# Functional Analysis and Applications

## Lecture notes for MATH 797fn

**Luc Rey-Bellet**
**University of Massachusetts Amherst**

The functional analysis, usually understood as the *linear* theory, can be described as

**Extension of linear algebra to infinite-dimensional vector spaces using topological concepts**

The theory arised gradually from many applications such as solving boundary value problems, solving partial differential equations such as the wave equation or the Schrödinger equation of quantum mechanics, etc... Such problems lead to a comprehensive analysis of function spaces and their structure and of linear (an non-linear) maps acting on function spaces. These concepts were then reformulated in abstract form in the modern theory of functional analysis. Functional analytic tools are used in a wide range of applications, some of which we will discuss in this class.

# Chapter 1

# Metric Spaces

## 1.1 Definitions and examples

One can introduce a *topology* on some set $M$ by specifying a *metric* on $M$.

**Definition 1.1.** A map $d(\cdot,\cdot) : M \times M \to \mathbb{R}$ is called a *metric* on the set $M$ if for all $\xi, \eta, \zeta \in M$ we have

1. (*positive definite*) $d(\xi,\eta) \geq 0$ and $d(\xi,\eta) = 0$ if and if $\xi = \eta$.
2. (*symmetric*) $d(\xi,\eta) = d(\eta,\xi)$.
3. (*triangle inequality*) $d(\xi,\eta) \leq d(\xi,\zeta) + d(\zeta,\xi)$

**Example 1.2.** Some examples of metric spaces

1. $M = C[a,b]$ with $d(f,g) = \int_a^b |f(t) - g(t)|^p \, dt$ with $0 < p < \infty$.
2. $M = C[a,b]$ with $d(f,g) = \max_{t\in[a,b]} |f(t) - g(t)|$
3. For any measure space $(X,\mu)$, $M = L^p(X,\mu)$ with $d(f,g) = \|f - g\|_p$ with $1 \leq p \leq \infty$.
4. Let $M$ be the set of all infinite sequences $\xi = (x_1, x_2, \cdots)$ with $x_i \in \mathbb{C}$. Then for $\eta = (y_1, y_2, \cdots)$

$$d(\xi,\eta) = \sum_{i=1}^{\infty} \left(\frac{1}{2}\right)^i \frac{|x_i - y_i|}{1 + |x_i - y_i|} \tag{1.1}$$

   defines a metric on $M$.

5. Let $M$ be the set of all infinite sequences of 0 and 1: $M = \{\xi = (x_1, x_2, \cdots) \, ; \, x_i \in \{0,1\}\}$. Then

$$\begin{aligned} d(\xi,\eta) &= \sum_{i=1}^{\infty} (x_i + y_i \ (\mathrm{mod})2) \\ &= \text{number of indices } j \text{ at which } \xi \text{ and } \eta \text{ differ} \end{aligned} \tag{1.2}$$

   defines a metric on $M$ and is called the *Hamming distance* and is used in coding theory.

6. A metric can be defined on arbitrary space $M$ (without any linear structure), for example set

$$d(\xi, \eta) = \begin{cases} 0 & \text{if} \quad \xi = \eta \\ 1 & \text{if} \quad \xi \neq \eta \end{cases} \tag{1.3}$$

and $d$ is called the *discrete metric* and $M$ a discrete space.

In a metric space $(X, d)$ ones introduces naturally a concept of *convergence* as well as a *topology*.

**Definition 1.3.** 1. A sequence $\{\xi_i\}$ is called a *Cauchy sequence* if for any $\epsilon > 0$ there exists $N$ so that $d(\xi_i, \xi_j) < \epsilon$ for all $i, j \geq N$ (or shorter if $\lim_{i,j\to\infty} d(\xi_i, \xi_j) = 0$).

2. We say that $\xi$ is the limit of the sequence $\{\xi_i\}$ (or that $\xi_i$ converges to $\xi$, or that $\lim_{i\to\infty} \xi_i = \xi$) if $\lim_{i\to\infty} d(\xi_i, \xi) = 0$.

**Definition 1.4.** A metric space $(M, d)$ is called complete if every Cauchy sequence $\{\xi_i\}$ has a limit $\xi \in M$.

**Theorem 1.5.** *The following metric spaces are complete.*

*1. Let $M$ be a finite dimensional vector space with (arbitrary norm) $\|\cdot\|$ and metric $d(\xi, \eta) = \|\xi - \eta\|$.*

*2. $M = L^p(X, \mu)$ with $d(f, g) = \|f - g\|_p$*

*3. $M = C[a, b]$ with $d(f, g) = \max_{t\in[a,b]} |f(t) - g(t)|$.*

*Proof.* Consult your class notes for Math 624. ∎

**Definition 1.6.** A brief reminder on some topological concepts.

- In a metric space $(X, d)$

$$B_r(\xi) := \{\eta : d(\xi, \eta) < r\} \tag{1.4}$$

is the open ball of radius $r$ around $\xi$ and

$$\overline{B}_r(\xi) := \{\eta : d(\xi, \eta) \leq r\} \tag{1.5}$$

is the closed ball of radius $r$ around $\xi$

- A set $N \subset M$ is called *bounded* if there exist a ball $B$ such that $N \subset B$.
- A set $N \subset M$ is called *open* if for every point $\xi \in N$ there exists a open ball around $\xi$ contained in $N$.
- A set $N \subset M$ is called *closed* if $M \setminus N$ is open.
- For a set $N \subset M$ the set $\overline{N}$ is the smallest closed set which contains $N$. The set $\overline{N}$ is called the *closure* of $N$.
- A set $N \subset M$ is called *dense* (in $M$) if $\overline{N} = M$.
- A set $K \subset M$ is called *compact* (in $M$) if every sequence in $K$ contains a convergent subsequence with limit in $K$.

## 1.2 Banach fixed point theorem

**Problem 1.7.** As a motivation imagine we want to solve the fixed point problem

$$F(\xi) \, = \, \xi \tag{1.6}$$

where $F : M \to M$ is some map (not necessarily linear).

The idea of the solution is simple. Pick a point $\xi_0$ and define the sequence $\xi_n$ inductively by

$$\xi_{n+1} = F(\xi_n)\,. \tag{1.7}$$

If this sequence converges to $\xi$ and $F$ is continuous we have then

$$\xi \, = \, \lim_{n\to\infty} \xi_{n+1} \, = \, \lim_{n\to\infty} F(\xi_n) \, = \, F(\lim_{n\to\infty} \xi_n) \, = \, F(\xi) \tag{1.8}$$

and so $\xi$ is a solution of the fixed point problem.

We have the following

**Theorem 1.8.** *(**Banach Fixed Point Theorem**) Let $(M, d)$ be a complete metric space and let $F \, : \, M \to M$ be a contraction, i.e., there exists $q \in [0, 1)$ such that for all $\xi$, $\eta \in M$*

$$d(F(\xi), F(\eta)) \leq q d(\xi, \eta)\,. \tag{1.9}$$

*Then $F$ has exactly one fixed point $\xi = \lim_n F^n(\xi_0)$ for arbitrary $\xi_0$.*

*Proof.* We have

$$d(\xi_{n+1}, \xi_n) \, \leq \, q d(F(\xi_n), F(\xi_{n-1})) \, \leq \, \cdots \, \leq \, q^n d(\xi_1, \xi_0)\,. \tag{1.10}$$

Using that $q < 1$ we have then

$$\begin{aligned} d(\xi_{n+m}, \xi_n) \quad &\leq \quad \sum_{k=1}^{m} d(\xi_{n+k}, \xi_{n+k-1}) \\ &\leq \quad \sum_{k=1}^{m} q^{n+k-1} d(\xi_1, \xi_0) \\ &\leq \quad \frac{q^n}{1-q} d(\xi_1, \xi_0)\,. \end{aligned} \tag{1.11}$$

Since $(M, d)$ is complete, $\xi = \lim_{n\to\infty} \xi_n$ exists.

To show that $\xi$ is a fixed point we note that

$$d(f(\xi), \xi) \, \leq \, d(f(\xi), f(\xi_n)) + d(\xi_{n+1}, \xi) \leq q d(\xi, \xi_n) + d(\xi, \xi_{n+1}) \tag{1.12}$$

and the right hand side can be made arbitrarily small for $n$ large enough.

Finally to show uniqueness, if $\xi$ and $\eta$ are two fixed points then

$$d(\xi, \eta) \, = \, d(F(\xi), F(\eta)) \, \leq \, q d(\xi, \eta) < d(\xi, \eta)\,. \tag{1.13}$$

and this implies $\xi = \eta$. ■

Let us give a few applications of this theorem.

**Example 1.9.** Let us try to solve the set of linear equation

$$\sum_{i=1}^{n} a_{ik} x_k \;=\; z_k \tag{1.14}$$

where $z_k$, $k = 1, \cdots, n$ are given and $x_k$, $k = 1, \cdots, n$ are unknown. We rewrite it as a fixed point equation

$$x_i = \sum_{k=1}^{n} (a_{ik} + \delta_{ik}) x_k - z_i \tag{1.15}$$

or

$$\xi = F(\xi) \tag{1.16}$$

where $\xi = (x_1, \cdots, x_n) \in \mathbb{R}^n$ (or $\mathbb{C}^n$) and

$$F(\xi) = C\xi + \zeta \tag{1.17}$$

where $\zeta = (z_1, \cdots, z_n)$ and $C$ is the matrix with $c_{ik} = a_{ik} + \delta_{ik}$.

To apply the Banach fixed point theorem we pick a metric on $\mathbb{C}^n$ such that $(\mathbb{C}^n, d)$ is complete. For example we can take $d(\xi, \eta) \;=\; \|\xi - \eta\|_p$ with $p \geq 1$ and then we have

$$d(F(\xi), F(\eta)) \;=\; \|C(\xi - \eta)\|_p \tag{1.18}$$

For example if $p = 1$ we have

$$\|C\xi\|_1 \quad = \quad \sum_i \left| \sum_k c_{ik} x_k \right| \;\leq\; \sum_i \sum_k |c_{ik}||x_k| \leq \underbrace{\max_k \sum_I |c_{ik}|}_{:= q_1} \|\xi\|_1 \,. \tag{1.19}$$

If we can find a norm such that $\|C(\xi)\| \leq q\|\xi\|$ then the equation (1.9) has a unique solution.

The fixed point equation has the form $\xi \;=\; C\xi + \zeta$ or $(\mathbf{1} - C)\xi \;=\; \zeta$ which gives formally using a Neumann series (which we will justify later)

$$\xi \;=\; (\mathbf{1} - C)^{-1}\zeta \;=\; (1 + C + C^2 + \cdots)\zeta \,. \tag{1.20}$$

Note also that the Banach fixed point algorithms gives the sequence

$$\xi_{n+1} \;=\; C\xi_n + \zeta \tag{1.21}$$

and $\xi$ is the limit of $\xi_n$ independent of the starting point $\xi_0$. This iteration is easy to solve and gives

$$\xi_n \;=\; C^n \xi_0 + \sum_{k=1}^{n-1} C^k \zeta \tag{1.22}$$

and thus, formally, we find

$$\xi \;=\; \lim_{n\to\infty} C^n \xi_0 + \sum_{k=1}^{\infty} C^k \zeta \tag{1.23}$$

The second term is exactly the Neumann series and the first term should go to $0$.

**Example 1.10. (Fredholm integral equation)** A very similar argument applies to the equation

$$f(t) \;=\; \lambda \int_a^b k(t,s) f(s)\, ds + h(t) \tag{1.24}$$

where we use the metric space $C[a,b]$ with the maximum metric $d(f,g) \;=\; \int_a^b |f(t) - g(t)| dt$. This is a fixed point equation $f = F(f)$ where

$$F(f)(t) \;=\; h(t) + \lambda \int_a^b k(t,s) f(s)\, ds\,. \tag{1.25}$$

We have

$$\begin{aligned} d(F(f),F(g)) \;&=\; \max_{t\in[a,b]} |F(f)(t) - F(g)(t)| \;=\; \max_{t\in[a,b]} \left| \int_a^b \lambda k(t,s)(f(s) - g(s))\, ds \right| \\ &\leq\; |\lambda| \left( \max_{t\in[a,b]} \int_a^b |k(t,s)|\, ds \right) d(f,g)\,. \end{aligned} \tag{1.26}$$

From the Banach fixed point theorem we deduce that the Fredholm integral equation has a unique solution provided

$$|\lambda| \;\leq\; \left( \max_{t\in[a,b]} \int_a^b |k(t,s)|\, ds \right)^{-1} \tag{1.27}$$

**Example 1.11. (Solving differential equations)** Consider an ordinary differential equation (initial value problem)

$$\frac{dx(t)}{dt} \;=\; f(x(t))\,, x(a) = x_0 \tag{1.28}$$

where $f : \mathbb{R} \to \mathbb{R}$. It is easy to see that $x(t)$ is solution of (1.28) on the interval $[a,b]$ if and only if we have for $t \in [a,b]$

$$x(t) \;=\; x_0 + \int_a^t f(x(s))\, ds\,, \quad a \leq t \leq b\,. \tag{1.29}$$

We can interpret this equation as a fixed point equation in $M \;=\; \{x \in C[a,b] \;:\; x(0) = x_0\}$ which is a closed subspace of a Banach space and thus itself a Banach space. We define $F : M \to M]$ by

$$F(x)(t) \;=\; x_0 + \int_a^t f(x(s))\, ds\,, \quad a \leq t \leq b \tag{1.30}$$

and thus $x(t)$ is a solution of (1.28) on the interval $[a,b]$ if and only if $F(x)(t) = x(t)$ for $t \in [a,b]$.

