+B)^* = A^* + B^*$,
4) $(AB)^* = B^* A^*$,
5) $A^{**} = A$,
6) $(\mathrm{Ran}\, A)^{\perp} = \mathrm{Ker} A^*$,
7) $(\mathrm{Ran}\, A^*)^{\perp} = \mathrm{Ker} A$,
8) A *is invertible* $\Leftrightarrow A^*$ *is invertible* $\Leftrightarrow \|Av\| \geq C\|v\|$ *and* $\|A^*v\| \geq C\|v\|$, *moreover,*
   $(A^{-1})^* = (A^*)^{-1}$.
9) $\mathrm{sp} A^* = \overline{\mathrm{sp} A}$.

5.8 Numerical range

Definition 5.29 *Let* $\mathfrak{t}$ *be a quadratic form on* $\mathcal{X}$*. The numerical range of* $\mathfrak{t}$ *is defined as*

$$\mathrm{Num}\,\mathfrak{t} := \{\mathfrak{t}(x) \in \mathbb{C} \;:\; x \in \mathcal{X},\; \|x\| = 1\}.$$

Theorem 5.30 (1) *In a twodimensional space the numerical range of a quadratic form is always an elipse together with its interior.*

(2) $\mathrm{Num}\,\mathfrak{t}$ *is a convex set.*

(3) $\mathrm{Num}(\alpha\mathfrak{t} + \beta) = \alpha\mathrm{Num}(\mathfrak{t}) + \beta$.

(4) $\mathrm{Num}\,\mathfrak{t}^* = \overline{\mathrm{Num}\,\mathfrak{t}}$.

(5) $\mathrm{Num}(\mathfrak{t} + \mathfrak{s}) \subset \mathrm{Num}\,\mathfrak{t} + \mathrm{Num}\,\mathfrak{s}$.

Proof. (1) We write Let $\mathfrak{t}(v) = \mathrm{Re}\mathfrak{t}(v) + \mathrm{Im}\mathfrak{t}(v)$. We diagonalize the imaginary part of $\mathfrak{t}$. Thus if $\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ is the matrix of $\mathfrak{t}$, then $a_{12} = \overline{a}_{21}$. By multiplying one of the basis vectors with a phase factor we can guarantee that $a_{12} = a_{21}$ is real.

Now $\mathfrak{t}$ is given by a matrix of the form

$$c\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} \lambda & \mu \\ \mu & -\lambda \end{bmatrix} + \mathrm{i}\begin{bmatrix} \gamma & 0 \\ 0 & -\gamma \end{bmatrix}$$

Any normalized vector up to a phase factor equals $v = (\cos\alpha, \mathrm{e}^{\mathrm{i}\phi}\sin\alpha)$ and

$$\mathfrak{t}(v) - c = \lambda\cos 2\alpha + \mu\cos\phi\sin 2\alpha + \mathrm{i}\gamma\cos 2\alpha \tag{5.17}$$

Now it is an elementary exercise to check that the range of $x + \mathrm{i}y$ given by (5.17), equals

$$(\gamma x - \lambda y)^2 + \mu^2 y^2 \leq \gamma^2\mu^2.$$

(2) follows immediately from (1). □

Let $\mathcal{V}$ be a Hilbert space. If A is an operator on $\mathcal{V}$, then the numerical range of A is defined as the numerical range of the form $(v|Aw)$, that is

$$\mathrm{Num}A := \{(v|Av) \in \mathbb{C} \;:\; v \in \mathcal{V},\; \|v\| = 1\}.$$

Theorem 5.31 *Let* $A \in B(\mathcal{V})$*. Then*

(1) $\mathrm{sp}A \subset (\mathrm{Num}A)^{\mathrm{cl}}$.

(2) *For* $z \notin (\mathrm{Num}A)^{\mathrm{cl}}$,

$$\|(z - A)^{-1}\| \leq \mathrm{dist}(z, \mathrm{Num}A)^{-1}.$$

Proof. Let $(z_0 \notin \mathrm{Num}A)^{\mathrm{cl}}$. Recall that $\mathrm{Num}(A)$ is convex. Hence, replacing A wih $\alpha A + \beta$ we can assume that $z_0 = \mathrm{i}\nu$ with $\nu = \mathrm{dist}(z, \mathrm{Num}(A))$ and $\mathrm{Num}A \subset \{\mathrm{Im}z \leq 0\}$. Now

$$\begin{aligned} \|(z_0 - A)v\|^2 &= (Av|Av) + i\nu(v|Av) - \mathrm{i}\nu(Av|v) + |\nu|^2\|v\|^2 \\ &= (Av|Av) - 2\nu\mathrm{Im}(v|Av) + |\nu|^2\|v\|^2 \\ &\geq |\nu|^2\|v\|^2. \end{aligned}$$

Next, $\mathrm{Num}A^* \subset \{\mathrm{Im}z \geq 0\}$.

$$\begin{aligned}\|(\overline{z}_0 - A^*)v\|^2 &= (A^*v|A^*v) - i\nu(v|A^*v) + \mathrm{i}\nu(A^*v|v) + |\nu|^2\|v\|^2 \\ &= (A^*v|A^*v) - 2\nu\mathrm{Im}(v|Av) + |\nu|^2\|v\|^2 \\ &\geq |\nu|^2\|v\|^2.\end{aligned}$$

Hence $z_0 - A$ is invertible and $z \in \mathrm{rs}A$. □

5.9 Self-adjoint operators

Theorem 5.32 *Let $A \in B(\mathcal{V})$. The following conditions are equivalent:*

(1) $A = A^*$.

(2) $(Aw|v) = (w|Av),\ \ w, v \in \mathcal{V}$.

(3) $(w|Av) = \overline{(v|Aw)},\ \ w, v \in \mathcal{V}$.

(4) $(v|Av) \in \mathbb{R}$.

Proof. (1)⇔(2)⇔(3)⇒(4) is obvious. To show (4)⇒(3) we use the polarization identity:

$$\begin{aligned}(w|Av) &= \tfrac{1}{4}\sum_{j=0}^{3}(-\mathrm{i})^j(w + \mathrm{i}^j v|A(w + \mathrm{i}^j v)), \\ \overline{(v|Aw)} &= \tfrac{1}{4}\overline{\sum_{j=0}^{3}(-\mathrm{i})^j(v + \mathrm{i}^j w|A(v + \mathrm{i}^j w))} \\ &= \tfrac{1}{4}\sum_{j=0}^{3}(-\mathrm{i})^j(w + \mathrm{i}^j v|A(w + \mathrm{i}^j v)).\end{aligned}$$

□

Definition 5.33 *An operator $A \in B(\mathcal{V})$ satisfying the conditions of Theorem 5.32 is called self-adjoint. An operator $A \in B(\mathcal{V})$ such that $(v|Av) \geq 0$ is called a positive operator.*

By Theorem 5.32, positive operators are self-adjoint.

Clearly, if $A \in B(\mathcal{V})$, then A is self-adjoint iff $\mathrm{Num}A \subset \mathbb{R}$ and positive iff $\mathrm{Num}A \subset [0, \infty[$.

Theorem 5.34 *Let A be self-adjoint. Then* $\mathrm{sp}A \subset \mathbb{R}$.

Proof. Let $\mu \neq 0$, $\mu, \lambda \in \mathbb{R}$. We have

$$\|(A - (\lambda + \mathrm{i}\mu))v\|^2 = \|(A - \lambda)v\|^2 + \mu^2\|v\|^2 \geq \mu^2\|v\|^2.$$

Besides, $(A - (\lambda + i\mu))^* = A - (\lambda - i\mu)$. Hence

$$\|(A - (\lambda + \mathrm{i}\mu))^* v\|^2 = \|(A - \lambda)v\|^2 + \mu^2\|v\|^2 \geq \mu^2\|v\|^2.$$

So $A - (\lambda + \mathrm{i}\mu)$ is invertible. □.

Theorem 5.35 *The operator A^*A is positive and*

$$\|A^*A\| = \|A\|^2.$$

Proof. A^*A is positive because

$$(v|A^*Av) = \|Av\|^2 \geq 0.$$

$$\begin{aligned}\|A\|^2 = \|A^*\|\|A\| \geq \|A^*A\| &\geq \sup_{\|v\|=1}(v|A^*Av)\\ &= \sup_{\|v\|=1}\|Av\|^2 = \|A\|^2.\end{aligned}$$

□

The following facts will follow immediately from the spectral theorem. In particular, Theorem 5.18 will follow for an arbitrary normal operator.

Lemma 5.36 *Let A be self-adjoint. Then*

$$\|A\| = \sup_{\|v\|\leq 1} |(v|Av)|.$$

Proof. Let $w, v \in \mathcal{V}$. We will show first that

$$|(w|Av)| \leq \frac{1}{2}(\|w\|^2 + \|v\|^2) \sup_{\|y\|\leq 1}(y|Ay).$$

Replacing w with $\mathrm{e}^{i\alpha}w$ we can suppose that $(w|Av)$ is positive. Then

$$\begin{aligned}&(w|Av) = \tfrac{1}{2}((w|Av) + (v|Aw))\\ &= \tfrac{1}{4}\left((w+v|A(w+v) - (w-v|A(w-v))\right)\\ &\leq \tfrac{1}{4}\left(\|v+w\|^2 + \|v-w\|^2\right)\sup_{\|y\|=1}|(y|Ay)|\\ &= \tfrac{1}{2}\left(\|v\|^2 + \|w\|^2\right)\sup_{\|y\|=1}|(y|Ay)|\end{aligned}$$

Hence $\|w\| = \|v\| = 1$ implies

$$|(w|Av)| \leq \sup_{\|y\|=1} |(y|Ay)|.$$

But

$$\|A\| = \sup_{\|v\|=\|w\|=1} |(w|Av)|.$$

□

Theorem 5.37 *If A is self-adjoint, then*

$$(\mathrm{Num}A)^{\mathrm{cl}} = \mathrm{CH}(\mathrm{sp}A). \tag{5.18}$$

Proof. Step 1. Let A be self-adjoint and

$$-\inf(\mathrm{sp}A) = \sup(\mathrm{sp}A) =: a. \tag{5.19}$$

Clearly, $\mathrm{CH}(\mathrm{sp}A) = [-a, a]$ and $a = \|A\|$. By Lemma 5.36, $(\mathrm{Num}A)^{\mathrm{cl}} \subset [-a, a]$. Hence, $(\mathrm{Num}A)^{\mathrm{cl}} \subset \mathrm{CH}(\mathrm{sp}A)$. The converse inclusion follows from Theorem 5.31.

Step 2. Let A be self-adjoint. Let $a_- := \inf(\mathrm{sp}A)$, $a_+ := \sup(\mathrm{sp}A)$. Then $\tilde{A} := A - \frac{1}{2}(a_- + a_+)$ is self-adjoint and satisfies (5.19). Hence (5.18) holds for $\tilde{A}$. Hence (5.18) holds for A as well. □

5.10 Projectors

Theorem 5.38 *Let $P \in B(\mathcal{V})$ be an idempotent. The following conditions are equivalent:*

(1) *P is self-adjoint.*

(2) $\mathrm{Ker}P = (\mathrm{Ran}\, P)^{\perp}$.

An idempotent P satisfing these conditions with $\mathrm{Ran}\, P = \mathcal{W}$ will be called the projector onto $\mathcal{W}$.

If $(w_i)_{i\in I}$ is an orthogonal basis in $\mathcal{W}$, then

$$Pv = \sum_{i\in I} \frac{(w_i|v)}{\|w_i\|^2} w_i.$$

Proposition 5.39 (Gramm-Schmidt ortogonalization) *Let $y_1, y_2, \dots$ be a linearly independent system. Let P_n be a projection onto the (n-dimensional) space $\mathrm{Span}\{y_1, \dots, y_n\}$. Then*

$$w_n := (1 - P_{n-1}) y_n$$

is an orthogonal system. An equivalent definition:

$$w_1 = y_1, \qquad w_n := y_n - \sum_{j=1}^{n-1} \frac{(w_j|y_n)}{\|w_j\|^2} w_j.$$

Theorem 5.40 *Let $P^* = P$ and $P^2 = P^3$. Then P is a projector.*

Proof. $(P^2 - P)^*(P^2 - P) = 0$, hence $P = P^2$. □

5.11 Orthogonal polynomials

Theorem 5.41 *Let μ be a Borel measure on $\mathbb{R}$ such that for some $\epsilon > 0$*

$$\int \mathrm{e}^{\epsilon|x|} \mathrm{d}\mu(x) < \infty.$$

Then polynomials are dense in $L^2(\mathbb{R}, \mu)$.

Proof. Let $\phi \in L^2(\mathbb{R}, \mu)$. Clearly, for $\mathrm{Re}z| < \epsilon$, $\phi(x)\mathrm{e}^{z|x|}$ be;ongs to $L^2(\mathbb{R}, \mu)$. Define $F(z) := \int \mathrm{e}^{-zx}\phi(x)\mathrm{d}\mu(x)$. It is analytic for $|\mathrm{Re}z| < \epsilon$ Now

$$(x^n|\phi) = \int x^n \phi(x) \mathrm{d}\mu(x) = (-\mathrm{i})^n \frac{\mathrm{d}^n}{\mathrm{d}z^n} F(0).$$

If ϕ is orthogonal to polynomials, all the derivatives of F at zero vanish. Hence $F(z) = 0$ in the whole strip. In particular, $F(\mathrm{i}y) = 0$, $y \in \mathbb{R}$. Therefore, $\phi(x)\mathrm{d}\mu(x)$ is zero. □

5.12 Isometries and partial isometries

Definition 5.42 *An operator U is called a partial isometry if U^*U and are UU^* projectors.*

Theorem 5.43 *U is a partial isometry iff U^*U is a projector.*

Proof. We check that $(UU^*)^3 = (UU^*)^2$. □

Proposition 5.44 *If U is a partial isometry, then UU^* is a projector onto* $\mathrm{Ran}\, U$ *and U^*U is the projector onto* $(\mathrm{Ker}U)^\perp$.