In order to apply the Banach fixed point theorem we assume that $f$ is *globally Lipschitz*, i.e., there exists a constant $L > 0$ such that for all $x, y \in \mathbb{R}$

$$|f(x) - f(y)| \leq L|x - y| \tag{1.31}$$

Then if we equip $C[a,b]$ with the maximum metric we have

$$\begin{aligned} d(F(x),F(y)) \;&=\; \sup_{t\in[a,b]} |F(x(t)) - F(y(t))| \;\leq\; \sup_{t\in[a,b]} \int_a^t |f(x(s)) - f(y(s))|\, ds \\ &\leq\; L \int_a^t |x(s) - y(s)| \;\leq\; L(b-a)\, d(x,y)\,. \end{aligned} \tag{1.32}$$

We can apply the Banach fixed point theorem provided $q = L(b-a) < 1$. This means there is a unique solution of the differential equation on a suitable interval $[a,b]$ is $b$ is sufficiently close to $a$.

There are many generalizations of this result to systems of differential equations or even partial differential equations and the Lipschitz condition can be somewhat relaxed too.

## 1.3 Exercises

**Exercise 1.** Let $M$ be a complete metric space and let $F : M \mapsto M$. Suppose there exists a sequence $a_n$ is a sequence of non-negative numbers with $\sum_{n=1}^{\infty} a_n < \infty$ and

$$d(F^n(x), F^n(y)) \le a_n d(x,y)\,. \tag{1.33}$$

Show that $F$ has a unique fixed point.
*Hint: Modify the proof of Banach fixed point theorem*

**Exercise 2.** In this problem we consider the Volterra integral equation given by the fixed point equation $f = F(f)$ where

$$F(t)(t) \;=\; \lambda \int_a^t k(t,s) f(s)\, ds + h(s) \tag{1.34}$$

where $h(t) \in C[a,b]$ and $k(t,s) \in C[a,b] \times C[a,b]$ are given. We will show that this equation has a unique solution for any value of $\lambda$ (compare with the Fredholm integral equation).

Let $K$ the Volterra integral kernel be given by

$$Kf(t) \;=\; \lambda \int_a^t k(t,s) f(s)\, ds \tag{1.35}$$

1. Show that $F^n(f) \;=\; \sum_{k=0}^{n} K^k h + K^n f$.
2. Prove that $|K^n f(t)| \;\le\; C^n \frac{(t-a)^n}{n!} \sup_t |f(t)|$ for a suitable constant $C$.
3. Use the previous exercise to show the existence of a unique solution in $C[a,b]$ to the Volterra integral equation.

**Exercise 3.** Consider the differential equation (initial value problem)

$$\frac{dx(t)}{dt} \;=\; f(x(t))\,, x(0) = x_0 \tag{1.36}$$

where $f : \mathbb{R} \to \mathbb{R}$ is Lipschitz. Consider the metric space $C[0,\infty)$ with the metric

$$d(x,y) \;=\; \sup_{t \ge 0} e^{-Dt} |f(t)|\,. \tag{1.37}$$

Use this metric space for a suitable choice of $D$ to show that the initial value problem (1.36) has a unique solution for $t \in [0,\infty)$.

# Chapter 2

# Normed Vector Spaces

For general metric spaces $(M, d)$ the set $M$ has no structure besides the topology induced buy the norm. We concentrate now on the special case where

$$M = V = \text{ vector space over } \mathbb{K} \text{ with } \mathbb{K} = \mathbb{R} \text{ or } \mathbb{C}$$

## 2.1 Some concepts from linear algebra

We recall some basic concepts from linear algebra slightly generalized to vector spaces which may have infinite dimension. In this section we do not use any topological concepts yet.

**Definition 2.1.** 1. A set $M \subset V$ is called *linearly independent* if every *finite* subset of $M$ is linearly independent.

2. The set $E \subset V$ is called a *Hamel basis* (*algebraic basis*) of $V$ if $E$ is linearly independent and every vector $\xi \in V$ can be written uniquely as *finite linear combination* of elements in $E$.

Using Zorn's Lemma one can prove that

**Theorem 2.2.** *Let $M \subset V$ be linearly independent. Then $V$ has a Hamel basis which contains $M$.*

We use the following notations: for $N, M \subset V$, $\xi \in V$ and $a \in \mathbb{K}$

$$\begin{aligned} \xi + M &= \{\xi + \eta \,;\, \eta \in M\}, \\ N + M &= \{\xi + \eta \,;\, \xi \in N, \eta \in M\}, \\ aM &= \{a\xi \,;\, \xi \in M\}. \end{aligned}$$

**Definition 2.3.** 1. If $M$ and $N$ are subspaces of $V$ and $M \cap N = \{0\}$ then $M + N$ is a subspace and one write $M + N$ as $M \oplus N$ (*direct sum*).

2. If $V = M \oplus N$ we say that $M$ and $N$ are *complementary subspaces*.

**Proposition 2.4.** *To each subspace $M \subset V$ there exists a complementary subspace $N$ (not uniquely defined).*

*Proof.* Let $E_M$ a Hamel basis of $M$ and $E$ a Hamel basis of $V$ which contains $E_M$ (see the proof of Theorem 2.2). Then $E \setminus E_M$ generates a subspace $N$ which is complementary to $M$. ∎

**Definition 2.5.** If $V = M \oplus N$ and $\dim N = n < \infty$ then we say that $M$ has *codimension* $n$, i.e.,

$$\operatorname{codim} M = \dim M\,. \tag{2.1}$$

One check easily that

**Proposition 2.6.** *If* $V = M \oplus N_1 = M \oplus N_2$. *Then* $\dim N_1 = \dim \mathrm{N}_2$.

**Definition 2.7.** Suppose $V = M \oplus N$. Then any $\xi \in V$ has a unique decomposition

$$\xi \;=\; \alpha + \beta \tag{2.2}$$

with $\alpha \in M$, $\beta \in N$. The *projection of* $\xi$ *on* $M$ *along* $N$ is given by

$$P\xi = \alpha\,. \tag{2.3}$$

One verifies easily that

**Lemma 2.8.** *We have* $P^2 = P$.

Conversely we have

**Lemma 2.9.** *Suppose* $P\ :\ V \to V$ *is linear map such that* $P^2 = P$. *Then* $V = M \oplus N$ *where* $M = PV$ *and* $N = (\mathbf{1} - P)V$.

*Proof.* For any $\xi \in V$ we have

$$\xi = \underbrace{P\xi}_{\in M} + \underbrace{(\mathbf{1} - P)\xi}_{\in N} \tag{2.4}$$

and thus $V = M + N$. The sum is direct, since if

$$\xi \in M \cap N = PV \cap (\mathbf{1} - P)V \tag{2.5}$$

we have, on one hand, $\xi = P\alpha$ and so

$$P\xi = P^2\alpha = P\alpha = \xi\,. \tag{2.6}$$

On the other hand $\xi = (\mathbf{1} - P)\beta$ and so we obtain

$$\xi = P\xi = P(\mathbf{1} - P)\beta \;=\; (P - P^2)\beta \;=\; 0\,. \tag{2.7}$$

Thus $M \cap N = \{0\}$. ∎

Using projections we can prove

**Theorem 2.10.** *Suppose* $T\ : V \to W$ *is a linear map. Let* $P$ *a projection* **on the nullspace of** $T$ *and* $Q$ *a projection* **along the range of** $T$. *Then there exists a linear map* $S : W \to V$ *such that*

$$ST \;=\; \mathbf{1}_V - P\,, \quad TS \;=\; \mathbf{1}_W - Q\,. \tag{2.8}$$

*The map* $T$ *is bijective if and only* $P = Q = 0$.

*Proof.* Let us denote $N = \mathcal{N}(T) \subset V$ the nullspace (or kernel) of $T$ and and $M = \mathcal{R}(T) \subset W$ the range of $T$. We have

$$V = N \oplus V_1 , \quad W = W_1 \oplus M . \tag{2.9}$$

and $P$ is the projection on $N$ along $V_1$ and $Q$ the projection on $W_1$ along $M$.

We define $T_0 : V_1 \to M$ by $T_0\xi = T\xi$. Then $T_0$ is bijective and so $T_0^{-1} : M \to V_1$ exists. If

$$S = T_0^{-1}(\mathbf{1} - Q) \, : \, W \to V_1 \tag{2.10}$$

we have

$$ST\xi \, = \begin{cases} 0 & \text{if } \xi \in V \\ \xi & \text{if } \xi \in V_1 \end{cases} \tag{2.11}$$

and so

$$ST = \mathbf{1}_V - P . \tag{2.12}$$

Arguing similarly we have

$$TS = \mathbf{1}_W - Q . \tag{2.13}$$

∎

## 2.2 Norm

Suppose we have a metric $d$ on a vector space $V$. It is natural to ask whether the metric respects the linear structure of $V$. By that we mean that

1. $d$ is translation invariant, i.e., $d(\xi + \alpha, \eta + \alpha) \, = \, d(\xi, \eta)$ for all $\alpha \in V$.

2. Under scalar multiplication we have $d(a\xi, a\eta) \, = \, |a| d(\xi, \eta)$.

Property 1. implies that $d(\xi, \eta) = d(\xi - \eta, 0)$. If we set then $\|\xi\| := d(\xi, 0)$ we have then from the properties of the distance that

(N1) $\|\xi\| \geq 0$ and $\xi = 0$ if and only if $\xi = 0$.

(N2') $\|\xi\| = \| - \xi\|$ (symmetry)

(N3) $\|\xi + \eta\| \leq \|\xi\| + \|\eta\|$

while from Property 2., instead of (N2') we obtain the stronger

(N2) $\|a\xi\| = |a|\|\xi\|$

**Definition 2.11. (Normed vector spaces)**

1. A map $\| \cdot \| : V \to \mathbb{R}$ is called a *norm* on $V$ if it satisfies the condition (N1), (N2), and (N3).

2. The pair $(V, \| \cdot \|)$ is called a *normed vector space*.

3. A complete normed vector space is called a *Banach space*.

Note that in a normed vector space we have convergence in the sense of the metric defined by $d(\xi, \eta) = \|\xi - \eta\|$ or equivalently $\xi_n$ converges to $\xi$ if and only if $\lim_{n\to\infty} \|\xi - \xi_n\| = 0$.

**Example 2.12.** We give examples of normed vector spaces (see Math 623-624 for details and proofs).

1. For $p \geq 1$,
$$l^p \;=\; \big\{\xi = (x_1, x_2, x_3, \cdots)\, x_i \in \mathbb{C} \text{ with } \|\xi\|_p := (\sum_i |x_i|^p)^{1/p} < \infty\big\} \tag{2.14}$$
is a Banach space. The completeness is non-trivial as is the triangle inequality (a.k.a Minkowsky inequality).

2. The space
$$l^\infty \;=\; \big\{\xi = (x_1, x_2, x_3, \cdots)\, x_i \in \mathbb{C} \text{ with } \|\xi\|_\infty := \sup_i |x_i| < \infty\big\} \tag{2.15}$$
is a Banach space as well as
$$c_0 \;=\; \big\{\xi \in l^\infty\,, \xi = (x_1, x_2, x_3, \cdots)\; \lim_i x_i = 0\big\} \tag{2.16}$$
and
$$c \;=\; \big\{\xi \in l^\infty\,, \xi = (x_1, x_2, x_3, \cdots)\; \lim_i x_i \text{ exists}\big\} \tag{2.17}$$

3. The space
$$C^n[a,b] \;=\; \big\{f : [a,b] \to \mathbb{R} \,:\, f \text{ n times continuously differentiable } \big\} \tag{2.18}$$
is a Banach space with the norm
$$\|f\| \;=\; \sum_{k=0}^{n} \max_{t \in [a,b]} |f^{(k)}(t)| \tag{2.19}$$
where $f^{(k)}$ is the $k$-th derivetive of $f$.

4. The space
$$BV[a,b] \;=\; \big\{f : [a,b] \to \mathbb{R} \,:\, f \text{ of bounded variation}\big\} \tag{2.20}$$
is a Banach space with the norm
$$\|f\| \;=\; |f(a)| + V(f) \tag{2.21}$$
where $V(f)$ is the variation of $f$ on $[a,b]$ given by
$$V(f) \;=\; \sup_{\mathcal{P}} \sum_{k=1}^{N} |f(x_k) - f(x_{k-1}| \tag{2.22}$$
where $\mathcal{P}$ is set of partition of $[a,b]$: $a = x_0 < x_1 < \cdots < x_n = b$.

The following facts are very easy but also very important

**Proposition 2.13.** *Let $(V, \|\cdot\|)$ be a normed vector space.*

1. *The linear operations are continuous*
2. *The map $\xi \to \|\xi\|$ is continuous.*

*Proof.* If $\xi_n \to \xi$ and $\eta_n \to \eta$ then $\xi_n + t\eta_n \to \xi + t\eta$ since

$$\|(\xi + t)\eta - (\xi_n + t\eta_n\| \leq \|\xi - \xi_n\| + |t|\|\eta - \eta_n\| \,. \tag{2.23}$$

If $\xi_n \to \xi$ then $\|\xi_n\| \to \|\xi\|$ since

$$|\|\xi\| - \|\xi_n\|| \leq \|\xi - \xi_n\| \tag{2.24}$$

by the (reverse) triangle inequality. ■

**Definition 2.14.** (**Comparison of norms**) Suppose two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are given on a vector space $V$.

1. The norm $\|\cdot\|_1$ is *stronger* than $\|\cdot\|_2$ if there exists $C > 0$ such that for all $\xi \in V$.

$$\|\xi\|_2 \;\leq\; C\|\xi\|_1 \tag{2.25}$$

2. The The norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are *equivalent* if there exists constants $C$ and $D$ such that

$$D\|\xi_1\| \leq \|\xi\|_2 \;\leq\; C\|\xi\|_1 \,. \tag{2.26}$$

Clearly equivalent norm induce the same topology since convergence of one sequence in one norm implies the convergence of the sequence in the other norm.

**Theorem 2.15.** *If* $\dim(V) < \infty$ *then all norms on* $V$ *are equivalent.*

*Proof.* : This is left as an exercise. Use Bolzano-Weierstrass.

## 2.3 Continuous linear maps

Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be two normed vector space. In the sequel, whenever there is no risk of confusion we shall drop the index $V$ or $W$ from the norm and simply denote it by $\|\cdot\|$. We also consider a *linear map*

$$T \,:\, V \to W \,. \tag{2.27}$$

Unless explicitly specified we will deal only with linear maps in the sequel. We also use the notation $T\xi$ for $T(\xi)$

**Definition 2.16.** Let $T \,:\, V \to W$ be a linear map.

1. The map $T$ is *bounded* if there exists a constant $C \geq 0$ such that

$$\|T\xi\| \;\leq\; C\|\xi\| \tag{2.28}$$

   for all $\xi \in V$.

2. The *norm* of $T$, denoted by $\|T\|$ is the smallest $C$ such that (2.28) holds, i.e.

$$\|T\| \;:=\; \sup_{\xi \in V, \xi \neq 0} \frac{\|T\xi\|}{\|\xi\|} \;=\; \sup_{\xi \in V, \|\xi\|=1} \|T\xi\| \,. \tag{2.29}$$

3. The set of bounded linear maps is denoted by

$$\mathcal{L}(V,W) \;=\; \big\{T\,:\,V\to W,\, T \text{ linear and bounded}\big\} \tag{2.30}$$

and we write

$$\mathcal{L}(V) := \mathcal{L}(V,V)\,. \tag{2.31}$$

**Theorem 2.17.** *Let $T$ be a linear map. Then $T$ is bounded if and only if $T$ is continuous.*

*Proof.* Suppose $T$ is bounded, then we have

$$\|T\xi - T\,\eta\| \;=\; \|T(\xi-\eta)\| \;\le \|T\|\|\xi-\eta\| \tag{2.32}$$

and this implies that $T$ is (Lipschitz) continuous.

Conversely let us assume that $T$ is continuous but not bounded. Then there exists a sequence $\xi_n$ such that

$$\|T\xi_n\| \ge n\|\xi_n\|\,. \tag{2.33}$$

Let us set

$$\eta_n = \frac{\xi_n}{\sqrt{n}\|\xi_n\|}\,.$$

Then we have

$$\|\eta_n\| = \frac{1}{\sqrt{n}}\,, \quad \|T\eta_n\| > n\|\eta_n\| \;=\; \sqrt{n}\,. \tag{2.34}$$

This means that $\eta_n \to 0$ but $T\eta_n$ is divergent which contradicts the continuity of $T$. ∎

**Example 2.18.** Let us consider some examples of bounded (and not bounded operators). Many more examples to come.

1. The identity operator $\mathbf{1}$ defined by $\mathbf{1}\xi = \xi$ is bounded with $\|\mathbf{1}\| = 1$.

2. The differentiation operator is not bounded. Take for example the space $V$ which consists of polynomials $p(t)$ on $[0,1]$ with the sup-norm and set $Tp(t) = p'(t)$. Then if $p_n(t) = x^n$ we have $\|p_n\| = 1$ for all $n$ but $\|Tp_n\| = \|nt^{n-1}\| = n$ and this shows that $T$ is not bounded.

   Note that differentiation is a very natural operation so to include it in our consideration we will consider unbounded linear operator later on.

3. The integral operator

$$Tf(t) \;=\; \int_a^b k(t,s)f(s)\,ds \tag{2.35}$$

   is bounded on $C[a,b]$ equipped with the sup-norm if $k(t,s) \in C([a,b]\times[a,b])$. Indeed we $Tf(t)$ is continuous in $t$ by e.g. the dominated convergence theorem and

$$\|Tf\| \;=\; \sup_t |\int_a^b k(t,s)f(s)\,ds| \;\le\; \sup_s |f(s)| \underbrace{\sup_t \int_a^b |k(t,s)|\,ds}_{\equiv C} \tag{2.36}$$

and $C$ is finite since $k$ is continuous. Actually we can show that $\|T\| = \sup_t \int_a^b |k(t,s)|\,ds$. If $k(t,s)$ is nonnegative then we simply take $f = 1$ and we have $\|T1\| = \sup_t \int_a^b |k(t,s)|\,ds$. For a general $k$ we pick $t_0$ such that $\sup_t \int_a^b |k(t,s)|\,ds = \int_a^b |k(t_0,s)|\,ds$. Then we would like to pick the function $f(s) = \operatorname{sign} k(t_0,s)$ so that $Tf(t_0) = \int_a^b k(t_0,s)f(s)ds = \int_a^b |k(t_0,s)|\,ds$. But this $f$ is not continuous and we need to use an approximation argument. Consider the function $\phi_n(t)$ which is piecewise linear, increasing, continuous, with $\phi(t) = \operatorname{sign} t$ if $|t| > 1/n$. We set $f_n(s) = \phi_n(k(t_0,s))$. We have

$$Tf_n(t_0) \;=\; \int_a^b k(t_0,s)f_n(s)\,ds \;=\; \int_a^b |k(t_0,s)|\,ds - \frac{1}{n}(b-a)\,. \tag{2.37}$$

From this it follows that $\|T\| = \int_a^b |k(t_0,s)|\,ds$. ∎

4. The Fourier transform is defined by

$$\hat{f}(k) \;=\; \int_{\mathbb{R}} f(x)e^{-i2\pi kx}\,, \tag{2.38}$$

and we write $T(f) = \hat{f}$. In Math 623-624 ones proves (after some effort) that $T : L^1(\mathbb{R}, dx) \to C_0(\mathbb{R})$ is a bounded operator with norm 1. Here $C_0(\mathbb{R})$ is the Banach space of continuous functions such that $\lim_{|x|\to\infty} |f(x)| \to 0$ equipped with the sup-norm. One also proves (after some more effort) that $T : L^2(\mathbb{R}, dx) \to L^2(\mathbb{R}, dx)$ defines an unitary transformation (i.e. $T$ is invertible and $\|Tf\| = \|f\|$ for al $f$).

**Theorem 2.19.** *$\mathcal{L}(V,W)$ is a normed vector space.*

*Proof.* Let $T$, $S \in V$ and $a$ a scalar.

- (N1) We have $\|T\| \geq 0$. If $\|T\| = 0$ then we have $\|T\xi\| \leq \|T\|\|\xi\| = 0$ and so $T = 0$. Conversely if $T = 0$ then obviously $\|T\| = 0$.
- (N2) $\|aT\| \;=\; \sup_{\|\xi\|=1} \|aT\xi\| \;=\; |a| \sup_{\|\xi\|=1} \|T\xi\| \;=\; |a|\|T\|$
- (N3) $\|(T+S)\xi\| \;\leq\; \|T\xi\| + \|S\xi\|$ and thus $\|T+S\| \leq \|T\| + \|S\|$.

∎

**Theorem 2.20.** *If $W$ is a Banach space then $\mathcal{L}(V,W)$ is a Banach space.*

*Proof.* Let $T_n$ be Cauchy sequence in $\mathcal{L}(V,W)$. For any $\xi \in V$ we have

$$\|T_n\xi - T_m\xi\| \;\leq\; \|T_n - T_m\|\|\xi\| \tag{2.39}$$

and thus $T_n\xi$ is a Cauchy sequence in $W$. Since $W$ is complete this sequence has a limit in $W$ and we set

$$\eta \;=\; \lim_{n\to\infty} T_n\xi \tag{2.40}$$

We define then $T : V \to W$ by $T\xi = \eta$. The linearity of $T_n$ immediately implies that $T$ is linear. What remains to prove is that $T$ is bounded and $T_n \to T$. Given $\epsilon > 0$ pick $N$ such that for $n, m > N$ we have

$$\|T_n\xi - T_m\xi\| \leq \|T_n - T_m\|\|\xi\| \leq \epsilon\|\xi\| \tag{2.41}$$

Using the continuity of the norm and taking $m \to \infty$ in (2.41) we have

$$\|T_n\xi - T\xi\| \leq \epsilon\|\xi\| \tag{2.42}$$

for any $n \geq N$. This implies that $\|T - T_N\| \leq \epsilon$ and so $\|T\| \leq \|T - T_N\| + \|T_N\| < \infty$ and so $T$ is bounded. This also implies that $\|T - T_n\| \leq \epsilon$ for $n \geq N$ and so $T_n \to T$. ■

**Theorem 2.21.** *For $S \in \mathcal{L}(U, V)$ and $T \in \mathcal{L}(V, W)$ we have $TS := T \circ S \in \mathcal{L}(U, W)$ and*

$$\|TS\| \leq \|T\|\|S\| \tag{2.43}$$

*Proof.* $\|TS\| = \sup_{\|\xi\|=1} \|TS\xi\| \leq \|T\| \sup_{\|\xi\|=1} \|S\xi\| = \|T\|\|S\|$. ■

**Corollary 2.22.** *For $T \in \mathcal{L}(V)$ we have*

1. $\|T^n\| \leq \|T\|^n$
2. *The limit $r(T) = \lim_{n\to\infty} \|T^n\|^{1/n}$ exists and is called the spectral radius of $T$.*
3. $r(T) = \inf_n \|T^n\|^{1/n} \leq \|T\|$

*Proof.* 1. is immediate. For 2. and 3. set $c_n = \|T^n\|/\|T\|^n$. Then we have $0 \leq c_n \leq 1$ and $c_{n+m} \leq c_n c_m$ and this implies that $c_n$ is a bounded decreasing sequence and so $c = \lim_{n\to\infty c_n}$ exists. Therefore $c_n^{1/n}$ is also a decreasing with limit

$$\lim_{n\to\infty} \frac{\|T^n\|^{1/n}}{\|T\|} = \inf_n \frac{\|T^n\|^{1/n}}{\|T\|} \tag{2.44}$$

■

The notation "spectral radius" will become clear in the sequel.

As an application we consider the **Neumann series** which is is a generalization of the geometric series. Suppose $V$ is a Banach space, $T \in \mathcal{L}(V)$ with $\|T\| = \delta < 1$. Then the series $\sum_{k=0}^{\infty} T^k$ converges since since the partial sum $S_n = \sum_{k=0}^{\infty} T^k$ satisfy for $n > m$

$$\|S_n - S_m\| \leq \sum_{k=m+1}^{n} \|T^k\| \leq \sum_{k=m+1}^{n} \delta^k \leq \delta^{m+1}\frac{1}{1-\delta} \tag{2.45}$$

and thus form a Cauchy sequence and we have $\lim_n S_n = \sum_{k=0\infty} T^k$.