Proof.

$$\begin{aligned} v\in \mathrm{Ker}U &\Leftrightarrow Uv=0\\ &\Leftrightarrow 0=(Uv|UV)=(v|U^*Uv)=(v|U^*UU^*Uv)=(U^*Uv|U^*Uv)\\ &\Leftrightarrow v\in\mathrm{Ker}U^*U. \end{aligned}$$

This proves that U^*U is the projector onto $(\mathrm{Ker}U)^\perp$. □

Theorem 5.45 *Let $U\in B(\mathcal{V},\mathcal{W})$. The following properties are equivalent:*
1) $U^*U=1$,
2) $(Uv|Uw)=(v|w)$, $v,w\in\mathcal{V}$,
3) U is an isometry, that means $\|Uv\|=\|v\|$.

Definition 5.46 *An operator U satisfying the properties of Theorem 5.45 is called a linear isometry.*

Proof. 1)⇔2) is obvious, and so is2)⇒3). 3)⇒2) follows by the polarization identity:

$$\begin{aligned} (Uw|Uv) &= \tfrac{1}{4}\sum_{j=0}^{3}(-\mathrm{i})^j(Uw+\mathrm{i}^jUv|Uw+\mathrm{i}^jUv),\\ (w|v) &= \tfrac{1}{4}\sum_{j=0}^{3}(-\mathrm{i})^j(v+\mathrm{i}^jw|v+\mathrm{i}^jw). \end{aligned}$$

□

Theorem 5.47 *Let V be isometric. Then* $\mathrm{sp}V\subset\{|z|\leq 1\}$.

Proof. We have $\|V\|^2=\|V^*V\|=\|1\|=1$. Hence, $\mathrm{sp}V\subset\{|z|\leq 1\}$.□

5.13 Unitary operators

Theorem 5.48 *Let $U\in B(\mathcal{V},\mathcal{W})$. The following properties are equivalent:*
1) $U^*U=UU^*=1$;
2) U is a surjective isometry;
3) U is bijective and $U^*=U^{-1}$.

Definition 5.49 *An operator satisfing the properties of Theorem 5.48 is called unitary.*

Proposition 5.50 *Let $\mathcal{V}$ be finite dimensional and $V\in B(\mathcal{V})$ isometric. Then V is unitary.*

Proof. We have $\dim\mathrm{Ker}V+\dim\mathrm{Ran}\,V=\dim\mathcal{V}$. $\mathrm{Ker}V=\{0\}$, since V is isometric. Hence $\dim\mathrm{Ran}\,V=\dim\mathcal{V}$. But $\mathcal{V}$ is finite dimensional, hence $\mathrm{Ran}\,V=\mathcal{V}$. □

Example 5.51 *Let (e_i), $i=1,2,\dots$ be the canonical basis in $L^2(\mathbb{N})$. Put*

$$Te_i := e_{i+1}.$$

Then T is isometric but not unitary. It is called the unitaral shift.

Theorem 5.52 (1) *U is unitary iff U is normal and* $\mathrm{sp}U\subset\{z\ :\ |z|=1\}$.

(2) *A is self-adjoint iff A is normal and* sp$A \subset \mathbb{R}$.

Proof. (1)$\Rightarrow$. Clearly, U is normal.

U is an isometry, hence sp$U \subset \{|z| \leq 1\}$.

U^{-1} is also an isometry, hence sp$U^{-1} \subset \{|z| \leq 1\}$. This implies sp$U \subset \{|z| \geq 1\}$.

(1)$\Leftarrow$ Since U is normal and $|z| = 1$ on sp$U$, by the spectral mapping theorem, $\|U\| = 1$. U^{-1} is normal as well and by the spectral mapping theorem $|z| = 1$ on spU^{-1}, hence $\|U^{-1}\| = 1$.

Suppose that $\|Uv\| < \|v\|$. Then for $w := Uv$, $\|Uw\| > \|w\|$. Thus U and U^{-1} are isometries. Hence U is unitary.

(2)$\Rightarrow$ was proven in Theorem 5.34. Let us prove (2)$\Leftarrow$ Let A be normal and sp$A \subset \mathbb{R}$. We can find $\lambda > 0$ such that $\lambda\|A\| < 1$. Hence $1 + \mathrm{i}\lambda A$ is invertible. It is easy to check that $U := (1 - \mathrm{i}\lambda A)(1 + \mathrm{i}\lambda A)^{-1}$ is normal. By the spectral mapping theorem, sp$U \subset \{|z| = 1\}$. Hence, by (1), it is unitary. Now

$$\begin{aligned} A = -\mathrm{i}\lambda^{-1}(1 - U)(1 + U) \quad &= \mathrm{i}\lambda^{-1}(U - U^*U)(U + U^*U)^{-1} \\ &= \mathrm{i}\lambda^{-1}(1 - U^*)(1 + U^*)^{-1} = A^*. \end{aligned}$$

□

5.14 Convergence

Let (A_λ) be a net of operators in $B(\mathcal{V}, \mathcal{W})$.

(1) We say that (A_λ) is norm convergent to A iff $\lim_\lambda \|A_\lambda - A\| = 0$. In this case we write

$$\lim_\lambda A_\lambda = A.$$

(2) We say that (A_λ) is strongly convergent to A iff $\lim_\lambda \|A_\lambda v - Av\| = 0$, $v \in \mathcal{V}$. In this case we write

$$\mathrm{s-}\lim_\lambda A_\lambda = A.$$

(3) We say that (A_λ) is weakly convergent to A iff $\lim_\lambda |(w|A_\lambda v) - (w|Av)| = 0$, $v \in \mathcal{V}$, $w \in \mathcal{W}$. In this case we write

$$\mathrm{w-}\lim_\lambda A_\lambda = A.$$

Theorem 5.53 *Let (U_λ) be a net of unitary operators*

(1) *If (U_λ) is norm convergent, then its limit is unitary.*

(2) *If (U_λ) is strongly convergent, then its limit is isometric.*

(3) *If (U_λ) is weakly convergent, then its limit is a contraction.*

Theorem 5.54 *Let (A_λ) be a net of operators in $B(\mathcal{V})$ weakly convergent to A. Then*

$$\mathrm{Num}A \subset \bigcap_{\mu \in \Lambda} \bigcup_{\lambda > \mu} \mathrm{Num}A_\lambda.$$

In particular, if A_λ are self-adjoint, then so is A; if A_λ are positive, then so is A.

Theorem 5.55 *Let (A_n) be a weakly convergent sequence of operators in $B(\mathcal{V})$. Then it is uniformly bounded.*

Proof. This follows from the uniform boundedness principle. □

Example 5.56 *In* $L^2(\mathbb{N})$*, let* $(e_1, e_2, \dots)$ *be the canonical basis. Set*

$$\begin{aligned} U_n e_j &= e_{j+1}, \quad j = 1, \dots, n-1; \\ U_n e_n &= e_1; \\ U_n e_j &= e_j, \quad j = n+1, \dots; \\ U e_j &= e_{j+1}, \quad j = 1, \dots. \end{aligned}$$

Then U_n *are unitary,* $\mathrm{s-}\lim_{n\to\infty} U_n = U$ *is not. Moreover.* $\mathrm{sp}U_n = \{\exp(\mathrm{i}2\pi/n) \;:\; j = 1, \dots, n\}$ *and* $\mathrm{sp}U = \{|z| \leq 1\}$.

Example 5.57 *In* $L^2(\mathbb{Z})$*, let* e_i*,* $i \in \mathbb{Z}$ *be the canonical basis. Set* $U_n \mathrm{e}_j = e_{j+n}$*,* $j \in \mathbb{Z}$*. Then* U_n *are unitary,* $\mathrm{w-}\lim_{n\to\infty} U_n = 0$*. Moreover,* $\mathrm{sp}U_n = \{|z| = 1\}$*,* $\mathrm{sp}U = \{0\}$.

5.15 Monotone convergence of selfadjoint operators

Theorem 5.58 *Let* $\{A_\lambda \;:\; \lambda \in \Lambda\}$ *be an increasing net of self-adjoint operators, which is uniformly bounded. Then there exists the smallest self-adjoint operator* A *such that* $A_\lambda \leq A$*. We will denote it* $\mathrm{lub}A_\lambda$ *(the least upper bound). We have*

$$A = \mathrm{s-}\lim_{\lambda} A_\lambda.$$

Proof. Let $\|A_\lambda\| \leq c$. For each $v \in \mathcal{V}$, $(v|A_\lambda v)$ is an increasing net bounded by $c\|v\|^2$. Hence it is convergent. Using the polarization identity we obtain the convergence of $(v|A_\lambda w)$. Thus we obtain a sesquilinear form

$$\lim_{\lambda} (v|A_\lambda w) \tag{5.20}$$

It is bounded by c, hence it is given by a bounded operator, which we denote by A, so that (5.20) equals $(v|Aw)$. It is evident that A is the smallest self-adjoint operator greater than A_λ.

Since $A - A_\lambda \geq 0$, we have

$$(A - A_\lambda)^2 = (A - A_\lambda)^{\frac{1}{2}} (A - A_\lambda)(A - A_\lambda)^{\frac{1}{2}} \leq \|A - A_\lambda\|(A - A_\lambda).$$

Besides, $\|A - A_\lambda\| \leq 2c$. Now

$$\|(A - A_\lambda)v\|^2 = (v|(A - A_\lambda)^2 v) \leq \|A - A_\lambda\|(v|(A - A_\lambda)v) \to 0.$$

□

6 Spectral theorems

In this section we prove various forms of the spectral theorem. We avoid using the Gelfand theory, which makes our approach slightly more elementary than that of most contemporary literature.

6.1 Holomorphic spectral theorem for normal operators

Theorem 6.1 *Let $A \in B(\mathcal{V})$ be normal. Then*

$$\mathrm{sr}(A) = \|A\|.$$

Proof.

$$\|A^2\|^2 = \|A^{2*}A^2\| = \|(A^*A)^2\| = \|A^*A\|^2 = \|A\|^4.$$

Thus $\|A^{2^n}\| = \|A\|^{2^n}$. Hence, using the formula for the spectral radius of A we get $\|A^{2^n}\|^{2^{-n}} = \|A\|$. □

If K is a compact subset of $\mathbb{C}$ let $C_{\mathrm{hol}}(K)$ be the completion of $\mathrm{Hol}(K)$ in $C(K)$.

The following version of the spectral theorem follows easily from Theorem 6.1. It will be improved in next subsection so that the functional calculus will be defined on the whole $C(\mathrm{sp}A)$.

Note that in the case A is self-adjoint or unitary, $C(\mathrm{sp}A) = C_{\mathrm{hol}}(\mathrm{sp}A)$.

Theorem 6.2 *Let $A \in B(\mathcal{V})$ be normal. Then there exists a unique continuous isomorphism*

$$C_{\mathrm{hol}}(\mathrm{sp}(A)) \ni f \mapsto f(A) \in B(\mathcal{V}),$$

such that

(1) $\mathrm{id}(A) = A$ *if* $\mathrm{id}(z) = z$.
Moreover, we have

(2) *If* $f \in \mathrm{Hol}(\mathrm{sp}(A))$, *then* $f(A)$ *coincides with* $f(A)$ *defined in (4.6).*

(3) $\mathrm{sp}(f(A)) = f(\mathrm{sp}(A))$.

(4) $g \in C_{\mathrm{hol}}(f(\mathrm{sp}(A))) \Rightarrow g \circ f(A) = g(f(A))$.

(5) $\|f(A)\| = \sup |f|$.

6.2 Commutative C^*-algebras

Let X be a compact Hausdorff space. Then $C(X)$ with the norm $\|\cdot\|_\infty$ is a commutative C^*-algebra.

Note that if $A \subset U \subset X$ where U is open and A is closed, then there exists $F \in C(X)$ with $F = 1$ on A, $0 \leq F \leq 1$ and $\{F \neq 0\} \subset U$.

Let $Y \subset X$ and Let $C_Y(X)$ denote the set of functions vanishing on Y.

The following fact is well known from topology.

Theorem 6.3 *Let X be a compact Hausdorff space.*

(1) *Let Y be a closed subset of X. Then $C_Y(X)$ is a closed ideal of $C(X)$.*

(2) *Let $\mathfrak{N}$ be a closed ideal of $C(X)$. Set*

$$Y := \bigcap_{F \in \mathfrak{N}} F^{-1}(0).$$

Then Y is closed and $\mathfrak{N} = C_Y(X)$.

(3)

$$C(X)/C_Y(X) \ni F + C_Y(X) \mapsto F\Big|_Y \in C(Y)$$

is an isometric $*$-homomorphism.

6.3 Spectrum of a $*$-homomorphisms of $C(X)$

Let X be a compact Hausdorff space. Let $\mathcal{V}$ be a Hilbert space and $\gamma : C(X) \to B(\mathcal{V})$ a homomorphism. We say that it is a $*$-homomorphism iff $\gamma(\overline{F}) = \gamma(F)^*$.

Theorem 6.4 *Every $*$-homomorphism is a contraction.*

Proof. Let $z \notin F(X)$. Then $(z - F)^{-1} \in C(X)$. Thus $\gamma((z - F)^{-1})$ is the inverse of $z - \gamma(F)$. Thus $\mathrm{sp}\gamma(F) \subset F(X)$, and hence $\mathrm{sr}\gamma(F) \leq \|F\|_\infty$.

Clearly, $\gamma(F)$ is normal, and hence $\|\gamma(F)\| = \mathrm{sr}\gamma(F)$. □

In what follows $\gamma : C(X) \to B(\mathcal{V})$ is a unital $*$-homomorphism.