Furthermore we have

$$(\mathbf{1} - T)S_n = S_n(\mathbf{1} - T) = \mathbf{1} - T^{n+1} \to \mathbf{1} \tag{2.46}$$

and thus we conclude that $\mathbf{1} - T$ is invertible and

$$(\mathbf{1} - T)^{-1} = \sum_{k=0}^{\infty} T^k\,. \tag{2.47}$$

One can prove the stronger result

**Theorem 2.23.** *Let $V$ be a Banach space and $T \in \mathcal{L}(V)$ with $r(T) < 1$. Then $(\mathbf{1} - T)^{-1}$ exists and is given by the Neumann series*

$$(\mathbf{1} - T)^{-1} \;=\; \sum_{k=0}^{\infty} T^k \tag{2.48}$$

*Proof.* Note first that $\sum T^K$ converges whenever $\sum_k \|T^k\|$ does. If $r(T) = \delta < 1$ then for any $\delta'$ with $\delta < \delta' < 1$ there exists $N$ such that for $n > N$ we have $\|T^n\|^{1/n} \le \delta'$. This means that $\|T^n\| \le (\delta')^n$ and so $\sum_k \|T^k\|$ converges. The rest is as before. ∎

Linear operators occur naturally as *derivative* and this is true in infinite-dimensional spaces as well.

**Definition 2.24.** Let $V$ and $W$ be normed vector spaces and $U \subset V$ open. The map $F : U \to W$ is said to be *differentiable* at $\eta \in U$ if there exists a continuous linear map $F'(a)$ such that

$$F(\xi) = F(\eta) + F'(\eta)(\xi - \eta) + R(\xi)\|\xi - \eta\|\,, \tag{2.49}$$

where $R : U \to W$ satisfy $\lim_{\xi \to \eta} R(\xi) = 0$.

Let is consider a few examples

**Example 2.25.** 1. Let $V = W = C[a, b]$ and let $k \in C[0, 1] \times [0, 1]$ and $g$ be twice continuously differentiable. Let us consider the map

$$F(x)(t) \;=\; \int_a^b k(t, s) g(x(s))\, ds \tag{2.50}$$

To compute the derivative we pick $h \in C[0, 1]$ and using the mean value theorem we obtain

$$\begin{aligned} F(x + h) - F(x) \;&=\; \int_a^b k(t, s) \left[g(x(s) + h(s)) - g(x(s))\right]\, ds \\ &=\; \int_a^b k(t, s) \left[g'(x(s))h(s) + \frac{1}{2} g''(a(s))h(s)^2\right]\, ds \end{aligned} \tag{2.51}$$

where $a(s)$ is a value between $x(s)$ and $x(s) + h(s)$. This computation shows that the linear $F'(x) : C[0, 1] \to C[0, 1]$ given by

$$F'(x)(h)(t) \;=\; \int_a^b k(t, s) g'(x(s)) h(s)\, ds \tag{2.52}$$

is a good candidate for the derivative. Clearly this map is bounded since $k(t, s)g'(x(s)) \in C[0, 1] \times [0, 1]$. Moreover we have the bound

$$\sup_t \left| \int_a^b k \frac{1}{2} k(t, s) g''(a(s)) h(s)^2\, ds \right| \;\le\; C\|h\|^2 \tag{2.53}$$

since $g''$ is bounded by assumption. This shows that $F$ is differentiable. ∎

2. Let $E$ be a Banach space and consider the Banach space $\mathcal{L}(E)$. We say that $T \in \mathcal{L}(E)$ is an isomorphism if $T$ is linear, bijective, continuous and if $A^{-1}$ is continuous. We will see in fact in Section 2.8 that by the open map theorem the continuity of $A^{-1}$ follows from the other assumptions.

   In any case the set
$$\mathcal{GL}(E) \,=\, \{T \in \mathcal{L}(E) \,:, T \text{ an isomorphism}\} \tag{2.54}$$
   is an open subset of $\mathcal{L}(E)$. Indedd if $T$ is an isomorphism then $T + H$ is an isomorphism provided $\|H\| \le \|T^{-1}\|^{-1}$. This follows from $T + H \,=\, T(\mathbf{1} + T^{-1}H)$, from $\|T^{-1}H\| \le \|T^{-1}\|\|H\| < 1$ and from the geometric series.

   Let us consider the map $F : \mathcal{GL}(E) \to \mathcal{L}(E)$ given by
$$F(T) \,=\, T^{-1}\,. \tag{2.55}$$
   We claim that the $F$ is differentiable and that we have
$$F'(T)H \,=\, -T^{-1}HT^{-1}\,. \tag{2.56}$$
   The continuity of $F'(T)$ is clear since we have $\|F'(T)H\| \le \|A^{-1}\|^2\|H\|$. Furthermore we have
$$\begin{aligned} F(T+H) - F(T) \quad &= \quad (T+H)^{-1} - T^{-1} \,=\, \left[(\mathbf{1} + T^{-1}H) - \mathbf{1}\right] T^{-1} \\ &= \quad \sum_{j=1}^{\infty}(-T^{-1}H)^j T^{-1} \end{aligned} \tag{2.57}$$
   if $\|H\| \le \|T^{-1}\|^{-1}$ (geometric series). The first term in the series is $-T^{-1}HT^{-1}$ and the remainder can be bounded by $\|T^{-1}\|^2\|H\|(1 - \|T^{-1}\|\|H\|)^{-1}$. The reminder divided by $\|H\|$ tends to $0$, showing the differentiability. ∎

**Pointwise (strong) convergence:** In $\mathcal{L}(V, W)$, in addition to norm convergence ( $\equiv$ uniform convergence) there is the weaker notion of point wise convergence.

**Definition 2.26.** Let $\{T_n\}$ a sequence in $\mathcal{L}(V, W)$ and $T \in \mathcal{L}(V, W)$. We say that $T_n$ *converges strongly* to $T$ and write
$$T \,=\, s - \lim_{n\to\infty} T_n \tag{2.58}$$
if for any $\xi \in V$ we have
$$\lim_{n\to\infty} T_n\xi \,=\, T\xi\,. \tag{2.59}$$

Whenever we need to differentiate between different types of convergence we will write $n - \lim_{n\to\infty} T_n$ for the convergence in norm.

The following two lemmas are very easy and the proof is left to the reader.

**Lemma 2.27.** *If* $n - \lim_{n\to\infty} T_n \,=\, T$ *then* $s - \lim_{n\to\infty} T_n \,=\, T$.

**Lemma 2.28.** *The linear operations are continuous with respect to strong convergence.*

However note that if $n - \lim_{n\to\infty} T_n \,=\, T$ then $T_n\| \to \|T\|$ but this does not necessarily hold in the case of strong convergence. Similarly if $n - \lim_{n\to\infty} T_n \,=\, T$ and $n - \lim_{n\to\infty} S_n \,=\, S$ then we have $n - \lim_{n\to\infty} S_nT_n \,=\, ST$ but this does not necessarily hold for strong convergence.

**Example 2.29.** Let $V = l^p$ and for $\xi = (x_1, x_2, \cdots)$ let us set

$$T_n\xi \;=\; (x_1, \cdots, x_n, 0, 0, \cdots)\,. \tag{2.60}$$

Then we have $T_n\xi - \xi \to 0$ and so $s - \lim_{n\to\infty} T_n \;=\; \mathbf{1}$. But since

$$(\mathbf{1} - T_n)\xi \;=\; (0, \cdots, 0, x_{n+1}, x_{n+2}, \cdots) \tag{2.61}$$

for any $n$ we can find $\eta_n$ with $\|\eta_n = 1\| = 1$ and $(\mathbf{1} - T_n)\eta_n = \eta_n$. This implies that $\|\mathbf{1} - T_n\| = 1$ for all $n$ and so $n - \lim_{n\to\infty} T_n$ is certainly not equal to $\mathbf{1}$.

## 2.4 Linear functionals and dual spaces

**Definition 2.30.** If $V$ is a Banach space over $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$ then $\mathcal{L}(V, \mathbb{K})$ is a Banach space. We write

$$V' \;:= \mathcal{L}(V, \mathbb{K}) \tag{2.62}$$

and $V'$ is called the *dual space*. The elements of $V'$ are called *linear functionals on* $V$ and for $\lambda \in V'$ the norm is given by

$$\|\lambda\| \;=\; \sup_{\substack{\xi\in V\\ \|\xi\|=1}} |\lambda(\xi)|\,. \tag{2.63}$$

Let us work out an example in detail

**Example 2.31.** If $V = l^p$ for $p > 1$ then $V' = l^q$ where $p^{-1} + q^{-1} = 1$.

*Proof.* Note first that if $\eta = (y_1, y_2, \cdots) \in l^q$ then $\lambda_\eta(\xi)$ defined by

$$\lambda_\eta(\xi) \;=\; \sum_{k=1}^{\infty} x_k y_k \tag{2.64}$$

defines a bounded linear functional on $l^p$ since by Hölder's inequality we have

$$|\lambda_\eta(\xi)| \;\le\; \|\xi\|_p \|\eta\|_q\,, \tag{2.65}$$

and so $\|\lambda_\eta\| \le \|\eta\|_q$.

Let us denote

$$\epsilon_k \;=\; \{\delta_{ik}\}_{i=1}^{\infty} \in l^p\,. \tag{2.66}$$

Any $\xi \in l^p$ can be written as the convergent series in $l^p$

$$\xi \;=\; \sum_{k=1}^{\infty} x_k \epsilon_k\,. \tag{2.67}$$

Let $\lambda \in V$, since it is continuous, we have

$$\lambda(\xi) \;=\; \sum_k x_k y_k \text{ with } y_k = \lambda(\epsilon_k)\,. \tag{2.68}$$

and we set $\eta = \{y_k\}$ and $\lambda$ has the from $\lambda_\eta$.

Let us pick $\xi^{(n)} = \{x_k^{(n)}\}$ with

$$x_k^{(n)} = \begin{cases} \frac{|y_k|^q}{y_k} & \text{if } k \leq n \text{ and } y_k \neq 0 \\ 0 & \text{otherwise} \end{cases} . \tag{2.69}$$

We have then

$$\begin{aligned} \lambda(\xi^{(n)}) &= \sum_{k=1}^{n} |y_k|^q \\ &\leq \|\lambda\| \|\xi^{(n)}\|_p \\ &= \|\lambda\| \left( \sum_{k=1}^{n} |x_k^{(n)}|^p \right)^{1/p} \\ &= \|\lambda\| \left( \sum_{k=1}^{n} |y_k|^{(q-1)p} \right)^{1/p} \\ &= \|\lambda\| \left( \sum_{k=1}^{n} |y_k|^q \right)^{1-1/q} . \end{aligned} \tag{2.70}$$

We obtain then

$$\left( \sum_{k=1}^{n} |y_k|^q \right)^{1/q} \leq \|\lambda . \tag{2.71}$$

Since $n$ is arbitrary we conclude that $\eta \in l^q$ and $\|\eta\|_q \leq \lambda$. Combining with Hölder's inequality shows that $\|\Lambda_\eta\| = \|\eta\|_q$ and that the map $\lambda \mapsto \eta$ from $(l_p)'$ to $l^q$ is a norm preserving isomorphism. ∎

**Example 2.32.** In a similar way one shows that $L^p(X,\mu)' = L^q(X,\mu)$ for an arbitrary measure space $(X,\mu)$ and $1 < p < \infty$.

**Example 2.33.** With some minor modifications the same proof works for $l^1$ and we have $(l^1) = l^\infty$. However it is not true that $(l^\infty)' = l^1$ but rather we have

$$l^1 = (c_0)' \tag{2.72}$$

*Proof.* Let $\eta \in l^1$ and $\xi \in c_0$. Then the series $\lambda_\eta(\xi) = \sum_k x_k y_k$ converges and by Hölder's equality we have

$$|\lambda_\eta(\xi)| \leq \|\xi\|_\infty \|\eta\|_1 . \tag{2.73}$$

Thus we $\|\lambda_\eta\| \leq \|\eta\|_1$.

For any $\epsilon > 0$, let $N$ be such that $\sum_{k \geq N} |y_k| < \epsilon$. Pick then $\xi^{(N)} = \{x_k^{(N)}\}$ with

$$x_k^{(N)} = \begin{cases} \frac{|y_k|}{y_k} & \text{if } k \leq n \text{ and } y_k \neq 0 \\ 0 & \text{otherwise} \end{cases} . \tag{2.74}$$

We have then

$$|\lambda_\eta(\xi)| = \sum_{k \geq N} |y_k| \geq (\|\eta\|_1 - \epsilon)\|\xi\| . \tag{2.75}$$

Therefore $\|\lambda_\eta\| \geq \|\eta\|_1$ and so $\|\lambda_\eta\| = \|\eta\|_1$.

It remains to show that every linear functional on $c_0$ has the above form. If $\xi \in c_0$ then we can write $\xi = \sum_{k=1}^{\infty} x_k \epsilon_k$ (not true in $l^\infty$!) and so for any $\lambda \in c_0'$ there exists $\eta = \{y_k\}$ such that $\lambda(\xi) = \lambda_\eta(\xi) = \sum_{k=1}^{\infty} x_k y_k$. Picking $\xi^{(n)}$ as in (2.74) we have $\|\xi^{(n)}\|_\infty = 1$ and

$$\|\lambda_\eta(\xi^{(n)}) \;=\; \left(\sum_{k=1}^{n} |y_k|\right) \|\xi^{(n)}\|_\infty \tag{2.76}$$

from which we conclude that $\|\lambda_\eta\| \geq \|\eta\|_1$, i.e. $\eta \in l^1$. ∎

**Schauder basis and separability:**

**Definition 2.34.** A metric space is called *separable* if it contains a countable dense set.

One shows that $l^p$ is separable but $l^\infty$ is not separable.

In the construction of dual spaces we used the fact that any $\xi \in l^p$, $1 \leq p < \infty$ can be written uniquely as

$$\xi \;=\; \sum_{k=1} x_k \epsilon_k\,, \quad \text{with } \|\xi\|_p \;=\; \left(\sum_{k=1} |x_k|^p\right)^{1/p} < \infty \tag{2.77}$$

In general we have

**Definition 2.35.** A countable subset $B = \{\epsilon_k\}$ of a normed vector space $V$ is called a *Schauder Basis* of $V$ if each vector $\xi \in V$ can be written uniquely as $\xi = \sum_{k=1}^{\infty} x_k \epsilon_k$.

Without difficulty one shows that

**Lemma 2.36.** *A normed vector space which has a Schauder basis is separable.*

It is a remarkable and deep fact that the converse does not hold: there exists Banach spaces which are separable but do not have a Schauder basis. (See P. Enflo. A counterexample to the approximation problem in Banach spaces, Acta Math **130**, 309, (1973).)

## 2.5 Hahn-Banach theorem

As we will see a very important question in functional analysis is how to extend a functional defined on a subspace $W$ of a vector space $V$ to all $V$ while respecting some properties.

**Theorem 2.37. Hahn-Banach (real vector spaces)** *Let $V$ a vector space over $\mathbb{R}$ and $p : V \to \mathbb{R}$ a convex function on $V$, i.e. we have*

$$p(a\xi + (1-a)\eta) \leq ap(\xi) + (1-a)p(\eta) \tag{2.78}$$

*for any $a \in [0,1]$ and all $\xi, \eta \in V$.*

*Let $\phi$ be a linear functional defined on the subspace $W \subset V$ such that we have*

$$\phi(\xi) \;\leq\; p(\xi) \quad \text{for all } \xi \in W\,. \tag{2.79}$$

*Then there exists a linear functional $\Phi$ defined on $V$ such that*

$$\begin{aligned} \Phi(\xi) &= \phi(\xi) \quad \text{for all } \xi \in W\,. \\ \Phi(\xi) &\leq p(\xi) \quad \text{for all } \xi \in V\,. \end{aligned} \tag{2.80}$$

A very important example of a convex function $p$ is a norm (or a semi-norm) on $V$, i.e., $p(\xi) \;=\; \|\xi\|$.

*Proof.* The idea of the proof is to show that one can extend $\phi$ from $W$ to $W + \mathbb{R}\eta$ for any $\xi \notin W$. The rest follows from an applications of Zorn's lemma.

Let us choose $\eta \notin W$ and let denote $W_1 = W + \mathbb{R}\eta$. By linearity it is enough to specify $c := \phi(\eta)$ to define an extension $\phi_1$ on $W_1$ since we have then

$$\phi_1(\xi + a\eta) \;=\; \phi_1(\xi) + a\phi_1(\eta) = \phi(\xi) + ac\,. \tag{2.81}$$

To obtain the desired bound on $\phi_1$ we need that for all $\xi \in W$ and $a \in \mathbb{R}$ we must have

$$\phi_1(\xi + a\eta) \;=\; \phi(\xi) + ac \;\leq\; p(\xi + a\eta)\,. \tag{2.82}$$

We restrict ourselves to $a > 0$. Then we looking for a $c$ such that

$$\begin{aligned} \phi(\xi) + ac &\leq p(\xi + a\eta) \\ \phi(\xi) - ac &\leq p(\xi - a\eta) \end{aligned}$$

or

$$\begin{aligned} c &\leq \frac{1}{a}\left(p(\xi + a\eta) - \phi(\xi)\right) \\ c &\geq \frac{1}{a}\left(-p(\xi - a\eta) + \phi(\xi)\right) \end{aligned} \tag{2.83}$$

Such a $c$ exists provided we can prove that

$$\sup_{\xi \in W\,,\, a>0} \frac{1}{a}\left(-p(\xi - a\eta) + \phi(\xi)\right) \;\leq\; \inf_{\xi \in W\,,\, a>0} \frac{1}{a}\left(p(\xi + a\eta) - \phi(\xi)\right) \tag{2.84}$$

that is

$$\frac{1}{a_1}\left(-p(\xi_1 - a_1\eta) + \phi(\xi_1)\right) \;\leq\; \frac{1}{a_2}\left(p(\xi_2 + a_2\eta) - \phi(\xi_2)\right) \tag{2.85}$$

for all $a_1, a_2 \geq 0$ and $\xi_1, \xi_2 \in W$. Since $a_1$ and $a_2$ are non-negative this is equivalent to

$$a_2\phi(\xi_1) + a_1\phi(x_2) \leq a_1 p(\xi_2 - a_2\eta) + a_2 p(\xi_1 - a_1\eta)\,. \tag{2.86}$$

On the other hand we have use the assumption

$$\begin{aligned} a_2\phi(\xi_1) + a_1\phi(x_2) &= \phi(a_2\xi_1 + a_1\xi_2) \\ &= (a_1+a_2)\phi\left(\frac{a_2}{a_1+a_2}\xi_1 + \frac{a_1}{a_1+a_2}\xi_2\right) \\ &\leq (a_1+a_2)p\left(\frac{a_2}{a_1+a_2}\xi_1 + \frac{a_1}{a_1+a_2}\xi_2\right) \\ &= (a_1+a_2)p\left(\frac{a_2}{a_1+a_2}(\xi_1 - a_1\eta) + \frac{a_1}{a_1+a_2}(\xi_2 + a_2\eta)\right) \\ &\leq a_2 p(\xi_1 - a_1\eta) + a_1 p(\xi_2 + a_2\eta)\,. \end{aligned} \tag{2.87}$$

This is exactly (2.86) and thus we have proved the existence of $c := \phi_1(\eta)$.

We now use Zorn's lemma. Let $A$ the set of all extension $\psi$ won $\phi$ auf some subspace $W_\psi$ with the property that $\phi = \psi$ on $W$ and $\psi(\xi) \leq p(\xi)$ for all $\xi \in W_\psi$. In $A$ we can define a partial ordering through $\psi_1 \prec \psi_1$ if $W_{\psi_1} \subset W_{\psi_2}$ and $\psi_1 = \psi_2$ on $W_{\psi_1}$. Suppose $B \subset A$ is totally ordered and define then $\Psi$ on $\cup_{\psi \in B} W_\psi$ through $\Psi(\xi) = \psi(\xi)$ on $W_\psi$. By construction $\psi \prec \Psi$ for all $\psi \in B$ and so $B$ has an upper bound. By Zorn's lemma $A$ has a maximal element $\Phi$ which satisfies $\Phi = \phi$ on $W$ and $\Phi(\xi) \leq p(\xi)$. Finally $W_\Phi$ must be $V$ since otherwise one could extend $\Phi$ as before and this contradicts maximality. ■

This theorem as an extension to complex vector spaces.

**Theorem 2.38. Hahn-Banach (real vector spaces)** *Let $V$ a vector space over $\mathbb{C}$ and $p : V \to \mathbb{R}$ such that we have $p(a\xi + b\eta) \leq |a|p(\xi) + |b|p(\eta)$ for any $a, b \in \mathbb{C}$ with $|a| + |b| = 1$ and all $\xi, \eta \in V$.*

*Let $\phi$ be a complex linear functional defined on the subspace $W \subset V$ such that we have*

$$|\phi(\xi)| \leq p(\xi) \quad \text{for all } \xi \in W\,. \tag{2.88}$$

*Then there exists a linear functional $\Phi$ defined on $V$ such that*

$$\begin{aligned} \Phi(\xi) &= \phi(\xi) \quad \text{for all } \xi \in W\,. \\ |\Phi(\xi)| &\leq p(\xi) \quad \text{for all } \xi \in V\,. \end{aligned} \tag{2.89}$$

*Proof.* The functional $\phi_r(\xi) = \mathrm{Re}\phi(\xi)$ is a real functional on $V$ viewed as a vector field over $\mathbb{R}$. In addition

$$\phi_r(i\xi) = \mathrm{Re}\phi(i\xi) = \mathrm{Re}i\phi(\xi) = -\mathrm{Im}\phi(\xi)\,. \tag{2.90}$$

and so

$$\phi(\xi) = \phi_r(\xi) - i\phi_r(i\xi)\,. \tag{2.91}$$

Conversely given any real linear functional $\Phi_r$ on $V$ let us define $\Phi(\xi) = \Phi_r(\xi) - i\Phi_r(i\xi)$. It is certainly linear over $\mathbb{R}$ and we have

$$\Phi(i\xi) = \Phi_r(i\xi) - i\Phi_r(-\xi) = \Phi_r(i\xi) + i\Phi_r(\xi) = i\Phi(\xi) \tag{2.92}$$

and so $\Phi$ is linear over $\mathbb{C}$.

By the real Hahn-Banach theorem we can extend $\phi_r$ to $\Phi_r$ on $V$ such that $\Phi_r(\xi) \leq p(\xi)$ and $\Phi(\xi) = \Phi_r(\xi) - i\Phi_r(i\xi)$ is complex linear. With $\theta = arg\Phi(\xi)$ we find

$$|\Phi(\xi)| = e^{-i\theta}\Phi(\xi) = \Phi(e^{-i\theta}\xi) = \Phi_r(e^{-i\theta}) \leq p(e^{-i\theta}\xi) \leq |e^{-i\theta}|p(\xi) = p(\xi). \tag{2.93}$$

∎

It is hard to overemphasize how important this theorem is for the foundations of functional analysis. Let us first derive some corollaries and then discuss a number of applications.

**Corollary 2.39.** *Let $(V, \|\cdot\|)$ be a normed vector space, $W \subset V$ a subspace, and $\phi \in W'$. Then there exists $\lambda \in V'$ such that $\lambda = \phi$ on $W$ and $\|\lambda\| = \|\phi\|$.*

*Proof.* Take $p(\xi) = \|\phi\|\|\xi\|$ and apply Hahn-Banach. ∎

**Corollary 2.40.** *Let $(V, \|\cdot\|)$ be a normed vector space, $0 \neq \xi \in V$. Then there exists $\lambda \in V'$ such that $\lambda(\xi) = \|\xi\|$ and $\|\lambda\| = 1$.*

*Proof.* Let $W = \mathbb{K}\xi$ be the subpsace spanned by $\xi$ and set $\phi(a\xi) = a\|\xi\|$. This is a linear functional with $\|\phi\| = 1$. Now use Corollary 2.39. ∎

**Corollary 2.41.** *Let $(V, \|\cdot\|)$ be a normed vector space, $W \subset V$ a subspace. Let $\xi \in V$ be such that*

$$\inf_{\eta \in W} \|\xi - \eta\| = \delta > 0\,. \tag{2.94}$$

*Then there exists $\lambda \in V'$ such that $\|\lambda\| = 1$, $\lambda(\xi) = \delta$ and $\lambda(\eta) = 0$ for $\eta \in W$.*

*Proof.* Consider the subspace $W_1 = W \oplus \mathbb{K}\xi$ and let define $\lambda_1$ on $W_1$ by

$$\lambda_1(\eta + a\xi) = a\delta \tag{2.95}$$

Clearly we have $\lambda_1(\eta) = 0$ for $\eta \in W$ and $\lambda_1(\xi) = \delta$. The functional $\lambda_1$ is linear and bounded with norm 1: using the assumption (2.94) we have

$$\|\eta + a\xi\| = \| - a(-\frac{1}{a}\eta - \xi)\| = |a|\| - \frac{1}{a}\eta - \xi\| \geq |a|\delta = |\lambda_1(\eta + a\xi)\| \tag{2.96}$$

and so $\|\lambda_1\| \leq 1$. On the other hand for $\epsilon > 0$ there exists $\eta \in W$ such that

$$\delta \leq \|\xi - \eta\| \leq \delta(1 + \epsilon)\,. \tag{2.97}$$

Then we have

$$\lambda_1(\xi - \eta) = \delta \geq \frac{1}{1+\epsilon}\|\xi - \eta\|\,, \tag{2.98}$$

that is $\|\lambda_1\| \geq (1 + \epsilon)^{-1}$. So $\|\lambda_1\| = 1$. Now use Corollary 2.39 for any $\|\xi\|$ there exists

## 2.6 Applications of Hahn-Banach theorem

**Dual and bidual spaces** The Hahn-Banach space is useful to derive the properties of the dual space $V'$ as well as the bidual $V'' = (V')'$ of a Banach space. First we derive a dual representation of the norm

**Theorem 2.42.** *Let $V$ be a normed vector space. Then we have*

$$\|\xi\| = \sup_{\lambda \in V', \lambda \neq 0} \frac{|\lambda(\xi)|}{\|\lambda\|} \tag{2.99}$$

*In particular if $\xi_0$ is such that $\lambda(\xi_0) = 0$ for all $\lambda \in V'$ then $\xi_0 = 0$.*

*Proof.* On one hand from $|\lambda(\xi)| \leq \|\lambda\|\|\xi\|$ we have

$$\|\xi\| \geq \sup_{\lambda \in V', \lambda \neq 0} \frac{|\lambda(\xi)|}{\|\lambda\|} . \tag{2.100}$$

Using corollary 2.40 for a given fixed $\xi$ there exists $\lambda_\xi \in V'$ such that $\|\lambda\| = 1$ and

$$\|\xi\| = \lambda_\xi(\xi) = \frac{|\lambda_\xi(\xi)|}{\|\lambda_\xi\|} . \tag{2.101}$$

■

The next results describe the relation between a Banach space $V$ and his bidual $V''$.