We define the spectrum of the homomorphism γ as

$$\mathrm{sp}\gamma = \bigcap_{F \in \mathrm{Ker}\gamma} F^{-1}(0).$$

Clearly, $\mathrm{sp}\gamma$ is a closed subset of X and $\mathrm{Ker}\gamma = C_{\mathrm{sp}\gamma}(X)$.

Theorem 6.5 *Let $F \in C(X)$ Then the following is true:*

(1) $F(\mathrm{sp}\gamma) = \mathrm{sp}\gamma(F)$.

(2) $\|\gamma(F)\| = \mathrm{sr}\gamma(F) = \|F\big|_{\mathrm{sp}\gamma}\|_\infty$.

(3) *γ is injective iff $X = \mathrm{sp}\gamma$ and then it is isometric.*

(4) *There exists a unique $*$-isomorphism $\gamma_{\mathrm{red}} : C(\mathrm{sp}\gamma) \to B(\mathcal{V})$ such that*

$$\gamma(F) = \gamma_{\mathrm{red}}\left(F\Big|_{\mathrm{sp}\gamma}\right), \quad F \in C(X)$$

γ_{red} *is injective.*

Proof. (1) $F(\mathrm{sp}\gamma) \subset \mathrm{sp}\gamma(F)$.

Suppose that $z \in \mathrm{rs}(\gamma(F))$. Then there exists $A \in B(\mathcal{V})$ with $A\gamma(z - F) = 1$. Let $x \in X$ with $F(x) = z$. We can find $G \in C(X)$ such that $0 \leq G \leq 1$, $G = 1$ on a neighborhood U of x and on $\mathrm{supp}G$ we have $|F - z| < \frac{1}{2}\|A\|^{-1}$. Then $A\gamma((z - F)G) = \gamma(G)$, hence

$$\begin{aligned} \|\gamma(G)\| &\leq \|A\|\|\gamma((z - F)G)\| \\ &\leq \|A\|\|(z - F)G\|_\infty < \tfrac{1}{2}. \end{aligned}$$

Let $H \in C(X)$, $\mathrm{supp}H \subset U$, $H(x) = 1$. Then for any n, $H = HG^n$. Hence

$$\gamma(H) = \gamma(H)\gamma(G)^n.$$

Therefore,

$$\|\gamma(H)\| \leq \|\gamma(H)\|2^{-n} \to 0.$$

Thus $\|\gamma(H)\| = 0$. Hence $\gamma(H) = 0$. But $x \notin H^{-1}(0)$. Hence $x \notin \mathrm{sp}\gamma$. This proves (2).

(1) $F(\mathrm{sp}\gamma) \supset \mathrm{sp}\gamma(F)$.

Let $z \notin F(\mathrm{sp}\gamma)$. $Z := \{x \in X \ : \ F(x) = z\}$ is a closed subset of X disjoint from $\mathrm{sp}\gamma$. Hence, there exists a function $G \in C(X)$ such that $G = 1$ on $\mathrm{sp}\gamma$ and $G = 0$ on a neighborhood of Z. Clearly, $G - 1 \in C_{\mathrm{sp}\gamma}(X)$, hence $\gamma(G) = 1$. Now $G(z - F)^{-1} \in C(X)$. We have

$$\gamma(z - F)\gamma((z - F)^{-1}G) = \gamma(G) = 1.$$

Hence $\gamma((z - F)^{-1}G)$ is the inverse of $z - \gamma(F)$. This means that $z \in \mathrm{rs}\gamma(F)$.

(2) follows from (1) and the normality of $\gamma(F)$.

The kernel of γ is a closed ideal of $C(X)$, hence it equals $C_Y(X)$ for some closed $Y \subset X$. γ is injective iff its kernel equals $\{0\}$. $C_Y(X) = C(X)$ iff $Y = X$. This together with (2) proves (3). □

6.4 Functional calculus for a single normal operator

Lemma 6.6 *Let* $(A_i\ :\ i\in I)$ *be a family of self-adjoint commuting operators in a* $B(\mathcal{V})$*. Then there exists a unique unital* $*$*-homomorphism*

$$C\Big(\underset{i\in I}{\times}\,\mathrm{sp}A_i\Big)\ni F\mapsto F(A_i\ :\ i\in I)\in B(\mathcal{V}) \tag{6.21}$$

such that $\mathrm{id}_j(A_i\ :\ i\in I)=A_i$*, where* $\mathrm{id}_j(x_i:i\in I)=x_j$*.*

Proof. On polynomials we define (6.21) in the obvious way. By the weak spectral mapping theorem of Theorem 4.12,

$$\mathrm{sp}F(A_i\ :\ i\in I)\subset F\Big(\underset{i\in I}{\times}\,\mathrm{sp}A_i\Big).$$

Hence, $\mathrm{sr}F(A_i\ :\ i\in I)\le\|F\|_\infty$. But $F(A_i\ :\ i\in I)$ is normal and hence $\|F(A_i\ :\ i\in I)\|\le\|F\|_\infty$.

By the Stone-Weierstrass Theorem, polynomials are dense in continuous functions, therefore we can extend the definition of (6.21) from polynomials to $C\Big(\underset{i\in I}{\times}\,\mathrm{sp}A_i\Big)$. □

We define the joint spectrum of the family of operators $(A_i\ :\ i\in I)$, denoted $\mathrm{sp}(A_i\ :\ i\in I)$, to be the spectrum of the homomorphism (6.21).

Theorem 6.7 *Let* $A\in B(\mathcal{V})$ *be normal. Then there exists a unique continuous isomorphism*

$$C(\mathrm{sp}(A))\ni f\mapsto f(A)\in B(\mathcal{V}), \tag{6.22}$$

such that

(1) $\mathrm{id}(A)=A$ *if* $\mathrm{id}(z)=z$.
Moreover, we have

(2) *If* $f\in\mathrm{Hol}(\mathrm{sp}(A))$*, then* $f(A)$ *coincides with* $f(A)$ *defined in (4.6).*

(3) $\mathrm{sp}(f(A))=f(\mathrm{sp}(A))$.

(4) $g\in C(f(\mathrm{sp}(A)))\Rightarrow g\circ f(A)=g(f(A))$.

(5) $\|f(A)\|=\sup|f|$.

Proof. $B^{\mathrm{R}}:=\frac{1}{2}(B+B^*)$ and $B^{\mathrm{I}}:=\frac{1}{2\mathrm{i}}(B-B^*)$ are commuting self-adjoint operators. Therefore, we have the homomorphisms

$$C(\mathrm{sp}B^{\mathrm{R}})\ni f\mapsto f(B^{\mathrm{R}})\in B(\mathcal{V}),$$
$$C(\mathrm{sp}B^{\mathrm{I}})\ni f\mapsto f(B^{\mathrm{I}})\in B(\mathcal{V}),$$

with commuting ranges. We can construct the product of these homomorphism,

$$C(\mathrm{sp}B^{\mathrm{R}}\times\mathrm{sp}B^{\mathrm{I}})\ni F\mapsto F(B^{\mathrm{R}},B^{\mathrm{I}}). \tag{6.23}$$

Define $\mathbb{R}^2\ni(x,y)\mapsto j(x,y):=x+\mathrm{i}y\in\mathbb{C}$. We claim that

$$j(\mathrm{sp}(B^{\mathrm{R}},B^{\mathrm{I}}))=\mathrm{sp}B.$$

Let $(x_0,y_0)\notin\mathrm{sp}(B^{\mathrm{R}},B^{\mathrm{I}})$. The function

$$(x,y)\mapsto(x_0+\mathrm{i}y_0-x-\mathrm{i}y)^{-1}$$

is well defined outside of (x_0,y_0). In particular, it s well defined on $\mathrm{sp}(B^{\mathrm{R}},B^{\mathrm{I}})$. Hence

$$(x_0+\mathrm{i}y_0-B^{\mathrm{R}}-\mathrm{i}B^{\mathrm{I}})^{-1}=(x_0+\mathrm{i}y_0-B)^{-1}$$

exists. Therefore, $x_0 + \mathrm{i}y_0 \in \mathrm{rs}(B)$.

Let $x_0 + \mathrm{i}y_0 \notin \mathrm{sp}B$ and $(x_0, y_0) \in \mathrm{sp}(B^{\mathrm{R}}, B^{\mathrm{I}})$. Let $\epsilon > 0$ and $f \in C(\mathrm{sp}(B^{\mathrm{R}}, B^{\mathrm{I}}))$ with $f(x_0, y_0) = 1$ and $\{f \neq 0\} \subset \{(x - x_0)^2 + (y - y_0)^2 < \epsilon^2\}$. Clearly,

$$f(B^{\mathrm{R}}, B^{\mathrm{I}}) = f(B^{\mathrm{R}}, B^{\mathrm{I}})(x_0 + \mathrm{i}y_0 - B)(x_0 + \mathrm{i}y_0 - B)^{-1}.$$

Hence,

$$\|f(B^{\mathrm{R}}, B^{\mathrm{I}})\| \leq \|f(B^{\mathrm{R}}, B^{\mathrm{I}})(x_0 + \mathrm{i}y_0 - B)\| \|(x_0 + \mathrm{i}y_0 - B)^{-1}\| \leq \epsilon \|(x_0 + \mathrm{i}y_0 - B)^{-1}\|.$$

By choosing ϵ small enough we can demand that the right hand side is less than 1. But $f(x_0, y_0) = 1$. This is a contradiction.

Now the $*$-homomrphism (6.22) is simply the reduced $*$-homorphism (6.23), where we identify $\mathrm{sp}(B^{\mathrm{R}}, B^{\mathrm{I}})$ with a subset of $\mathbb{C}$ with the map j. $\square$

6.5 Functional calculus for a family of commuting normal operators

Theorem 6.8 (Fuglede) *Let $A, B \in \mathfrak{A}$ and let B be normal. Then $AB = BA$ implies $AB^* = B^*A$.*

Proof. For $\lambda \in \mathbb{C}$, the operator $U(\lambda) := \mathrm{e}^{\lambda B^* - \overline{\lambda} B} = \mathrm{e}^{-\overline{\lambda} B} \mathrm{e}^{\lambda B^*}$ is unitary. Moreover, $A = \mathrm{e}^{\overline{\lambda} B} A \mathrm{e}^{-\overline{\lambda} B}$. Hence

$$\mathrm{e}^{-\lambda B^*} A \mathrm{e}^{\lambda B^*} = U(-\lambda) A U(\lambda) \tag{6.24}$$

is a uniformly bounded analytic function. Hence is constant. Differentiating it wrt λ we get $[A, B^*] = 0$. $\square$

Suppose that $\{B_i \ : \ i \in I\}$ is a family of commuting normal operators in $B(\mathcal{V})$. Set $B_i^{\mathrm{R}} := \frac{1}{2}(B_i + B_i^*)$ and $B_i^{\mathrm{I}} := \frac{1}{2\mathrm{i}}(B_i - B_i^*)$. Then by the Fuglede theorem, $\{B_i^{\mathrm{R}}, B_i^{\mathrm{I}} \ : \ i \in I\}$ is a family of commuting self-adjoint operators. Thus we have a $*$-homomorphism

$$\prod_{i \in I} \mathrm{sp} B_i^{\mathrm{R}} \times \mathrm{sp} B_i^{\mathrm{I}} \ni G \mapsto G(B_i^{\mathrm{R}}, B_i^{\mathrm{I}} \ : \ i \in I) \in B(\mathcal{V}) \tag{6.25}$$

Recall that the joint spectrum

$$\mathrm{sp}(B_i^{\mathrm{R}}, B_i^{\mathrm{I}} \ : \ i \in I) \tag{6.26}$$

(the spectrum of the $*$-homomorphism (6.25)) is defined as a subset of $\prod_{i \in I} \mathrm{sp}(B_i^{\mathrm{R}}, B_i^{\mathrm{I}})$. By Theorem **??**, we can identify $\mathrm{sp}(B_i^{\mathrm{R}}, B_i^{\mathrm{I}}) \subset \mathbb{R}^2$ with $\mathrm{sp}B_i \subset \mathbb{C}$. Thus $\prod_{i \in I} \mathrm{sp}(B_i^{\mathrm{R}}, B_i^{\mathrm{I}})$ can be identified with $\prod_{i \in I} \mathrm{sp}B_i$. The image of (6.26) under this identification is called the joint spectrum of the family $(B_i \ : \ i \in I)$ and denoted $\mathrm{sp}(B_i \ : \ i \in I)$. Note that this generalizes the definition from the self-adjoint case.

Theorem 6.9 *Let $\{B_i \ : \ i \in I\}$ be a family of commuting normal operators in a $B(\mathcal{V})$. Then*

(1) *$\{z_i \ : \ i \in I\} \in \prod_{i \in I} \mathrm{sp}B_i$ does not belong to $\mathrm{sp}\{B_i \ : \ i \in I\}$ iff there exists a finite subset $\{i_1, \dots, i_n\} \subset I$ and functions $f_{i_j} \in C(\mathrm{sp}B_{i_j})$, with $f_{i_j}(z_{i_j}) \neq 0$, $j = 1, \dots, n$ such that $f_{i_1}(B_{i_1}) \cdots f_{i_n}(B_{i_n}) = 0$.*

(2) *There exists a unique continuous unital $*$-homomorphism*

$$C(\mathrm{sp}(B_i \ : \ i \in I) \ni g \mapsto g(B_i \ : \ i \in I) \in B(\mathcal{V}) \tag{6.27}$$

such that if $\mathrm{id}_j(z_i \ : \ i \in I) = z_j$, then

$$\mathrm{id}_j(B_i \ : \ i \in I) = B_j.$$

(3) *(6.27) is injective and satisfies*

$$\|g(B_i \ : \ i \in I)\| = \|g\|_\infty.$$

Proof. It is obvious that the spectrum of (6.21) is contained in the set described in (1). We need to prove the converse inclusion.

Suppose that $\{z_i \ : \ i \in I\}$ does not belong to the spectrum of (6.21). Then there exists $F \in C(\prod_{i\in I} \mathrm{sp}A_i)$ such that $F(z_i \ : \ i \in I) = 1$ and $F(B_i \ : \ i \in I) = 0$. Changing F into $\min(2F, 1)$ we can guarantee that $F = 1$ on a neighborhood of $\{z_i \ : \ i \in I\}$. This means that there exists a finite subset $\{i_1, \dots, i_n\} \subset I$ and open sets U_{i_j}, $z_{i_j} \in U_{i_j}$, $U_{i_j} \subset \mathrm{sp}B_{i_j}$ such that $F = 1$ on $\prod_{i\in I} W_i$ where $W_i = \mathrm{sp}A_i$ for $i \notin \{i_1, \dots, i_n\}$ and $W_{i_j} = U_{i_j}$, $j = 1, \dots, n$. We can find $g_{i_j} \in C(\mathrm{sp}B_{i_j})$ with $g_{i_j}(z_{i_j}) \neq 0$ and $\{g_{i_j} \neq 0\} \subset U_{i_j}$. Now

$$0 = F(B_i \ : \ i \in I)g_{i_1}(B_{i_1}) \cdots g_{i_n}(B_{i_n}) = g_{i_1}(B_{i_1}) \cdots g_{i_n}(B_{i_n}).$$

This ends the proof of (1).

To see (2) and (3) we reduce the homomrphism (6.25) and use Theorem 6.5 implies (2) and (3). □

6.6 Projector valued (PV) measures

Le $(X, \mathcal{F})$ be a set with a σ-field. Let $\mathcal{V}$ be a Hilbert space. We say that

$$\mathcal{F} \ni A \mapsto P_A \in \mathrm{Proj}(\mathcal{V}) \tag{6.28}$$

is a projector valued measure on $\mathcal{V}$ iff

(1) $P_\emptyset = 0$;

(2) If $A_1, A_2, \ldots \in \mathcal{F}$ are disjoint, and $A = \cup_{i=1}^\infty A_i$, then $P_A = \mathrm{s-}\lim_{n\to\infty} \sum_{j=1}^n P_{A_j}$.

We call P_X the support of the projector valued measure (6.28).

Theorem 6.10 *For any $A, B \in \mathcal{F}$ we have*

$$P_A P_B = P_{A\cap B}.$$

Proof. First consider the case $A \cap B = \emptyset$. By (2)

$$P_{A\cup B} = P_A + P_B.$$

Hence $P_A + P_B$ is a projector. Hence $(P_A + P_B)^2 = P_A + P_B$. This implies

$$P_A P_B + P_B P_A = 0. \tag{6.29}$$

Multiplying from both sides by P_B we get $2P_B P_A P_B = 0$ Multiplying (6.29) from the left by P_B we get $P_B P_A = -P_B P_A P_B$. Thus $P_B P_A = 0$.

Next consider the case $A \subset B$. Then

$$P_B = P_A + P_{B\setminus A}.$$

Using $P_A P_{B\setminus A} = 0$ we see that $P_B P_A = P_A$.

Finally, consider arbitrary A, B. Then

$$P_A P_B = (P_{A\setminus B} + P_{A\cap B})(P_{B\setminus A} + P_{A\cap B}) = P_{A\cap B}.$$

□

Theorem 6.11 *Let $\mathcal{F} \ni A \mapsto P_A$ be a projector valued measure and let $\mathcal{L}^\infty(X)$ denote the space of bounded measurable functions on X. Then there exists a unique contractive $*$-homomorphism*

$$\mathcal{L}^\infty(X) \ni f \mapsto \int f(x)\mathrm{d}P(x) \in B(\mathcal{V})$$

such that $\int 1_A(x)\mathrm{d}P(x) = P_A$, $A \in \mathcal{F}$.

Proof. If f is an elementary function, that is a finite linear combination of characteristic functions of measurable sets

$$f = \sum_{j=1}^{n} \lambda_j 1_{A_j},$$

then clearly

$$\int f(x)\mathrm{d}P(x) = \sum_{j=1}^{n} \lambda_j P_{A_j}.$$

For such functions the multiplicativity of γ is obvious.

Then we use the fact that elementary functions are dense in $\mathcal{L}^\infty(X)$ in the supremum norm. □

Let us give an alternative equivalent definition of the spectral integral, which uses directly concepts from measure theory. For any $w \in \mathcal{V}$

$$\mathcal{F} \ni A \mapsto \mu_w(A) := (w|P_A w)$$

is a finite measure. Likewise, for any $w, v \in \mathcal{V}$,

$$\mathcal{F} \ni A \mapsto \mu_{w,v}(A) := (w|P_A v)$$

is a finite charge.

Theorem 6.12 *For any $f \in \mathcal{L}^\infty(X)$,*

$$\int f(x)\mathrm{d}\mu_w(x) = \left(w| \int f(x)\mathrm{d}P(x)w\right).$$

Here is a version of the Lebesgue dominated convergence theorem for spectral integrals:

Theorem 6.13 *If $f_n \to f$ pointwise, $|f_n| \leq c$, then* $\mathrm{s-}\lim_{n\to\infty} \int f_n(x)\mathrm{d}P(x) = \int f(x)\mathrm{d}P(x)$.

6.7 Projector valued Riesz-Markov theorem

Let X be a compact Hausdorff space, $\mathcal{V}$ a Hilbert space and $\gamma : C(X) \to B(\mathcal{V})$ a unital $*$-homomorphism.

We define the upper projector valued measure associated with γ as follows. For any open $U \subset X$ we define

$$P_U^{\mathrm{up}} := \sup\{\gamma(f) \ : \ 0 \leq f \leq 1_U, \ f \in C(X)\}.$$

For any $A \subset X$ we set

$$P_A^{\mathrm{up}} := \inf\{P_U^{\mathrm{up}} \ : \ U \text{ is open }, \ A \subset U\}.$$

We define the lower projector valued measure associated with γ as follows. For any closed $C \subset X$ we define

$$P_C^{\mathrm{low}} := \inf\{\gamma(f) \ : \ 1_C \leq f, \ f \in C(X)\}.$$

For any $A \subset X$ we set

$$P_A^{\mathrm{low}} := \sup\{P_C^{\mathrm{low}} \ : \ C \text{ is closed }, \ C \subset A\}.$$

We say that $A \subset X$ is γ-measurable if $P_A^{\mathrm{up}} = P_A^{\mathrm{low}}$. The family of γ-measurable sets is denoted $\mathcal{F}_\gamma$. For such sets A we set $P_A = P_A^{\mathrm{up}} = P_A^{\mathrm{low}}$.

Theorem 6.14 (1) P_A^{up} *and* P_A^{low} *are projectors for any* $A \subset X$.

(2) $\mathcal{F}_\gamma$ *is a* σ*-field containing Borel sets.*

(3) $\mathcal{F}_\gamma \ni A \mapsto P_A \in \mathrm{Proj}(\mathcal{V})$ *is a projector valued measure with support 1.*

(4) $C(X) \subset \mathcal{L}^\infty(X)$ *and if* $f \in C(X)$*, then* $\gamma(f) = \int f(x)\mathrm{d}P(x)$.

One can construct the spectral integral directly from γ as follows.

We define the upper integral as follows. If f is a lower semicontinuous function on X, we set

$$\int^{\mathrm{up}} f(x)\mathrm{d}P(x) := \sup\{\gamma(g) \ : \ g \in C(X), \quad g \leq f\}.$$

If f is an arbitrary function, we set

$$\int^{\mathrm{up}} f(x)\mathrm{d}P(x) := \inf\left\{\int^{\mathrm{up}} g(x)\mathrm{d}P(x) \ : \ g \text{ is lower semicontinuous and } f \leq g\right\}.$$

We define the lower integral as follows. If f is a upper semicontinuous function on X, we set

$$\int^{\mathrm{low}} f(x)\mathrm{d}P(x) := \inf\{\gamma(g) \ : \ g \in C(X), \quad f \leq g\}.$$

If f is an arbitrary function, we set

$$\int^{\mathrm{low}} f(x)\mathrm{d}P(x) := \sup\left\{\int^{\mathrm{low}} g(x)\mathrm{d}P(x) \ : \ g \text{ is upper semicontinuous and } g \leq f\right\}.$$

Theorem 6.15 *A function* f *on* X *is* $\mathcal{F}_\gamma$*-measurable iff*

$$\int^{\mathrm{up}} f(x)\mathrm{d}P(x) = \int^{\mathrm{low}} f(x)\mathrm{d}P(x) \tag{6.30}$$

and then (6.30) equals

$$\int f(x)\mathrm{d}P(x).$$

One can also construct the spectral integral using the Riesz-Markov for usual measures. For any $w \in \mathcal{V}$,

$$C(X) \ni f \mapsto (w|\gamma(f)w)$$

is a positive functional on X. By the Riesz-Markov theorem it defines a unique Radon measure on X, which we will call μ_w.

Theorem 6.16 *If* f *is* γ*-measurable, then it is measurable for measure* μ_w *for any* $w \in \mathcal{V}$*, and then,*

$$\left(w\middle| \int f(x)\mathrm{d}P(x) \ w\right) = \int f(x)\mathrm{d}\mu_w(x).$$

6.8 Absolute continuous, singular continuous and point spectrum

Let $(X, \mathcal{F})$ be a set with a σ-field and $\mathcal{F} \ni A \mapsto P(A) \in \mathrm{Proj}(\mathcal{V})$ a projection valued measure. Let $\mathcal{I} \subset \mathcal{F}$ be a σ-ring – an ideal in $\mathcal{F}$. We say that P is $\mathcal{I}$-singular if

$$P(A) = \sup\{P(B) \ : \ B \subset A, \ B \in \mathcal{I}\}, \ A \in \mathcal{F}.$$

We say that P is $\mathcal{I}$-continuous if

$$A \in \mathcal{I} \ \Rightarrow \ P(A) = 0. \tag{6.31}$$

Let

$$P_{\mathcal{I}\mathrm{s}} := \sup\{P(N) \ : \ N \in \mathcal{I}\}, \quad P_{\mathcal{I}\mathrm{c}} := 1 - P_{\mathcal{I}\mathrm{c}}.$$

Then

$$\begin{aligned} \mathcal{F} \ni A \quad &\mapsto \quad P_{\mathcal{I}\mathrm{c}}(A) := P_{\mathcal{I}\mathrm{c}} P(A), \\ \mathcal{F} \ni A \quad &\mapsto \quad P_{\mathcal{I}\mathrm{s}}(A) := P_{\mathcal{I}\mathrm{s}} P(A) \end{aligned}$$

are respectively $\mathcal{I}$-continuous and $\mathcal{I}$-singular PV measures.

In particular, let $\mathcal{I}_0$ be the σ-ring of countable sets. If $\mathcal{I}_0$ is contained in $\mathcal{F}$, then it is clearly an ideal in $\mathcal{F}$. Then instead of saying $\mathcal{I}_0$-continuous, we say simply continuous and instead of $P_{\mathcal{I}_0\mathrm{c}}$ we write P_c. Instead of saying $\mathcal{I}_0$-singular, we say point and instead of $P_{\mathcal{I}_0\mathrm{s}}$ we write P_p.

Theorem 6.17 *Suppose that $\mathcal{V}$ is separable. Let $A \mapsto P(A) \in B(\mathcal{V})$ be a PV measure. Then there exists a countable set $I \subset X$, such that $P_\mathrm{p} = P_I$.*

Assume now that $\mathcal{F}$ is a Borel σ-field on a subset of $\mathbb{R}$ or $\mathbb{C}$. Let $\mathcal{I}_1$ be the σ-ring of sets of the Lebesgue measure zero. Clearly, $\mathcal{I}_1 \cap \mathcal{F}$ is an ideal in $\mathcal{F}$. (If needed, in what follows we replace $\mathcal{I}$ with $\mathcal{I}_1 \cap \mathcal{F}$). Then instead of saying $\mathcal{I}_1$-continuous, we say absolutely continuous and instead of $P_{\mathcal{I}_1\mathrm{c}}$ we write P_ac. Then instead of saying $\mathcal{I}_1$-singular, we say singular and instead of $P_{\mathcal{I}_1\mathrm{s}}$ we write P_s.

Clearly, we have $P_\mathrm{ac} \leq P_\mathrm{c}$. A measure that is both singular wrt $\mathcal{I}_1$ and continuous is called singular continuous and we write $P_\mathrm{sc} := (1 - P_\mathrm{ac}) P_\mathrm{s}$. Thus we have the decompositions

$$1 = P_\mathrm{ac} + P_\mathrm{sc} + P_\mathrm{p}, \qquad P_\mathrm{s} = P_\mathrm{sc} + P_\mathrm{p}, \quad P_\mathrm{c} = P_\mathrm{ac} + P_\mathrm{sc}.$$

Theorem 6.18 *Suppose that $\mathcal{V}$ is separable. Let $A \mapsto P(A) \in B(\mathcal{V})$ be a PV measure. Then there exists a set $N \subset X$ of the Lebesgue measure zero and a countable set $I \subset X$ such that $P_\mathrm{sc} = P_N$ and $P_\mathrm{p} = P_I$.*

If B is a normal operator, then we can define the corresponding PV measure $A \mapsto 1_A(B)$, and correspondingly we obtain the projections $1_\mathrm{ac}(A)$, $1_\mathrm{sc}(A)$, $1_\mathrm{p}(A)$. Note that $1_\mathrm{p}(A)$ is the projection onto the closed span of eigenvectors of A. In this case one often introduces the absolutely continuous and singular continuous spectrum of A:

$$\mathrm{sp}_\mathrm{ac} A := \mathrm{sp}\left(A\Big|_{\mathrm{Ran}\, 1_\mathrm{ac}(A)} \right), \quad \mathrm{sp}_\mathrm{sc} A := \mathrm{sp}\left(A\Big|_{\mathrm{Ran}\, 1_\mathrm{sc}(A)} \right).$$

Note that we defined the point spectrum of A as the set of eigenvalues of A. Therefore,

$$\left(\mathrm{sp}_\mathrm{p} A\right)^\mathrm{cl} = \mathrm{sp}\left(A\Big|_{\mathrm{Ran}\, 1_\mathrm{p}(A)} \right). \tag{6.32}$$

However, some authors prefer to use the right hand side of (6.32) as the definition of of $\mathrm{sp}_\mathrm{p} A$.