**Theorem 2.43.** *Let $V$ be a normed vector space. For $\xi \in V$ define an element $\hat{\xi} \in V''$ by*

$$\hat{\xi}(\lambda) = \lambda(\xi) , \quad \lambda \in V' . \tag{2.102}$$

*Then the map*

$$\begin{aligned} J : \quad & V \to V'' \\ & \xi \mapsto \hat{\xi} \end{aligned} \tag{2.103}$$

*is an isometric isomorphism from $V$ to a subspace of $V''$.*

*Proof.* Since $|\hat{\xi}(\lambda)| = |\lambda(\xi)| \leq \|\lambda\|_{V'}\|\xi\|_V$ is $\hat{\xi}$ a bounded linear functional on $V'$ with

$$\|\hat{\xi}\|_{V''} \leq \|\xi\|_V . \tag{2.104}$$

This shows that $J(V) \subset V''$ and it remains to show that $\|\hat{\xi}\|_{V''} = \|\xi\|_V$. By corollary 2.40 given $\xi$ there exists $\lambda_\xi \in V'$ such that $\lambda_\xi(\xi) = \|\xi\|$. Then we have

$$\|\hat{\xi}\|_{V''} = \sup_{\lambda \in V', \|\lambda\|=1} |\hat{\xi}(\lambda)| \geq |\hat{\xi}(\lambda_\xi)| = \lambda_\xi(\xi) = \|\xi\|_V . \tag{2.105}$$

Therefore $J$ is an isometry of $V$ onto its range. ■

The following theorem suggest

**Definition 2.44.** A Banach space $V$ is called *reflexive* (or *self-dual*) if $J : V \to V''$ is bijective.

With respect to separability we have

**Theorem 2.45.** *A normed vector space $V$ is separable if $V'$ is separable.*

*Proof.* Consider a dense set $\{\lambda_k\}_{k\geq 1}$ in $V'$ and for each $k$ pick $\xi_k \in V$ with $\|\xi_k\| = 1$ and such that

$$\lambda_k(\xi_k) \geq \frac{1}{2}\|\lambda_k\| \,. \tag{2.106}$$

Now let $W$ be the countable set of finite linear combinations of the $\xi_k$ with rational coefficients.

By contradiction let us assume that $W$ is not dense and so there exists $\xi \in V$ such that

$$\inf_{\eta\in W} \|\xi - \eta\| = \delta > 0 \,. \tag{2.107}$$

By corollary 2.41 there exists $\lambda \in V'$ such that $\|\lambda\| = 1$, $\lambda(\xi) = \delta$ and $\lambda|_W = 0$. But since $\{\lambda_k\}$ is dense in $V'$ there exists a subsequence $\lambda_{k_i}$ such that $\lim_{i\to\infty} \|\lambda - \lambda_{k_i}\| = 0$. On the other hand we have

$$\|\lambda - \lambda_{k_i}\| \geq |(\lambda - \lambda_{k_i})(\xi_{k_i})| \;=\; |\lambda_{k_i}(\xi_{k_i}| \geq \frac{1}{2}\|\lambda_{k_i}\| \tag{2.108}$$

But this implies that $\lim_i \|\lambda_{k_i}\|$ and hence $\lambda = 0$. This is a contradiction since $0 \in W$. ∎

**Corollary 2.46.** *A separable Banach space $V$ with a non separable dual space $V'$ cannot be reflexive.*

*Proof.* If $V$ were reflexive then $V'' = V$ would be separable and hence $V'$ would be separable by Theorem 2.45. ∎

**Example 2.47.** Every Hilbert space is reflexive. From Riesz representation theorem (see Math 623-624) a Hilbert space $H$ is isometrically isomorphic to its dual $H'$ which is itself a Hilbert space. Hence $H$ is isometricaly isomorphic to its bi-dual $H''$.

**Example 2.48.** $l^p$ is reflexive and separable (see hwk). On the other hand $l^\infty$ is not separable (see HWK) from which it follows by corollary 2.46 that $l^1$ is not reflexive and hence $(l^1)'' = (l^\infty)' \neq l^1$.

**Example 2.49.** $C[a, b]$ is not reflexive since its dual space $(C[a, b])'$ is not separable. To see this note that for any $s \in [a, b]$ the functionals

$$\lambda_s(f) \;=\; f(s) \tag{2.109}$$

satisfy $\|\lambda_s\| = 1$ and $\lambda_s - \lambda_{s'}\| = 2$. Since there are uncountably such functionals separability is excluded.

We shall not prove the following important and classical result (see your measure theory class for details). Let us consider the Banach space $M([a, b])$ of all finite complex Borel measures on $[a, b]$ with total variation norm $\|\mu\|_{var} \;=\; \int_{[a,b]} d|\mu|$ as well as the Banach space $NBV[a, b]$ of function of bounded variation with $F(a) = 0$ with norm $\|F\|_{NBV} = V(F)$ where $V(F)$ is the variation of $F$.

**Theorem 2.50.** *Any bounded linear functional $\lambda$ on $C[a,b]$ can be written as*

$$\begin{aligned}\lambda(f) &= \int_{[a,b]} f d\mu \quad \mu \text{ Borel measure} \\ &= \int_{[a,b]} f dF \quad \text{Lebesgue} - \text{Stieljes integral}\end{aligned} \tag{2.110}$$

*with*

$$\|\lambda\| = \|\mu\|_{var} = Var(F)\,. \tag{2.111}$$

*Hence*

$$(C[a,b])' = NBV[a.b] = M[a,b]\,. \tag{2.112}$$

Finally we construct the dual or adjoint operator to a bounded operator using Hahn-Banach theorem.

**Definition 2.51.** Let $V$ and $W$ be normed vector spaces and $T : V \to W$ a bounded linear operator. Then the *adjoint operator* $T' : W' \to V'$ is defined by

$$(T'\lambda)(\xi) = \lambda(T\xi)\,, \quad \xi \in V, \lambda \in W'\,. \tag{2.113}$$

**Example 2.52.** Let $V = W = l^1$ and $T$ the shift operator defined for $\xi = (x_1, x_2, \cdots)$ by

$$T\xi = (0, x_1, x_2, x_3, \cdots) \tag{2.114}$$

We have $(l^1)' = l^\infty$ with $\eta(\xi) = \sum_{k=1}^\infty x_k y_k$. So

$$(T'\eta)(\xi) = \eta(T\xi) = \sum_{k=1}^\infty y_k x_{k-1} = \sum_{k=1}^\infty y_{k+1} x_k\,, \tag{2.115}$$

*hence we have*

$$T'\eta = (y_2, y_3, \cdots) \tag{2.116}$$

*It is easy to check that* $\|T\| = \|T'\| = 1$.

**Theorem 2.53.** *The adjoint operator $T'$ is linear bounded and we have*

$$\|T'\| = \|T\|\,. \tag{2.117}$$

*Proof.* The fact $T'$ is linear is easy and left to the reader. We have

$$\|T'\lambda(\xi)\| = \|\lambda(T\xi)\| \le \|\lambda\|\|T\xi\| \le \|\lambda\|\|T\|\|\xi\| \tag{2.118}$$

and this implies that $\|T'\lambda\| \le \|\lambda\|\|T\|$ and so $\|T'\| \le \|T\|$.

To prove that $\|T'\| \ge \|T\|$ we use that, by Corollary 2.40, for any $\xi \in V$ there exists $\lambda_\xi \in W'$ such that

$$\|\lambda_\xi\| = 1 \qquad \lambda_\xi(T\xi) = \|T\xi\| \tag{2.119}$$

We have then

$$\begin{aligned}\|T\xi\| &= \|\lambda_\xi(T\xi)\| \\ &= \|(T'\lambda_\xi)(\xi)\| \\ &\le \|T'\lambda_\xi\|\|\xi\| \\ &\le \|T'\|\|\lambda_\xi\|\|\xi\| \\ &= \|T'\|\|\xi\|\end{aligned} \tag{2.120}$$

and therefore $\|T\| \leq \|T'\|$. ■

**Remark 2.54.** It is straightforward to see that the map $T \mapsto T'$ satisfies

1. $(T+S)' = T' + S'$
2. $(\alpha T)' = \alpha T'$
3. $(ST)' = T'S'$

In the case where $V = H$ is a Hilbert space, then there is, by Riesz representation theorem, a natural antilinear isometry

$$\begin{aligned} J: \quad H' &\to H \\ \lambda &\mapsto \hat{\lambda} \end{aligned} \tag{2.121}$$

where

$$\lambda(\xi) \,=\, (\xi, \hat{\lambda})\,, \quad \xi \in H \tag{2.122}$$

(Recall $(\cdot,\cdot))$ is the scalar product in $H$ linear in the first argument, anti-linear in the second argument.)

**Definition 2.55.** *The (Hilbert space) adjoint operator $T^* : H \to H$ for $T \in \mathcal{L}(H)$ is defined by*

$$T^* \,=\, JT'J^{-1}\,. \tag{2.123}$$

*Then we have for any $\xi, \eta \in H$*

$$\underline{(T\xi, \eta)} \,=\, (R^{-1}\eta)(T\xi) \,=\, (T'R^{-1}\eta)(\xi) \,=\, (\xi, RT'R^{-1}\eta) \,=\, \underline{(\xi, T^*\eta)}\,. \tag{2.124}$$

Note that the map $T \mapsto T^*$ is antilinear, i.e. $cT^* = \bar{c}T^*$ for $c \in \mathbb{C}$. By slight abuse of notation $T^*$ is called the adjoint operator to $T$.

**Theorem 2.56.** *Let $H$ be a Hilbert space and $T, S \in \mathcal{L}(H)$. Then we have*

1. *$T \mapsto T^*$ is an anti linear isometry of $\mathcal{L}(H)$ onto itself.*
2. *$(T^*)^* = T$*
3. *$(TS)^* = S^*T^*$*
4. *If $T^{-1} \in \mathcal{L}(H)$ then $T^*{-1} \in \mathcal{L}(H)$ and $(T^*)^{-1} = (T^{-1})^*$.*
5. *$\|T^*T\| = \|T\|^2$*

*Proof.* 1. follows from Riesz representation theorem. 2. follows from the reflexivity of $H$ and $H = H'$. 3. and 4. are left to the reader. For 5. note that $\|T^*T\| \leq \|T^*\|\|T\| = \|T\|^2$. Conversely using theorem 2.42 we have

$$\|T^*T\|\|\xi\| \geq \|T^*T\xi\| = \sup_{\eta:\eta\neq 0} \frac{|(T^*T\xi,\eta)|}{\|\eta\|} \geq \frac{|(T^*T\xi,\xi)|}{\|\xi\|} = \frac{|(T\xi,T\xi)|}{\|\xi\|} = \frac{\|T\xi\|^2}{\|\xi\|} \tag{2.125}$$

and hence $\|T\|^2 \leq \|T^*T\|$. ■

## 2.7 Baire theorem and uniform boundedness theorem

The uniform boundedness theorem is quite useful in applications and as the open mapping theorem and closed graph theorem of next section it derives from a common source: the so-called Baire category theorem. Here, as opposed to the Hahn-Banach theorem the *completeness* of the space plays a crucial role.

**Definition 2.57.** A subset $M$ of a metric space $X$ is said to be

1. *nowhere dense in* $X$ if its closure $\overline{M}$ has no interior points.
2. *of the first category in* $X$ if $M$ is the union of countable many sets each of which is nowhere dense.
3. *of the second category in* $X$ if $M$ is not of the first category.

**Theorem 2.58. (Baire category theorem)** *A complete non-empty metric space* $X$ *is of the second category in itself. In particular if* $X \neq \emptyset$ *is complete and*

$$X = \bigcup_{k=1}^{\infty} A_k \,, \quad A_k \text{ closed} \tag{2.126}$$

*then at least one* $A_k$ *contains a nonempty open subset.*

*Proof.* Let us assume that $X$ is of the first category and so

$$X = \bigcup_{k=1}^{\infty} M_k \,, \tag{2.127}$$

with each $M_k$ nowhere dense. We will construct a Cauchy sequence $\xi_k$ whose limit belongs to no $M_k$, therefore contradicting the representation 2.127.

By assumption $M_1$ is nowhere dense so that $\overline{M_1}$ does not contain a nonempty open set. But $X$ does (e.g. $X$ itself). This implies that $M_1 \neq X$ and thus $\overline{M_1}^c = X \setminus \overline{M_1}$ is open and non-empty. So we pick $\xi_1 \in \overline{M_1}^c$ and an open ball around it

$$B_1 = B_{\epsilon_1}(\xi_1) \subset \overline{M_1}^c \qquad \epsilon_1 < \frac{1}{2} \tag{2.128}$$

By assumption $M_2$ is nowhere dense in $X$ so that $\overline{M_2}$ does not contain a nonempty open set. Hence it does not contain the open ball $B_{\epsilon_1/2}(\xi_1)$. Therefore $\overline{M_2}^c \cap B_{\epsilon_1/2}(\xi_1)$ is nonempty and open so that we may choose an open ball in this set , say,

$$B_2 = B_{\epsilon_2}(\xi_2) \subset \overline{M_2}^c \cap B_{\epsilon_1/2}(\xi_1) \qquad \epsilon_2 < \frac{\epsilon_1}{2} < \frac{1}{4} \tag{2.129}$$

By induction we find a sequence of balls

$$B_k = B_{\epsilon_k}(\xi_k) \qquad \epsilon_k < \frac{1}{2^k} \tag{2.130}$$

such that $B_k \cap M_k = \emptyset$ and

$$B_{k+1} \subset B_{\epsilon_k/2}(\xi_k) \subset B_k \,. \tag{2.131}$$

Since $\epsilon_k < 2^k$ the sequence $\xi_k$ is Cauchy and $\lim_{k\to\infty} \xi_k = \xi \in X$. Furthermore for every $n$ and $m \geq n$ we have $B_m \subset B_{\epsilon_n/2}(\xi_n)$ so that

$$d(\xi_n, \xi) \leq d(\xi_n, \xi_m) + d(\xi_m, \xi) < \frac{\epsilon_n}{2} + d(\xi_m, \xi) \longrightarrow \frac{\epsilon_n}{2}\,. \tag{2.132}$$

as $m \to \infty$. Hence for every $n$, $\xi \in B_N \subset \overline{M}_n{}^c$ and so $\xi \notin \cup_k M_k = X$. This is a contradiction. ■

An important immediate consequence of Baire's theorem is the uniform boundedness theorem (Banach-Steinhaus theorem). It is quite remarkable since it shows that if a sequence $T_n$ is point wise bounded, it is is bounded in norm!.

**Theorem 2.59. (Uniform boundedness theorem)** *Let $\{T_n\}$ be a sequence of bounded linear operators $T_n \in \mathcal{L}(V, W)$ from a Banach space $V$ into a normed vector space $W$. Assume that for ever $\xi \in V$ there exists a constant $c_\xi$ such that*

$$\sup_n \|T_n \xi\| \leq c_\xi\,. \tag{2.133}$$

*Then there exists a $c$ such that*

$$\sup_n \|T_n\| \leq c \tag{2.134}$$

*Proof.* Let us define

$$A_k = \{\xi \,:\, \sup_n \|T_n \xi\| \leq k\}\,, \tag{2.135}$$

and it is easy to see that $A_k$ is a closed set and we have $V = \bigcup_k A_k$. Since $V$ is complete, by Baire theorem some $A_k$ contains an open ball, say,

$$B_0 = B_r(\xi_0) \subset A_{k_0}\,. \tag{2.136}$$

For an arbitrary $\xi \neq 0$ let $\eta \in B_0$ be given by

$$\eta = \xi_0 + \frac{r}{2}\frac{\xi}{\|\xi\|}\,. \tag{2.137}$$

and since $\eta \in A_{k_0}$ we have $\sup_n \|T_n \eta\| \leq k_0$. We have

$$\xi = \frac{2\|\xi\|}{r}(\eta - \xi_0) \tag{2.138}$$

and for any $n$

$$\|T_n \xi\| = \frac{2\|\xi\|}{r}\|T_n(\eta - \xi_0)\| \leq, \frac{2\|\xi\|}{r}\|T_n \eta\| + \|T_n \xi_0)\| \leq \frac{4k_0}{r}\|\xi\| \tag{2.139}$$

and hence

$$\sup_n \|T_n\| \leq \frac{4k_0}{r}\,. \tag{2.140}$$

This concludes the proof. ■

**Example 2.60. Bilinear functionals** Let $V$ and $W$ be Banach spaces. Consider a bilinear functional

$$B : V \times W \to \mathbb{C} \tag{2.141}$$

which is continuous in each variable. That is for fixed $\xi \in V$

$$B(\xi, \cdot) : W \to \mathbb{C} \tag{2.142}$$

is linear and continuous and for fixed $\eta in W$

$$B(\cdot, \eta) : V \to \mathbb{C} \tag{2.143}$$

is linear and continuous. We claim that continuity in each variable implies continuity, i.e., if $(\xi_n, \eta_n) \to (0, 0)$ then $B(\xi_n, \eta_n) \to 0$.

*Proof.* Define $T_n : W \to \mathbb{C}$ by

$$T_n\eta \;=\; B(\xi_n, \eta)\,. \tag{2.144}$$

For any $n$, by the continuity in the second variable we have that $T_n$ is a bounded operator. For any $\eta in W$ by continuity in the first variable we have $\lim_n \|T_n\eta\| \;=\; 0$ and hence $\sup_n \|T_n\eta\| \le c_\eta < \infty$. By the uniform boundedness theorem there exists a $c > 0$ such that

$$\sup_n \|T_n\| \le c \tag{2.145}$$

or

$$|B(\xi_n, \eta_n)| = \|T_n\eta_n\| \le c\|\eta_n\| \longrightarrow 0 \tag{2.146}$$

as $n \to \infty$. Hence we have continuity. ∎

Note further that continuity of $B$ is equivalent to

$$|B(\xi, \eta)| \le c\|\xi\|\|\eta\| \tag{2.147}$$

for all $\xi, \eta$. And this proved exactly as for linear maps.

We apply this to symmetric operators in Hilbert spaces and we show that unbounded symmetric operators are necessarily defined only on a subspace of $H$ but not on all of $H$.

**Theorem 2.61. (Hellinger-Toeplitz)** *Let $H$ be a Hilbert space and let $T : H \to H$ be a linear operator defined on all $H$ and we have $(T\xi, \eta) \;=\; (\xi, T\eta)$ for all $\xi, \eta \in H$. Then $T$ is bounded.*

**Corollary 2.62.** *Let $H$ be a Hilbert space and let $T : D_T \to H$ be a linear operator defined on all $D_T \subset H$ and we have $(T\xi, \eta) \;=\; (\xi, T\eta)$ for all $\xi, \eta \in D_T$. Then $D_T \ne H$*

*Proof.* The bilinear map

$$B(\xi, \eta) \;=\; (T\xi, \eta) \tag{2.148}$$

is continuous in $\eta$ and since

$$B(\xi, \eta) \;=\; (\xi, T\eta) \tag{2.149}$$

it is also continuous in $\xi$. By the previous example $B$ is continuous in both variables jointly and thus there exists $c > 0$ such that

$$|(T\xi, \eta)| \;\le\; c\|\xi\|\|\eta\| \tag{2.150}$$

for all $\xi, \eta \in H$. In particular if $\eta = T\xi$ we have

$$\|T\xi\|^2 \,=\, (T\xi, T\xi) \,=\, (T\xi, \eta) \,\le\, c\|\xi\|\eta\| = c\|\xi\|\|T\xi\|\,. \tag{2.151}$$

Hence we have $\|T\| \le c$. ■

**Example 2.63. (Convergence of Fourier series)** Consider a function $f$ periodic of period $2\pi$. Its Fourier coefficients are

$$c_n \,=\, \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-int}\,dt\,, \quad n \in \mathbb{Z} \tag{2.152}$$

and the Fourier partial sums are

$$S_N(f)(t) \,=\, \sum_{|k|\le N} c_n e^{inx} \tag{2.153}$$

As one learn in an analysis class we have pointwise (or even uniform) convergence of $S_N(f)(x) to f(x)$ if the function $f$ is sufficiently smooth (say $f$ is $C^1$). At discontinuity points $f$ may or may not converge but interestingly enough even at points where $f$ is continuous $S_N(f)$ need not converge. One can construct explicit examples but we prove this here using the uniform boundedness theorem. First we recall that, using trigonometric formulas one can write

$$S_N(f)(t) \,=\, \frac{1}{2\pi}\int_0^{2\pi} D_N(t-s)f(s)\,ds \quad \text{with } D_N(t) \,=\, \frac{\sin((n+\frac{1}{2})t)}{\sin(\frac{1}{2}t)} \tag{2.154}$$

We apply the uniform boundedness theorem by consider the Banach space $X$ of continuous periodic of period $2\pi$ with $\|f\| \,=\, \sup_t |f(t)|$ and let us define the linear functional

$$\lambda_n(f) \,=\, S_n(f)(0) \tag{2.155}$$

One checks that

$$|\lambda_n(f)| \,\le\, \|f\|\frac{1}{2\pi}\int_0^{2\pi} |D_N(s)|\,ds \tag{2.156}$$

By using the same argument as for computing the norm of the Fredholm integral operator we obtain

$$\|\lambda_n\| \,=\, \frac{1}{2\pi}\int_0^{2\pi} |D_n(s)|\,ds \tag{2.157}$$

Now we use the inequality $|\sin(t)| < t$ on $[0, \pi]$ and obtain

$$\begin{aligned}
\|\lambda_n\| &= \frac{1}{2\pi}\int_0^{2\pi} \left|\frac{\sin((n+\frac{1}{2})t)}{\sin(\frac{1}{2}t)}\right| dt \\
&> \frac{1}{\pi}\int_0^{2\pi} \frac{\left|\sin((n+\frac{1}{2})t)\right|}{t} dt \\
&= \frac{1}{\pi}\int_0^{(2n+1)\pi} \frac{|\sin v|}{v} dt \\
&= \frac{1}{\pi}\sum_{k=0}^{2n}\int_{k\pi}^{(k+1)\pi} \frac{|\sin v|}{v} dt \\
&\geq \frac{1}{\pi}\sum_{k=0}^{2n}\frac{1}{(k+1)\pi}\int_{k\pi}^{(k+1)\pi} |\sin v|\ dt \\
&= \frac{2}{\pi^2}\sum_{k=0}^{2n}\frac{1}{(k+1)} \longrightarrow 0 \qquad (2.158)
\end{aligned}$$

as $n \to \infty$. Hence the sequence $\|\lambda_n\|$ is unbounded. By the uniform boundedness theorem this implies that $\sup_f |\lambda_n(f)|$ cannot be bounded and hence there exists at least one $f$ such that $\lambda_n(f) = S_n(f)(0)$ diverges. That is the Fourier series diverges at $t = 0$. ■

## 2.8 Open mapping and closed graph theorems

After the Hahn-Banach theorem and the uniform boundedness theorem we now attack the third ”big” theorem of functional analysis, the open mapping theorem. It is well-known that the continuity of a map $f : X \to Y$ between metric spaces is equivalent to the property that for any open set $O \subset Y$, the set $f^{-1}(0)$ is also open. By contrast let us define

**Definition 2.64.** Let $M$ and $N$ be metric spaces. The $F : D_F \to N$ with domain $D_F$ is called an *open mapping* if for every open set $O \in D_F$ the image $F(O)$ is an open set in $Y$.

**Remark 2.65.** In general continuous mapping are not open, e.g. $f(t) = \sin(t)$ maps the open set $(0, 2\pi)$ onto the closed set $[-1, 1]$.