6.9 L^2 spaces

Let $(X, \mathcal{F}, \mu)$ be a space with a measure. Let Y be a topological space and $f : X \to Y$ a Borel function. We say that $y \in Y$ belongs to the essential range of f, denoted $y \in \mathrm{essRan}\, f$, iff for any neighborhood U of y we have $\mu(f^{-1}(U)) \neq 0$. Note that if $f : X \to \mathbb{C}$ is Borel, then $\|f\|_\infty = \sup\{|f(x)| \ : \ x \in \mathrm{essRan}\, f\}$.

Let $f \in L^\infty(X)$. Then

$$L^2(X) \ni h \mapsto T_f h := fh \in L^2(X)$$

is a bounded normal operator with $\mathrm{sp}T_f = \mathrm{essRan}\, f$ and $\|T_f\| = \|f\|_\infty$. The operator T_f is self-adjoint iff $\mathrm{essRan}\, f \subset \mathbb{R}$.

Suppose that $(f_i \ : \ i \in I)$ is a family of functions in $L^\infty(X)$. Clearly, the operators T_{f_i} commute with one another.

We can introduce the function

$$X \ni x \mapsto (f_i(x) \ : \ i \in I) \in \prod_{i \in I} \mathbb{C}.$$

We have

$$\mathrm{sp}(T_{f_i} \ : \ i \in I\} = \mathrm{essRan}\,(f_i \ : \ i \in I).$$

6.10 Spectral theorem in terms of L^2 spaces

Theorem 6.19 *Let $\gamma : C(X) \to B(\mathcal{V})$ be a unital $*$-homomorphism. Then there exists a family of Radon measures μ_i, $i \in I$ on X and a unitary operator $U : \underset{i \in I}{\oplus} L^2(X, \mu_i)$ such that $\gamma(f) = U \underset{i \in I}{\oplus} T_f U^*$.*

Proof. Step 1. If $v \in \mathcal{V}$, the cyclic subspace for v is defined as $\mathcal{V}_v := \{Av \ : \ A \in \gamma(C(X))\}^{\mathrm{cl}}$. Note that $\mathcal{V}_v$ is a closed linear subspace invariant wrt $\gamma(C(X))$ and $\mathcal{V}_v^\perp$ is also invariant wrt $\gamma(C(X))$.

We easily see that there exists a family of nonzero vectors $\{v_i \ : \ i \in I\}$ such that $\mathcal{V} = \underset{i \in I}{\oplus} \mathcal{V}_{v_i}$.

Step 2. By the Riesz-Markov Theorem there exists a Radon measure μ_i on X such that $\int f \mathrm{d}\mu_i := (v_i | \gamma(f) v_i)$. The unitary operator U is defined by $Uh := \sum_{i \in I} \gamma(h) v_i$. □

7 Discrete and essential spectrum

7.1 Discrete and essential spectrum

Let $\mathcal{X}$ be a Banach space and $A \in B(\mathcal{X})$. We say that $e \in \mathrm{sp}A$ belongs to the discrete spectrum of A if it is an isolated point of $\mathrm{sp}A$ and $\dim 1_{\{e\}}(A) < \infty$. The discrete spectrum is denoted by $\mathrm{sp}_{\mathrm{d}}(A)$. The essential spectrum is defined as

$$\mathrm{sp}_{\mathrm{ess}} A := \mathrm{sp}A \backslash \mathrm{sp}_{\mathrm{d}} A.$$

Assume now that $\mathcal{H}$ is a Hilbert space and A is an operator on $\mathcal{H}$. Then

Theorem 7.1 *Let A be self-adjoint and $\lambda \in \mathrm{sp}A$. Then*

(1) $\lambda \in \mathrm{sp}_{\mathrm{d}} A$ *iff there exists* $\epsilon > 0$ *such that* $\dim 1_{[\lambda-\epsilon, \lambda+\epsilon]}(A) < \infty$.

(2) $\lambda \in \mathrm{sp}_{\mathrm{ess}}(A)$ *iff for every* $\epsilon > 0$ *we have* $\dim 1_{[\lambda-\epsilon, \lambda+\epsilon]}(A) = \infty$.

Theorem 7.2 *Let A be normal and $\lambda \in \mathrm{sp}A$. Then*

(1) $\lambda \in \mathrm{sp}_{\mathrm{d}} A$ *iff there exists* $\epsilon > 0$ *such that* $\dim 1_{B(\lambda, \epsilon)}(A) < \infty$.

(2) $\lambda \in \mathrm{sp}_{\mathrm{ess}}(A)$ *iff for every* $\epsilon > 0$ *we have* $\dim 1_{B(\lambda, \epsilon)}(A) = \infty$.

Remark 7.3 *If A is a closed operator, then the definitions of discrete and essential spectrum remain unchanged, as well Theorems 7.1 and 7.2.*

7.2 The mini-max and max-min principle

Let A be a bounded self-adjoint operator on a Hilbert space $\mathcal{V}$. We define for $n = 1, 2, \ldots$

$$\mu_n(A) := \inf\{\sup\{(v|Av) \ : \ \|v\| = 1, \ v \in \mathcal{L}\} \ : \ \mathcal{L} \ \text{ is an } n\text{-dim. subspace of } \mathcal{V}\}.$$

$$\Sigma(A) := \inf \mathrm{sp}_{\mathrm{ess}}(A).$$

Theorem 7.4 *Write μ_n, Σ for $\mu_n(A)$, $\Sigma(A)$.*

(1) *(μ_n) is an increasing sequence.*

(2) $\lim\limits_{n\to\infty} \mu_n = \Sigma$

(3) *Let $a \in \mathbb{R}$. Then*

$$\begin{aligned} a \leq \mu_1 \quad &\Leftrightarrow \quad 1_{]-\infty,a[}(A) = 0; \\ \mu_n < a \leq \mu_{n+1} \quad &\Leftrightarrow \quad 1_{]-\infty,a[}(A) = n; \\ \Sigma < a \quad &\Leftrightarrow \quad 1_{]-\infty,a[}(A) = \infty. \end{aligned}$$

Proof. For any $(n+1)$-dimensional space $\mathcal{L}$ there exists an n-dimensional space $\mathcal{L}'$ contained in $\mathcal{L}$. Clearly,

$$\sup\{(v|Av) \ : \ \|v\| = 1, \ v \in \mathcal{L}'\} \leq \sup\{(v|Av) \ : \ \|v\| = 1, \ v \in \mathcal{L}\}.$$

Therefore, $\mu_n \leq \mu_{n+1}$.

Let $\mathcal{L}_a := \mathrm{Ran}\, 1_{]-\infty,a[}(A)$ be n-dimensional.

For some $a_0 < a$, $\mathcal{L}_a = \mathrm{Ran}\, 1_{]-\infty,a_0]}(A)$. Now

$$\sup\{(v|Av) \ : \ \|v\| = 1, \ v \in \mathcal{L}_a\} \leq a_0 < a.$$

Thus $\mu_n < a$.

If $\mathcal{L}$ is $(n+1)$-dimensional, then $\mathcal{L} \cap \mathcal{L}_a^\perp \neq \{0\}$. Thus we can find normalized $v \in \mathcal{L} \cap \mathcal{L}_a^\perp$. Now $v \in \mathrm{Ran}\, 1_{[a,\infty[}(A)$, hence $(v|Av) \geq a$. Thus

$$\sup\{(v|Av) \ : \ \|v\| = 1, \ v \in \mathcal{L}\} \geq a.$$

Hence, $a \leq \mu_{n+1}$. □

Theorem 7.5 (The Reyleigh-Ritz method) *Let $\mathcal{W}$ be a linear subspace. Set $A_{\mathcal{K}} := P_{\mathcal{W}} A P_{\mathcal{W}}\big|_{\mathcal{W}}$, where $P_{\mathcal{W}}$ denotes the projector onto $\mathcal{W}$. Then $A_{\mathcal{W}}$ is a bounded self-adjoint operator and*

$$\mu_n(A) \leq \mu_n(A_{\mathcal{W}}).$$

Theorem 7.6 (1) *Let $A \leq B$. Then $\mu_n(A) \leq \mu_n(B)$.*

(2) $|\mu_n(A) - \mu_n(B)| \leq \|A - B\|$.

Remark 7.7 *The theorems of this subsection remain true if the operators are only bounded from below (but not necessarily bounded). In this case, if v does not belong to the form domain of A, then we set $(v|Av) = \infty$.*

Notice also that if $\mathcal{D}$ is an essential domain for the quadratic form generated by A, then

$$\mu_n(A) := \inf\{\sup\{(v|Av) \ : \ \|v\| = 1, \ v \in \mathcal{L} \cap \mathcal{D}\} \ : \ \mathcal{L} \text{ is an } n\text{-dim. subspace of } \mathcal{V}\}.$$

7.3 Singular values of an operator

Let A be a bounded operator on a Hilbert space $\mathcal{V}$. We define for $n = 1, 2, \dots$

$$\mathrm{s}_n(A) := \sup\{\inf\{(\|Av\| \ : \ \|v\| = 1, \ v \in \mathcal{L}\} \ : \ \mathcal{L} \ n\text{-dim. subspace of } \mathcal{V}\}.$$

Clearly, for $|A| := (A^*A)^{1/2}$,

$$\mathrm{s}_n(A) = \mathrm{s}_n(|A|) = -\mu_n(-|A|),$$

and $\mathrm{s}_1(A) = \|A\|$.

8 Compact operators

8.1 Finite rank operators

This subsection can be viewed as an elementary introduction to compact operators.

Definition 8.1 *An operator $K \in B(\mathcal{X}, \mathcal{Y})$ is called a finite rank operator iff* $\dim \mathrm{Ran}\, K < \infty$.

Theorem 8.2 *Let $K \in B(\mathcal{X}, \mathcal{Y})$ be a finite rank operator. Then*

$$\dim \mathrm{Ran}\, K = \dim X/\mathrm{Ker}K.$$

Proof. Let $y_1, \dots, y_n$ be a basis in $\mathrm{Ran}\, K$. We can find $x_1, \dots, x_n \in X$ such that $Kx_i = y_i$. Then $\mathrm{Span}\{x_1, \dots, x_n\} \cap \mathrm{Ker}K = \{0\}$. Assume that $z \in X$. Then $Kz = \sum c_i y_i$. Thus $z - \sum c_i x_i \in \mathrm{Ker}K$. Hence $z \in \mathrm{Span}\{x_1, \dots, x_n\} + \mathrm{Ker}K$. □

Theorem 8.3 *Let $K \in B(\mathcal{X})$ be a finite rank operator. Then $\mathrm{sp}K = \mathrm{sp_p}K$. Moreover, $\mathrm{sp_{ess}}K = \emptyset$ if $\dim \mathcal{X} < \infty$, otherwise $\mathrm{sp_{ess}}K = \{0\}$.*

Proof. Using the fact that $\dim \mathcal{X}/\mathrm{Ker}K$ is finite, we can find a finite dimensional subspace $\mathcal{Z}$ such that $\mathcal{X} = \mathrm{Ker}K \oplus \mathcal{Z}$. $\mathcal{Z}_1 := \mathcal{Z} + \mathrm{Ran}\, K$ is also finite dimensional. We have $K\mathcal{Z}_1 \subset \mathcal{Z}_1$. We can find a subspace $\mathcal{Z}_2$ such that $\mathcal{Z}_1 \oplus \mathcal{Z}_2 = \mathcal{X}$. Obviously, $\mathcal{Z}_2 \subset \mathrm{Ker}K$. □

8.2 Compact operators on Banach spaces I

Let $\mathcal{X}$, $\mathcal{Y}$ be Banach spaces.

Definition 8.4 *$K \in B(\mathcal{X}, \mathcal{Y})$ is called a compact operator iff for any bounded sequence $x_1, x_2, \dots \in \mathcal{X}$ we can find a convergent subsequence from the sequence $Kx_1, Kx_2, \dots \in \mathcal{Y}$.*

Equivalent definition: if $(\mathcal{X})_1$ denotes the unit ball in $\mathcal{X}$, then $(K(\mathcal{X})_1)^{\mathrm{cl}}$ is a compact set. The set of compact operators from $\mathcal{X}$ to $\mathcal{Y}$ will be denoted $B_\infty(\mathcal{X}, \mathcal{Y})$.

Theorem 8.5 (1) *Let K be a compact operator. Let $(x_i)_{i \in I}$ be a bounded net weakly convergent to x. Then $\lim\limits_{i \in I} Kx_i = Kx$. ($K$ is weak-norm continuous on the unit ball).*

(2) *Let K be a compact operator. Let (x_n) be a sequence weakly convergent to x. Then $\lim\limits_{n \to \infty} Kx_n = Kx$.*

(3) *If A is bounded, K is compact, then AK and KA are compact.*

(4) *If K_n are compact and $\lim_{n \to \infty} K_n = K$, then K is compact.*

(5) *If K is finite rank, then K is compact.*

Proof. (1) Let $(x_i)_{i\in I}$ be a bounded net weakly convergent to x. Then $\mathrm{w}-\lim_{i\in I} Kx_i = Kx$ (because K is bounded). Hence, if Kx_i is convergent in norm, its only limit can be Kx.

Suppose that Kx_i is not convergent. Then there exists a subnet x_{i_j} such that $\|Kx_{i_j} - Kx\| > \epsilon > 0$. By compactness, we can choose a subsubnet $x_{i_{j_m}}$ such that $Kx_{i_{j_m}}$ is convergent. But it can be convergent only to Kx, which is impossible.

(3) is obvious, if we note that A maps a ball into a ball and a convergent sequence onto a convergent sequence.

(4) Let $x_1, x_2, \dots$ be a bounded sequence so that $\|x_n\| \leq C$. Below we will construct a double sequence $x_{n,k}$ such that, for any n, $x_{n+1,1}, x_{n+1,2}, \dots$ is a subsequence of $x_{n,1}, x_{n,2}, \dots$ and

$$\|Kx_{n,m} - Kx_{n,k}\| < (\min(m,k,n))^{-1}.$$

Eventually, the sequence $x_{n,n}$ is a subsequence of x_n such that $Kx_{n,n}$ satisfies the Cauchy condition.

Suppose that we have constructed $x_{n,m}$ up to the index n. We can find N such that $\|K - K_N\| < \frac{1}{3C(n+1)}$. We put $x_{n+1,m} = x_{n,m}$ for $m = 1, \dots n$. For $m > n$, we choose $x_{n+1,m}$ as the subsequence of $x_{n,m}$ such that $\|K_N x_{n+1,m} - K_N x_{n+1,k}\| < \frac{1}{3(n+1)}$ for $k, m > n$. Then for $m > n$

$$\begin{aligned}\|Kx_{n+1,m} - Kx_{n+1,k}\| &\leq \|Kx_{n+1,m} - K_N x_{n+1,m}\| + \|K_N x_{n+1,m} - K_N x_{n+1,k}\| \\ &+\|K_N x_{n+1,k} - Kx_{n+1,k}\| \leq \tfrac{2C}{C3(n+1)} + \tfrac{1}{3(n+1)} = (n+1)^{-1}.\end{aligned}$$

(5) follows by the compactness of the ball in a finite dimensional space $\mathrm{Ran}\, K$. □

Note that $B_\infty(X)$ is a closed ideal of $B(\mathcal{X})$.

8.3 Compact operators on Banach spaces II

In this subsection we prove some properties of compact operators on Banach spaces. They will be proved again in the context of Hilbert spaces, so the reader interested only in Hilbert spaces can omit this subsection.

Theorem 8.6 (Schauder) *If K is compact, then $K^\#$ is compact.*

Let $K \in B(\mathcal{X}, \mathcal{Y})$ be compact. We treat $(\mathcal{Y}^\#)_1$ (the unit ball in $\mathcal{Y}^\#$) as a family of continuous functions on $\mathcal{X}$. It is equicontinuous on $\mathcal{Y}$ and bounded on $(\mathcal{Y})_1$. Therefore, $(\mathcal{Y}^\#)_1$ is equicontinuous and bounded on the compact metric space $(K(\mathcal{X})_1)^{\mathrm{cl}}$ (where $(\mathcal{X})_1$ is the unit ball in $\mathcal{X}$). Hence, by the Ascoli Theorem, from every sequence $v_n \in (\mathcal{Y}^\#)_1$ we can choose a subsequence v_{n_k} uniformly convergent on $(K(\mathcal{X})_1)^{\mathrm{cl}}$. Hence $\langle K^\# v_{n_k}, x\rangle$ converges uniformly for $x \in (\mathcal{X})_1$. Therefore, $K^\# v_{n_k}$ is norm convergent in $X^\#$. □

Theorem 8.7 (Riesz-Schauder) *Let K be a compact operator. Then $\mathrm{sp}_{\mathrm{ess}} K = \{0\}$ if the space is infinite dimensional and $\mathrm{sp}_{\mathrm{ess}} K = \emptyset$ otherwise.*

Lemma 8.8 $\mathrm{sp}_{\mathrm{p}}(K)\backslash\{0\}$ *is a discrete set.*

Proof. Let $\lambda_n \in \mathrm{sp}_{\mathrm{p}}(K)$, $n = 1, 2, \dots$, $\lambda_n \to \lambda \neq 0$, $x_n \in \mathcal{X}$, $Kx_n = \lambda_n x_n$. Let $M_n = \mathrm{Span}\{x_1, \dots, x_n\}$. Notice that M_n is a strictly increasing sequence of subspaces invariant for K. We can find a sequence of vectors $v_1, v_2, \dots$ such that $v_n \in M_n$, $\mathrm{dist}(v_n, M_{n-1}) \geq \frac{1}{2}$ and $\|v_n\| = 1$. For $m < n$ write

$$\lambda_n^{-1} K v_n - \lambda_m^{-1} K v_m = v_n - \left(\lambda_m^{-1} K v_m - \lambda_n^{-1}(K - \lambda_n) v_n\right). \tag{8.33}$$

We have $Kv_n \in M_m \subset M_{n-1}$ and $(K - \lambda_n)v_n \in M_{n-1}$. Hence the second term on the right of (8.33) has the norm $\geq \frac{1}{2}$. Thus $\lambda_n^{-1} K v_n$ does not contain a Cauchy subsequence. But it is a bounded sequence. Hence K is not compact. □

Lemma 8.9 *If $z \notin \mathrm{sp_p} K\backslash\{0\}$, then there exists $c > 0$ such that*

$$\|(z-K)x\| \geq c\|x\|.$$

Proof. Let the sequence x_n satisfy $\|x_n\| = 1$, $(z-K)x_n \to 0$. Passing to a subsequence, we can suppose that $Kx_n \to y$. Then $zx_n \to y$. We have 2 possibilities:
1) $y \neq 0$. Then $Ky = zy$. But this contradicts the assumption $z \notin \mathrm{sp_p} K$.
2) $y = 0$. Then $x_n \to 0$, which contradicts $\|x_n\| = 1$. □

Lemma 8.10 $\mathrm{sp} K = \mathrm{sp_p} K \cup \mathrm{sp_p} K^{\#} \cup \{0\}$.

Proof. Let $z \notin \mathrm{sp_p} K \cup \mathrm{sp_p} K^{\#} \cup \{0\}$. Using Lemma 8.9 and the compactness of $K^{\#}$, we obtain

$$\mathrm{Ker}(z-K) = \mathrm{Ker}(z-K^{\#}) = \{0\}.$$

Hence $(\mathrm{Ran}\,(z-K))^{\mathrm{cl}} = (\mathrm{Ker}(z-K^{\#}))^{\mathrm{an}} = \mathcal{X}$. Lemma 8.9 implies also that $\mathrm{Ran}\,(z-K)$ is closed, hence $\mathrm{Ran}\,(z-K) = \mathcal{X}$. By Lemma 8.9, $(z-K)$ has a bounded inverse. □

Proof of Theorem 8.7 Let $\lambda \in \mathrm{sp}K\backslash\{0\}$. Then λ is an isolated point of $\mathrm{sp}K$. Let γ be a closed curve around λ that does not encircle 0. Then

$$\begin{aligned} 1_{\{\lambda\}}(K) &= (2\pi\mathrm{i})^{-1} \int_\gamma (z-K)^{-1}\mathrm{d}z = (2\pi\mathrm{i})^{-1} \int_\gamma \left((z-K)^{-1} - z^{-1}\right)\mathrm{d}z \\ &= (2\pi\mathrm{i})^{-1} \int_\gamma (z-K)^{-1}Kz^{-1}\mathrm{d}z \end{aligned}$$

is compact. But a projection is compact iff it is finite dimensional. □

8.4 Compact operators in a Hilbert space

Theorem 8.11 *Let $\mathcal{X}$, $\mathcal{Y}$ be Hilbert spaces and $K \in B(\mathcal{X}, \mathcal{Y})$. TFAE:*

(1) *K is compact (i.e. $(K(\mathcal{X})_1)^{\mathrm{cl}}$ is compact).*

(2) *K maps bounded weakly convergent nets onto norm convergent sequences (K is weak-norm continuous on the unit ball).*

(3) *$K(\mathcal{X})_1$ is compact.*

(4) *Let (x_n) be a sequence weakly convergent to x. Then $\lim_{n\to\infty} Kx_n = Kx$.*

(5) *If $|K| := (K^*K)^{1/2}$, then $\mathrm{sp_{ess}}|K| \subset \{0\}$.*

(6) *There exist orthonormal systems $x_1, x_2, \ldots \in \mathcal{X}$ and $y_1, y_2, \ldots \in \mathcal{Y}$ and a sequence of positive numbers $k_1, k_2, \ldots$ convergent to zero such that*

$$K = \sum_{n=1}^{\infty} k_n |y_n)(x_n|.$$

(7) *There exists a sequence of finite rank operators K_n such that $K_n \to K$.*

Proof. (1)⇒(2), by Theorem 8.5, is true even in Banach spaces.

(2)⇒(3). In a Hilbert space $(\mathcal{X})_1$ is weakly compact. The image of a compact set under a continuous map is compact.

(3)⇒(1) is obvious.

(2)⇒(4) is obvious.

(4)⇒(5). Suppose (5) is not true. This means that for some $\epsilon > 0$, $\operatorname{Ran} 1_{[\epsilon,\infty[}(|K|)$ is infinite dimensional. Let $x_1, x_2, \dots$ be an infinite orthonormal system in $\operatorname{Ran} 1_{[\epsilon,\infty[}(A)$. Then x_n goes weakly to zero, but $\|Kx_n\| \geq \epsilon$.

(5)⇒(6). Let $x_1, x_2, \dots$ be an orthonormal system of eigenvectors of $|K|$ with eigenvalues k_n. Then set $y_n := k_n^{-1} K x_n$.

(6)⇒(7). It suffices to set $K_\epsilon := K 1_{[\epsilon,\infty[}(|K|)$. Then

$$\|K - K_\epsilon\| = \||K| 1_{[0,\epsilon]}(|K|)\| \leq \epsilon.$$

(7)⇒(1), by Theorem 8.5, is true for Banach spaces. □

(1)⇒(6) is sometimes called the *Hilbert-Schmidt Theorem.*

Corollary 8.12 (Schauder) *Let $\mathcal{X}$, $\mathcal{Y}$ be Hilbert spaces and $K \in B_\infty(\mathcal{X}, \mathcal{Y})$. Then $K^* \in B_\infty(\mathcal{Y}, \mathcal{X})$.*

Proof. It follows immediately from Theorem 8.11 (7).

8.5 The Fredholm alternative

Theorem 8.13 (Analytic Fredholm Theorem) *Let $\mathcal{V}$ be a Hilbert space, $\Omega \subset \mathbb{C}$ is open and connected. Let $\Omega \ni z \mapsto A(z) \in B_\infty(\mathcal{V})$ be an analytic function. Let $S := \{z \in \Omega \ : 1 - A(z) \ \text{ is not invertible } \}$ Then either*

(1) *$S = \Omega$, or*

(2) *S is discrete in Ω. Moreover, for $z \in S$, $\operatorname{Ker}(1 - A(z)) \neq \{0\}$ and the coefficients at the negative powers of the Laurent expansion of $(1 - A(z))^{-1}$ are of finite rank. In particular, the residuum is of finite rank.*

Proof. Let $z_0 \in \Omega$. We can find a finite rank operator F with $\|A(z_0) - F\| < 1/2$. Let $\epsilon > 0$ with $\|A(z) - A(z_0)\| < 1/2$ for $|z - z_0| < \epsilon$. Thus $\|A(z_0) - F\| < 1$ for $|z - z_0| < 1$.

Set $G(z) := F(1 + F - A(z))^{-1}$. We have

$$(1 - G(z))(1 + F - A(z)) = 1 - A(z).$$

Thus $1 - A(z)$ is invertible iff $1 - G(z)$ is invertible and $\operatorname{Ker}(1 - A(z)) = \{0\}$ iff $\operatorname{Ker}(1 - G(z)) = \{0\}$.

Let P be the projector onto $\operatorname{Ran} F$. Set

$$\begin{aligned} G_0(z) &:= G(z)P &&= PG(z)P, \\ G_1(z) &:= G(z)(1 - P) &&= PG(z)(1 - P). \end{aligned}$$

Then

$$1 - G(z) = 1 - G_0(z) - G_1(z) = (1 - G_1(z))(1 - G_0(z)),$$

and $(1 - G_1(z))^{-1} = 1 + G_1(z)$. Hence, $1 - G(z)$ is invertible iff $1 - G_0(z)$ is and $\operatorname{Ker}(1 - G(z)) = \{0\}$ iff $\operatorname{Ker}(1 - G_0(z)) = \{0\}$. Since $G_0(z)$ is an analytic function in a fixed finite dimensional space, $1 - G_0(z)$ is invertible iff $\det(1 - G_0(z)) \neq 0$ iff $\operatorname{Ker}(1 - G_0(z)) = \{0\}$. Thus $S = \{z \in \Omega \ : \ \det(1 - G_0(z)) \neq 0\}$.

Now we have

$$(1 - A(z))^{-1} = (1 + F - A(z))^{-1}(1 - G_0(z))^{-1}(1 + G_0(z)).$$

The first and third factor on the rhs are analytic in the neighborhood of z_0. Suppose that the middle term has a singularity at z_0. Then it is a pole of the order at most $\dim \mathrm{Ran}\, F$ and all the coefficients at the negative powers of its Laurent expansion are finite rank. □

Corollary 8.14 (Riesz-Schauder) *Let K be a compact operator on a Hilbert space. Then $\mathrm{sp_{ess}}K = \{0\}$ if the space is infinite dimensional and $\mathrm{sp_{ess}}K = \emptyset$ otherwise.*

Proof. We apply the Analytic Fredholm Theorem to $1 - z^{-1}K$. □

8.6 Positive trace class operators

Let $\{v_i\}_{i\in I}$ be an orthonormal basis of a Hilbert space $\mathcal{V}$. Let $A \in B(\mathcal{V})$ and $A \geq 0$. Define

$$\mathrm{Tr}A := \sum_{i\in I} (v_i|Av_i). \tag{8.34}$$

Theorem 8.15 *(8.34) does not depend on the basis.*

Proof. First note that if $A_\alpha \in B(\mathcal{V})$ is an increasing net, then

$$\sum_{i\in I} (v_i|Av_i) = \sup_\alpha \sum_{i\in I} (v_i|A_\alpha v_i). \tag{8.35}$$

Let $\{v_i \ : \ i \in I\}$ and $\{w_j \ : \ j \in J\}$ are orthonormal bases. Assume that $c < \sum\limits_{i\in I} (v_i|Av_i)$. By (8.35), we can find a finite subset $J_0 \subset J$ such that if P_0 is the projection onto $\mathrm{Span}\{w_j \ : \ j \in J_0\}$, then

$$c \leq \sum_{i\in I} (v_i|P_0AP_0v_i).$$

Now

$$\begin{aligned} \sum_{i\in I} (v_i|P_0AP_0v_i) &= \sum_{i\in I} \sum_{j,k\in J_0} (v_i|w_j)(w_j|Aw_k)(w_k|v_i) \\ &= \sum_{j\in J_0} (w_j|Aw_j) \leq \sum_{j\in J} (w_j|Aw_j). \end{aligned} \tag{8.36}$$

Above we used the fact that for any j, k

$$\sum_{i\in I} |(v_i|w_j)(w_j|Aw_k)(w_k|v_i)| \leq \|A\|,$$

which together with the finiteness of J_0 imples that the second sum in (8.36) is absolutely convergent, and also

$$\sum_{i\in I} (v_i|w_j)(w_k|v_i) = \delta_{j,k}.$$

This shows

$$\sum_{i\in I} (v_i|Av_i) \leq \sum_{j\in J} (w_j|Aw_j).$$

Of course, we can reverse the argument. □

We will write $B_+^1(\mathcal{V})$ for the set of $A \in B_+(\mathcal{V})$ such that $\mathrm{Tr}A < \infty$.

Theorem 8.16 (1) *If $A, B \in B_+(\mathcal{V})$, then $\mathrm{Tr}(A+B) = \mathrm{Tr}A + \mathrm{Tr}B$. If $\lambda \in [0, \infty[$, then $\mathrm{Tr}\lambda A = \lambda \mathrm{Tr}A$, where $0\infty = 0$.*

(2) *Let* $B \in B(\mathcal{V},\mathcal{W})$. *Then* $\mathrm{Tr}B^*B = \mathrm{Tr}BB^*$.

(3) *If* $A \in B^1_+(\mathcal{V})$, *and* $B \in B(\mathcal{W},\mathcal{V})$. *Then* $B^*AB \in B^1_+(\mathcal{W})$ *and* $\mathrm{Tr}B^*AB \leq \|B\|^2\mathrm{Tr}A$.

(4) *If* $A \in B^1_+(\mathcal{V})$, *then* A *is compact.*

(5) *Let* $(A_i\ i \in I)$ *be an increasing net in* $B_+(\mathcal{V})$ *and* $A = \mathrm{lub}A_i$. *Then*

$$\mathrm{Tr}A = \sup\{\mathrm{Tr}A_i \ : \ i \in I\}.$$

(6) $\mathrm{Tr}A = \sum_{n=1}^{\infty} s_n(A)$.

Proof. (2) Let (v_i) and (w_j) be bases of $\mathcal{V}$ and $\mathcal{W}$. Then

$$\begin{aligned}\mathrm{Tr}B^*B &= \sum_i\sum_j (v|B^*w_j)(w_j|Bv_i) \\ &= \sum_j\sum_i (w_j|Bv_i)(v_i|B^*w_j) \quad = \mathrm{Tr}BB^*,\end{aligned}$$

where all the terms in the sum are positive, which justifies the exchange of the order of summation.

(3) By (2), we have $\mathrm{Tr}B^*AB = \mathrm{Tr}A^{1/2}BB^*A^{1/2}$. Besides $A^{1/2}BB^*A^{1/2} \leq \|B\|^2A$.

(4) If A has continuous spectrum, then there exists an infinite dimensional projector P and $\epsilon > 0$ such that $A \geq \epsilon P$. Then $\mathrm{Tr}A \geq \epsilon\mathrm{Tr}P = \infty$.

Hence A has just point spectrum. We have $\mathrm{Tr}A = \sum\limits_{i\in I} a_i$, where a_i are eigenvalues of A (counting their multiplicities). □

8.7 Hilbert-Schmidt operators

For $A \in B(\mathcal{V},\mathcal{W})$ set

$$\|A\|_2 := (\mathrm{Tr}A^*A)^{\frac{1}{2}} = (\mathrm{Tr}AA^*)^{\frac{1}{2}}.$$

$B^2(\mathcal{V},\mathcal{W})$ denotes the set of operators with a finite norm $\|A\|_2$. Clearly,

$$\|A\|_2 = \left(\sum_{n=1}^{\infty} s_n(A)^2\right)^{1/2}.$$

If $(v_i)_{i\in I}$ and $(w_j)_{j\in J}$ are bases in $\mathcal{V}$ and $\mathcal{W}$, then

$$\|A\|_2 = \sum_{i\in I}\sum_{j\in J} |(w_j|Av_i)|^2. \tag{8.37}$$

$B^2(\mathcal{V},\mathcal{W})$ is equipped with the scalar product

$$(A|B)_2 = \sum_{i\in I}\sum_{j\in J} \overline{(w_j|Av_i)}(w_j|Bv_i), \tag{8.38}$$

where we used $(v_i)_{i\in I}$ and $(w_j)_{j\in J}$ orthonormal bases in $\mathcal{V}$ and $\mathcal{W}$.

Proposition 8.17 *(8.38) is finite and does not depend on a choice of bases.*

Proof. Clearly, (8.37) is the norm for (8.38). Hence the finiteness of(8.38) follows by the Schwarz inequality. $|(A|B)_2| \leq \|A\|_2\|B\|_2$.

Next note that

$$\|(A + \mathrm{i}^kB)v\|^2 \leq 2\|Av\|^2 + 2\|Bv\|^2.$$

Therefore,

$$\|(A+\mathrm{i}^kB)\|_2^2 \leq \|A\|_2^2+\|B\|_2^2+.$$

Hence if A, B are Hilbert-Schmidt, then so are $A+\mathrm{i}^kB$. Then we note that (8.38) equals

$$(A|B)_2 := \sum_{k=0}^{3}\frac{\mathrm{i}^k}{4}\mathrm{Tr}(A+\mathrm{i}^kB)^*(A+\mathrm{i}^kB), \tag{8.39}$$

which is basis independent.

Remark 8.18 *In the next subsection we extend the notion of trace and (8.39) will be written simply as* $\mathrm{Tr}A^*B$.

Theorem 8.19 (1) *If $A\in B^2(\mathcal{V},\mathcal{W})$, then A is compact.*

(2) *$B^2(\mathcal{V},\mathcal{W})$ is a Hilbert space.*

(3) *If $\{v_i\}_{i\in I}$ is a basis in $\mathcal{V}$ and $\{w_j\}_{j\in J}$ is a basis in $\mathcal{W}$, then $|w_j)(v_i|$ is a basis in $B^2(\mathcal{V},\mathcal{W})$.*

(4) *$B^2(\mathcal{V},\mathcal{W})\ni A\mapsto A^*\in B^2(\mathcal{W},\mathcal{V})$ is a unitary map.*

(5) *If $A\in B^2(\mathcal{V},\mathcal{W})$ and $B\in B(\mathcal{W},\mathcal{X})$, then $BA\in B^2(\mathcal{V},\mathcal{X})$.*

(6) *If (X,μ) and (Y,ν) are spaces with measurs and $\mathcal{V}=B^2(X,\mu)$, $\mathcal{W}=B^2(Y,\nu)$, then every operator $A\in B^2(\mathcal{H},\mathcal{K})$ has the integral kernel $A(\cdot,\cdot)\in B^2(Y\times X,\nu\otimes\mu)$, ie.*

$$(w|Av)=\int\int\overline{w(y)}A(y,x)v(x)\mathrm{d}\mu(y)\mathrm{d}\mu(x)$$

The transformation $B^2(\mathcal{V},\mathcal{W})\ni A\mapsto A(\cdot,\cdot)\in B^2(Y\times X,\nu\otimes\mu)$ that to an operator associates its integral kernel is unitary.

Proof. (1) The operator A^*A is trace class, hence is compact. We can represent A^*A as

$$A^*A=\sum_{j=1}^{\infty}b_j|v_j)(v_j|,$$

with $b_j\to 0$.

If we set $w_j := Av_j$, then

$$A=\sum_{j=1}^{\infty}a_j|w_j)(v_j|,$$

with $|a_j|^2=b_j$. Hence, $a_j\to 0$.

Let us show (2) and (3). Set $E_{ji} := |w_j)(v_i|$. We first check that it is an orthonormal system. If $A\in B^2(\mathcal{V},\mathcal{W})$ is orthogonal to all E_{ji}, then all its matrix elements are zero. Hence $A=0$.

Then we check that if a_{ji} belongs to $L^2(J\times I)$, then $\sum_{j\in J,\ i\in I}a_{ji}E_{ji}$ is the integral kernel of an operator in $B^2(\mathcal{V},\mathcal{W})$. Hence, $B^2(\mathcal{V},\mathcal{W})$ is isomorphic to $L^2(J\times I)$. Hence it is a Hilbert space and $\{E_{ij}\ :\ i\in I, j\in J\}$ is its orthonormal basis. This proves (2) and (3), □

Theorem 8.20 *Suppose that $f,g\in L^\infty(\mathbb{R}^d)$ converge to zero at infinity. Then the operator $g(D)f(x)$ is compact.*

Proof. Let

$$f_n(x) := \left\{ \begin{array}{ll} f(x), & |x| < n \\ 0 & |x| \geq n, \end{array} \right.$$

$$g_n(\xi) := \left\{ \begin{array}{ll} g(\xi), & |\xi| < n \\ 0 & |\xi| \geq n, \end{array} \right.$$

$$g(D)f(x) = \mathcal{F}^* g(x) \mathcal{F} f(x).$$

$$\begin{aligned} \|g(x)\mathcal{F}f(x) - g_n(x)\mathcal{F}f_n(x)\| &\leq \|(g(x) - g_n(x))\mathcal{F}f(x)\| \\ &+\|g_n(x)\mathcal{F}(f(x) - f_n(x)\| \to 0. \end{aligned}$$

It suffices to show the compactness of $g_n(x)\mathcal{F}f_n(x)$. But its integral kernel equals

$$(2\pi)^{-\frac{1}{2}d} g_n(x) \mathrm{e}^{-\mathrm{i}xy} f_n(y),$$

which is square integrable . □

8.8 Trace class operators

Lemma 8.21 *Let $A_+, A'_+ \in B^1_+(\mathcal{V})$, $A_-, A'_- \in B_+(\mathcal{V})$ satisfy $A_+ - A_- = A'_+ - A'_-$. Then*

$$\mathrm{Tr}A_+ - \mathrm{Tr}A_- = \mathrm{Tr}A'_+ - \mathrm{Tr}A'_-.$$

Proof. Clearly, $A_+ + A'_- = A_- + A'_+ \in B_+(\mathcal{V})$. Thus

$$\mathrm{Tr}A_+ + \mathrm{Tr}A'_- = \mathrm{Tr}(A_+ + A'_-) = \mathrm{Tr}(A_- + A'_+) = \mathrm{Tr}A_- + \mathrm{Tr}A'_+.$$

□

By Lemma 8.21, we can uniquely extend the definition of trace as a function with values in $[-\infty, \infty]$ to operators in $B_{\mathrm{sa}}(\mathcal{V})$ that admit a decomposition $A = A_+ - A_-$, where $A_+, A_- \in B_+(\mathcal{V})$ and either B_+ or B_- belongs to $B^1_+(\mathcal{V})$, by setting

$$\mathrm{Tr}(A_+ - A_-) := \mathrm{Tr}A_+ - \mathrm{Tr}A_-.$$

We define $B^1(\mathcal{V}) := \mathrm{Span}B^1_+(\mathcal{V})$. Clearly, $B_+(\mathcal{V}) \cap B^1(\mathcal{V}) = B^1_+(\mathcal{V})$.
Obviously, Tr is well defined and finite on $B^1(\mathcal{V})$.

Theorem 8.22 *Let $A \in B^1(\mathcal{V})$. Then for any orthonormal basis (v_i) in $\mathcal{V}$,*

$$\mathrm{Tr}A = \sum_{i \in I} (v_i | A v_i), \tag{8.40}$$

where the above series is absolutely convergent.

Proof. Let $A = A_+ - A_-$, where $A_+, A_- \in B^1_+(\mathcal{V})$. Then for any orthonormal basis $\sum_{i\in I}(v_i|A_\pm v_i)$ is finite, hence absolutely convergent. Thus (8.40) is the sum of two absolutely convergent series, and hence, absolutely convergent. □

Theorem 8.23 *$B, C \in B^2(\mathcal{V}, \mathcal{W})$ implies $B^*C \in B^1(\mathcal{V})$ and $(B|C)_2 = \mathrm{Tr}B^*C = \mathrm{Tr}CB^*$.*

Proof. We know that $B + \mathrm{i}^{-k}C \in B^2(\mathcal{V},\mathcal{W})$. Hence $B^*C \in B^1(\mathcal{V})$ follows immediately from (8.39). $(B|C)_2 = \mathrm{Tr}B^*C = \mathrm{Tr}CB^*$ also follows from (8.39).

Theorem 8.24 *If $A \in B^1(\mathcal{V})$ and $B \in B(\mathcal{V})$, then $AB, BA \in B^1(\mathcal{V})$ and*

$$\mathrm{Tr}AB = \mathrm{Tr}BA.$$

Proof. It suffices to assume that $A \in B^1_+(\mathcal{V})$. $A^{1/2}$ and $BA^{1/2}$ belong to $B^2(\mathcal{V})$. Hence, using Theorem 8.23, we obtain

$$\begin{aligned}\mathrm{Tr}BA = \mathrm{Tr}(BA^{1/2})A^{1/2} &= \mathrm{Tr}A^{1/2}(BA^{1/2}) \\ &= \mathrm{Tr}(A^{1/2}B)A^{1/2} \quad = \mathrm{Tr}A^{1/2}(A^{1/2}B) = \mathrm{Tr}AB.\end{aligned}$$

□

Theorem 8.25 *TFAE*

(1) *$A \in B^1(\mathcal{V})$.*

(2) *$|A| \in B^1_+(\mathcal{V})$.*

(3) *There exist $B, C \in B^2(\mathcal{V},\mathcal{W})$ such that $A = B^*C$.*

(4) *$\sum\limits_{n=1}^{\infty} s_n(A) < \infty$.*

(5) *For any orthonormal basis (v_i) in $\mathcal{V}$,*

$$\sum_{i\in I} |(v_i|Av_i)| < \infty.$$

Proof. Let $A = U|A|$ be the polar decomposition of A.

(1)⇒(2). Let $A \in B^1(\mathcal{V})$. Then $U^*A = |A| \in B^1(\mathcal{V})$. Since $|A| \in B_+(\mathcal{V})$, this also means that $A \in B^1_+(\mathcal{V})$.

(1)⇐(2). Let $A \in B(\mathcal{V})$ with $A \in B^1(\mathcal{V})$. Then $A = U|A|$ shows that $A \in B^1(\mathcal{V})$.

(2)⇒(3). $A = U|A|^{1/2}|A|^{1/2}$ with $U|A|^{1/2}, |A|^{1/2} \in B^2(\mathcal{V})$.

(2)⇐(3) is Theorem 8.23.

(1)⇒(5). Write $A = A_1 + \mathrm{i}A_1 - A_3 - \mathrm{i}A_4$, with $A_i \in B^1(\mathcal{V})$. We have $\sum(v_i|A_k v_i) < \infty$. Thus $(v_i|Av_i)$ is a linear combination of 4 absolutely convergent series.

(1)⇐(5). First assume that A is self-adjoint. Then $A = A_+ - A_-$ with $A_+A_- = A_-A_+ = 0$ and $A_+, A_- \in B_+(\mathcal{V})$. We have the decomposition $\mathcal{V} = \mathrm{Ran}\, 1_{]-\infty,0[}A) \oplus \mathrm{Ker}A \oplus \mathrm{Ran}\, 1_{]0,\infty[}A)$. Let $(v_1^-, v_2^-, \dots, v_1^0, v_2^0, \dots, v_1^+, v_2^+, \dots)$ be a basis that respects this decomposition. Then we compute that

$$\infty > \sum_{\epsilon=-,0,+} \sum_i |(v_i^\epsilon|Av_i^\epsilon)| = \mathrm{Tr}A_+ + \mathrm{Tr}A_-.$$

Thus $A_+, A_- \in B^1_+(\mathcal{V})$. Hence $A \in B^1(\mathcal{V})$.

If A is not necessarily self-adjoint, then consider $\mathrm{Re}A := \frac{1}{2}(A + A^*)$, $A := \frac{1}{\mathrm{i}2}(A - A^*)$. Then

$$\sum |(v_i|\mathrm{Re}Av_i)| + \sum |(v_i|\mathrm{Im}Av_i)| \leq 2\sum |(v_i|Av_i)| < \infty$$

Thus (5) is satisfied for $\mathrm{Re}A$, $\mathrm{Im}A$, and hence $\mathrm{Re}A, \mathrm{Im}A \in B^1(\mathcal{V})$. But $A = \mathrm{Re}A + \mathrm{i}\mathrm{Im}A$. □

For $A \in B^1(\mathcal{V})$ we set

$$\|A\|_1 := \mathrm{Tr}|A| = \sum_{n=1}^{\infty} \mathrm{s}_n(A).$$

Theorem 8.26 (1) *If* $A \in B^1(\mathcal{V})$, $B \in B(\mathcal{V})$, *then*

$$\|AB\|_1 \leq \|A\|_1\|B\|, \quad \|BA\|_1 \leq \|A\|_1\|B\|.$$

(2) $B^1(\mathcal{V})$ *is a Banach algebra.*

Proof. (1) Let $BA = W|BA|$ be the polar decomposition of BA and $A = U|A|$ be the polar decomposition of A. Note that $BU|A|^{1/2} \in B^2(\mathcal{V})$. Thus

$$\mathrm{Tr}|BA| = \mathrm{Tr}W^*BU|A|^{1/2}|A|^{1/2}| \leq \|W^*BU|A|^{1/2}\|_2\||A|^{1/2}\|_2.$$

Now

$$\||A|^{1/2}\|_2 = (\mathrm{Tr}|A|)^{1/2},$$
$$\|W^*BU|A|^{1/2}\|_2 \leq \|W^*BU\|\||A|^{1/2}\|_2 \leq \|B\|(\mathrm{Tr}|A|)^{1/2}.$$

(2) Let us prove the subadditivity. Let $A, B \in B^1(\mathcal{V})$ and $A + B = W|A + B|$ be the polar decomposition of $A + B$. Then, using $|A + B| = W^*(A + B)$,

$$\begin{aligned} \|A + B\|_1 = & \quad \mathrm{Tr}W^*(A + B) \\ \leq & \quad |\mathrm{Tr}W^*A| + \mathrm{Tr}W^*B| \quad \leq \|W^*\|\mathrm{Tr}|A| + \|W^*\|\mathrm{Tr}|B| = \mathrm{Tr}|A| + \mathrm{Tr}|B|. \end{aligned}$$

Thus $B^1(\mathcal{V})$ is a normed space.

Using $\|A\| \leq \|A\|_1$ we see, that (1) implies

$$\|AB\|_1 \leq \|A\|_1\|B\|_1.$$

Thus $B^1(\mathcal{V})$ is a normed algebra.

Let A_n be a Cauchy sequence in the $\|\cdot\|_1$ norm. Then it is also Cauchy in the $\|\cdot\|$ norm. Thus there exists $\lim_{n\to\infty} A_n =: A \in B(\mathcal{V})$. Let $A - A_n = U_n|A - A_n|$ be the polar decomposition of $A - A_n$. Let P be a finite projection. Clearly, for fixed n, $\|A_m - A_n\|_1$ is a Cauchy sequence and thus $\lim_{m\to\infty} \|A_m - A_n\|_1$ exists.

$$\begin{aligned} \|P|A - A_n|P\|_1 \quad & = \mathrm{Tr}PU^*(A - A_n)P \\ & = \lim_{m\to\infty} \mathrm{Tr}PU^*(A_m - A_n)P \quad \leq \lim_{m\to\infty} \|A_m - A_n\|_1. \end{aligned}$$

Since P was arbitrary,

$$\|A - A_n\|_1 \leq \lim_{m\to\infty} \|A_m - A_n\|_1 \to 0.$$

Hence $B^1(\mathcal{V})$ is complete. □

Theorem 8.27 *Let* $x_1, x_2, \ldots$ *and* $y_1, y_2, \ldots$ *be sequences of vectors with* $\sum_{n=1}^{\infty} \|x_n\|^2 < \infty$, $\sum_{n=1}^{\infty} \|y_n\|^2 < \infty$. *Then* $\sum_{n=1}^{\infty} |y_n)(x_n|$ *is trace class.*

Proof. Let $e_1, e_2, \ldots$ be an orthonormal system. Define $A := \sum_{n=1}^{\infty} |x_n)(e_n|$, $B := \sum_{n=1}^{\infty} |y_n)(e_n|$. Then $\mathrm{Tr}A^*A = \sum \|x_n\|^2$ and $\mathrm{Tr}B^*B = \sum \|y_n\|^2$. Hence A, B are Hilbert-Schmidt. But $C = BA^*$. □

8.9 Duality in operator spaces

Let $\mathcal{V}$ be a Hilbert space.

Theorem 8.28 *For $B \in B^1(\mathcal{V})$ define $\rho(B) \in B_\infty(\mathcal{V})^\#$ by*

$$\langle \rho(B)|A\rangle := \mathrm{Tr} BA.$$

Then $\rho : B^1(\mathcal{V}) \to B_\infty(\mathcal{V})^\#$ is an isometric isomorphism.

Proof. It suffices to prove that ρ is onto. Let $\psi \in B_\infty(\mathcal{V})^\#$. Then $\psi \in B^2(\mathcal{V})^\#$. Therefore, there exists $B \in B^2(\mathcal{V})$ such that

$$\psi(A) = \mathrm{Tr} B^* A.$$

For any finite projection P and if $A = U|A|$ is the polar decompoition of A, then

$$\mathrm{Tr} P|B|P = \mathrm{Tr}|B|P = \mathrm{Tr} B^* UP = \psi(UP) \leq \|\psi\|.$$

Hence $\mathrm{Tr}|B| \leq \|\psi\|$. Thus $B \in B^1(\mathcal{V})$ and $\psi = \rho(B^*)$. □

Theorem 8.29 *For $A \in B(\mathcal{V})$ define $\pi(A) \in B^1(\mathcal{V})^\#$ by*

$$\pi(A)|B\rangle := \mathrm{Tr} AB, \quad B \in B^1(\mathcal{V}).$$

Then $\pi : B(\mathcal{V}) \to B^1(\mathcal{V})^\#$ is an isometric isomorphism.

Proof. It suffices to prove that π is onto. Let $\phi \in B^1(\mathcal{V})^\#$. We define the sequilinear form

$$\mathcal{V} \times \mathcal{V} \ni (w, v) \mapsto \phi\big(|v|(w|\big) \in \mathbb{C}.$$

Clearly, $\||v)(w|\|_1 = \|v\|\|w\|$. Thus the boundedness of ϕ means that

$$\big|\phi\big(|v)(w|\big)\big| \leq \|\psi\|\|v\|\|w\|.$$

Hence, there exists a unique $A \in B(\mathcal{V})$ such that

$$\phi(|v)(w|) = (w|Av),$$

and $\phi = \pi(A)$. □

8.10 Fredholm operators

Let $\mathcal{V}, \mathcal{W}$ be Hilbert spaces.

Theorem 8.30 *Let $T \in B(\mathcal{V}, \mathcal{W})$. TFAE:*

(1) *There exists a unique $S \in B(\mathcal{W}, \mathcal{V})$ such that $TS = 1 - P$, where P is the projection on* $\mathrm{Ker} T^*$ *and* $\dim P < \infty$.

(2) *There exists $S \in B(\mathcal{W}, \mathcal{V})$ such that $TS - 1 \in B_\infty(\mathcal{W})$.*

(3) *The image of T in $B(\mathcal{V}, \mathcal{W})/B_\infty(\mathcal{V}, \mathcal{W})$ is right invertible.*

(4) *$T\mathcal{V}$ is closed and* $\dim(T\mathcal{V})^\perp < \infty$.

If the above conditions are satisfied, then we say that T is right Fredholm. We say that T is left Fredholm iff T^* is right Fredholm. We say that T is Fredholm iff it is left and right Fredholm.

Theorem 8.31 *T is Fredholm iff $(TT^*)^{1/2}$ is Fredholm.*

If T is Fredholm, we define

$$\mathrm{index}T := \dim T - \dim T^* \in \mathbb{Z}.$$

Theorem 8.32 (1) *The set of Fredholm operators is open in $B(\mathcal{V}, \mathcal{W})$.*

(2) *Fredholm operators of a given index for a connected component of $B(\mathcal{V}, \mathcal{W})$.*

(3) *If T is Fredholm and K is compact, then $T + K$ is Fredholm and* $\mathrm{index}(T + K) = \mathrm{index}T$.

(4) *If T is Fredholm, then so is T^*, and* $\mathrm{index}T^* = -\mathrm{index}T$.

(5) *If T, S are Fredholm, then so is TS, and* $\mathrm{index}TS = \mathrm{index}T + \mathrm{index}S$.

Recall that $B_\infty(\mathcal{V})$ is a closed ideal of $B(\mathcal{V})$. Thus $B(\mathcal{V})/B_\infty(\mathcal{V})$ is a C^*-algebra. and we have the canonical $*$-homomorphism $\phi : B(\mathcal{V}) \to B(\mathcal{V})/B_\infty(\mathcal{V})$. It is called the Calkin algebra. Note that Fredholm operators are mapped by ϕ onto the group of invertible operators in the Calkin algebra and the index is a homomorphism of this group onto $\mathbb{Z}$.

References

[RS1] Reed, M., Simon, B.: Methods of Modern Mathematics, vol. 1, Academic Press 1972

[RS2] Reed, M., Simon, B.: Methods of Modern Mathematics, vol. 4, Academic Press 1978

[Ru] Rudin, W.: Functional Analysis, 2nd edition, McGraw-Hill, 1991

[Pe] Pedersen, G.: Analysis now, corrected 2nd printing, Springer New York, 1995

[TL] Taylor, A., Lay, D.: Introduction to Functional Analysis, 2nd edition, John Wiley and Sons, 1980