We have the remarkable result

**Theorem 2.66. (Open mapping theorem)** *Let $V$ and $W$ be Banach spaces. A surjective bounded linear operator $T : V \to W$ is an open mapping.*

An immediate consequence of this theorem is

**Theorem 2.67. Inverse mapping theorem** *LetLet $V$ and $W$ be Banach spaces and let $T : V \to W$ be a bounded linear bijective map. Then $T^{-1}$ is continuous and thus bounded.*

The proof of theorem 2.66 relies on

**Proposition 2.68.** *A bounded linear operator $T : V \to W$ from a Banach space $V$ onto a Banch space $W$ has the property that the image of $T(B_0)$ of the open unit ball $B_0 = B_1(0) \subset V$ contains an open ball around $0 \in W$*

*Proof.* The proof consists of three steps:

1. The closure of the open ball $B_1 = B_{1/2}(0)$ contains an open ball $B^*$.
2. If $B_n = B_{2^{-n}}(0)$ then the closure $\overline{T(B_n)}$ contains an open ball $V_n$ around $0 \in Y$.
3. $T(B_0)$ contains an open ball around $0$.

Step 1.: Let $B_1 = B_{1/2}(0)$. Then we have

$$V \;=\; \bigcup_{k=1}^{\infty} kB_1 \,. \tag{2.159}$$

Since $T$ is surjective and linear,

$$W \;=\; T(V) \;=\; T\big(\bigcup_{k=1}^{\infty} kB_1\big) \;=\; \bigcup_{k=1}^{\infty} kT(B_1) \;=\; \bigcup_{k=1}^{\infty} \overline{kT(B_1)}\,, \tag{2.160}$$

where the last equality follows from the fact that the union is $Y$, hence we did not add any points by taking the closure. By Baire category theorem there exist some $k$ such that $\overline{kT(B_1)}$ contains an open ball and hence $\overline{T(B_1)}$ contains an open ball, say $B^* = B_\epsilon(\eta_0) \subset \overline{T(B_1)}$. Then we also have

$$B^* - \eta_0 = B_\epsilon(0) \subset \overline{T(B_1)} - \eta_0 \tag{2.161}$$

Step 2.: We prove that $B^* - \eta_0 \subset \overline{T(B_0)}$ where $B_0 = B_1(0)$ by proving that (see (2.161)) that

$$\overline{T(B_1)} - \eta_0 \subset \overline{T(B_0)}\,. \tag{2.162}$$

Let $\eta \in \overline{T(B_1)} - \eta_0$. Then $\eta + \eta_0 \in \overline{T(B_1)}$ and we remember that also $\eta_0 \in \overline{T(B_1)}$. Then there exists $\alpha_n = T\beta_n \in T(B_1)$ and $\delta_n = T\gamma_n \in T(B_1)$ such that

$$\lim_n \alpha_n = \eta + \eta_0, , \quad \lim_n \delta_n \;=\; \eta_0\,. \tag{2.163}$$

Since $\beta_n, \gamma_n \in B_1$ of radius 1/2 we have $\|\beta_n - \gamma_n\| < 1$ and $\beta_n - \gamma_n \in B_0$. From

$$\lim_{n\to\infty} T(\beta_n - \gamma_n) \;=\; \eta \tag{2.164}$$

we conclude that $\eta \in \overline{T(B_0)}$. This concludes the proof of (2.162) and thus

$$B^* - \eta_0 = B_\epsilon(0) \subset \overline{T(B_0)} \tag{2.165}$$

By the linearity of $T$ if $B_n = B_{2^{-n}}(0)$, we have $\overline{T(B_n)} = 2^{-n}\overline{T(B_0)}$ and thus

$$V_n \;=\; B_{\epsilon/2^n}(0) \subset \overline{T(B_n)} \tag{2.166}$$

Step 3: Finally we prove that

$$V_1 \,=\, B_{\epsilon/2}(0) \subset T(B_0)\,. \tag{2.167}$$

Let $\eta \in V_1$. By (2.166) with $n = 1$, for any $\epsilon > 0$ there exist $\xi_1 \in B_1$ such that

$$\|\eta - T\xi_1\| \leq \epsilon/4\,. \tag{2.168}$$

Then $\eta - T\xi_1 \in V_2$ and by (2.166) with $n = 2$ we see that $\eta - T\xi_1 \in V_2 \subset \overline{T(B_2)}$. Repeating the same argument we find $\xi_2 \in B_2$ such that

$$\|\eta - T\xi_1 - T\xi_2\| \leq \epsilon/8 \tag{2.169}$$

and hence $\eta - T\xi_1 - T\xi_2 \in V_3 \subset \overline{T(B_3)}$, an so on. In the $n$th step we select $\xi_n in B_n$ such that

$$\|\eta - \sum_{k=1}^{n} T\xi_k\| \leq \frac{\epsilon}{2^{n+1}} \tag{2.170}$$

Let us set $\zeta_n = \xi_1 + \cdots \xi_n$ then since $\|\xi_k\| \leq 2^{-k}$ $\zeta_n$ is Cauchy sequence and $\zeta_n \to \xi \in V$. Also $\xi \in B_0$ since $B_0$ has radius 1. Since $T$ is continuous $T\zeta_n \to Tx$ and by (2.170) we have $T\xi = \eta$ and so $\eta \in T(B_0)$. ■

*Proof of theorem 2.66.* This follows from the previous proposition by using linearity to translate and dilate balls. ■

The next theorem is also an immediate consequence of the open mapping theorem. Suppose $T : D(T) \to W$ where $D(T)$ is a subspace of $V$ and is called the domain of $T$, and $V, W$ are Banach spaces. We do not assume that $T$ is bounded and in general $D(T) \neq V$.

**Definition 2.69.** The *graph* of $T$ is the set

$$\Gamma(T) \,=\, \{[\xi, \eta] \in V \times W\,;\, \xi \in D_T, \eta = T\xi\} \tag{2.171}$$

Note that $\Gamma(T)$ is a subspace of $V \times W$ which we can make it into a normed vector space with the norm

$$\|[\xi, \eta]\| \,:=\, \|\xi\| + \|\eta\| \tag{2.172}$$

**Definition 2.70.** A map $T : D(T) \to W$ ($V$,$W$ Banach spaces) is *closed* if the graph $\Gamma(T)$ is closed. Equivalently $T$ is closed if for any sequence $\{\xi_n\}$ such that

$$\xi_n \longrightarrow \xi \quad \text{ and } T\xi_n \to eta \tag{2.173}$$

then

$$Tx \,=\, \eta \tag{2.174}$$

Clearly bounded operators are closed, but the converse is not true.

**Theorem 2.71.** **(Closed graph theorem)** *Let $V$ and $W$ be Banach spaces and $T : D(T) \to W$ a linear operator. If $D(T)$ is closed and if $T$ is closed then $T$ is bounded.*

**Corollary 2.72.** *Let $V$ and $W$ be Banach spaces and $T : V \to W$ a linear operator. Then $T$ is closed if and only if $T$ is bounded.*

*Proof.* If $D(T)$ and $\Gamma(T)$ are closed then they are Banach spaces and we define the map $P\ \Gamma(T) \to D(T)$ by

$$P([\xi, T\xi]) = \xi \tag{2.175}$$

i.e., $P$ is the projection on the first component. The map $P$ is a bijection and by the inverse mapping theorem $P^{-1}$ is bounded. This means

$$\|P^{-1}\xi\| \leq c\|\xi\| \text{ or } \|\xi\| + \|T\xi\| \leq C\|\xi\| \tag{2.176}$$

. Thus we have $\|T\xi\| \leq (C-1)\|xi\|$ and so $T$ is bounded. ∎

As an application we prove

**Theorem 2.73.** *Suppose than $V$ is a Banach space with respect to the two norms $\|\xi\|_1$ and $\|xi\|_2$ which are compatible in the sense that if a sequence $\{\xi_n\}$ converges in both norms then the two limits are equal. Then the two norms are equivalent in the sense that there exists constants cand $C$ such that $c\|\xi\|_1 \leq \|\xi\|_2 \leq C\|\xi\|_1$.*

*Proof.* Consider the identity map $\mathbf{1} : (V, \|\cdot\|_1) \to (V, \|\cdot\|_2)$ given by $\mathbf{1}(\xi) = \xi$. Compatibility means that the map $\mathbf{1}$ is closed. By the closed graph theorem it is bounded in both directions. ∎

We conclude with an example of a closed operator which is not bounded

**Example 2.74.** Let $V = C[0,1]$ with the sup-norm and let $Tf(t) = f'(t)$ be the differentiation operator with domain $D(T) = C^1[0,1]$ of continuously differentiable functions. The operator is unbounded (take $f_n = t^n$) and closed since if $f_n$ converges uniformly to $f$ and $f'_n$ converge uniformly to $h$ then we have using uniform convergence to interchange integral and limit

$$\int_0^t h(s)\,ds = \int_0^t \lim_n f'_n(s)\,ds = \int_0^t \lim_n f'_n(s)\,ds = f(t) - f(0)\,, \tag{2.177}$$

that is $f'(t) = h$. ∎

## 2.9 Exercises

**Exercise 4.** Prove that $l^p$ is a Banach space.

**Exercise 5.** Consider the normed vector space $BV[a,b]$ with $\|f\|_{BV} = f(a) + V(f)$ where $V(f)$ is the variation of $f$ on $[a,b]$ and let $\|f\|_\infty = \sup_t |f(t)|$.

1. Show that the norm $\|f\|_\infty$ is weaker than $\|f\|_{BV}$.
2. Show that $BV[a,b]$ with $\|\cdot\|_{BV}$ is a Banach space. (You may use part 1.)

**Exercise 6.** Prove Theorem 2.15.

**Exercise 7.** Compute the spectral radius of the Volterra integral $K$ operator given in Eq. (1.35) of Exercise 2

**Exercise 8.** Let $m$ denote Lebesgue measure. Show that the integral operator $Tf = \int_a^b k(t,s)f(s)\,ds$ defined a bounded operator on $L^2([a,b],m)$ provided $k \in L^2([a,b]\times[a,b], m \times m)$.

**Exercise 9.** Let $V$ be a Banach space.

1. Show that if $T \in \mathcal{L}(V)$ then $e^X$ defined by

$$e^X = \sum_{k=1}^{\infty} \frac{1}{k!} X^k\,, \tag{2.178}$$

   defines a linear operator in $\mathcal{L}(V)$.

2. Show that $e^{X+Y} = e^X e^Y$ whenever $X$ and $Y$ commute, i.e. $XY = YX$.

3. Show that $e^X$ is invertible and $(e^X)^{-1} = e^{-X}$.

4. Show that the (non-linear) map $X \mapsto e^X$ is differentiable and compute the derivative $(e^X)'$. Show that

$$(e^X)' \neq e^X\,. \tag{2.179}$$

**Exercise 10.** Show that $l^p$ $1 \leq p < \infty$ is separable but that $l^\infty$ is not separable.

**Exercise 11.** To show that $(l^\infty)' \neq l^1$ consider the subspace $c$ and define a function $\lambda$ on $c$ by

$$\lambda(\xi) = \lim_{n\to\infty} \xi_n\,. \tag{2.180}$$

Show that $\lambda$ extends to functional on $l^\infty$ and deduce from this that $(l^\infty)' \neq l^1$.

**Exercise 12.** Let $V$ and $W$ be normed vector spaces and $T \in \mathcal{L}(V,W)$. If $T^{-1}$ exists and is bounded show that $(T^{-1})' = (T')^{-1}$.

**Exercise 13.** To illustrate the Hahn-Banach theorem and its consequences:

1. For Corollary 2.39, consider the functional $\lambda$ on the euclidean plane $\mathbb{R}^2$ given by $\lambda(\xi) = \alpha_1 x_1 + \alpha_2 x_2$, its linear extensions $\tilde{\lambda}$ to $\mathbb{R}^3$ and the corresponding norms $\tilde{\lambda}$.

2. For Corollary 2.40, let $V = \mathbb{R}^2$, find the functional $\lambda$.

**Exercise 14.** Let $V$, $W$ be normed vector spaces and $\{T_n\}$ a sequence of bounded operators.

1. Suppose $V$ is a Banach space. Show that if $\{T_n\}$ converges *strongly* to $T$ then $T$ is a bounded operator.

2. If $V$ is not complete then $T$ need not to be bounded. To see this let $E \subset l^\infty$ be the subspace of sequence which contains only a finite number of nonzero terms and define $A$ by

$$T(x_1, x_2, \cdots) = (x_1, 2x_2, 3x_3, \cdots) \tag{2.181}$$

   Show that $A$ is not bounded but can be written as the strong limit of a sequence bounded operators.

**Exercise 15.** Let $V$, $W$ be Banach spaces and $\{T_n\}$ a sequence of bounded operators. Show that the sequence $\{T_n\}$ converges strongly to a bounded operator $T$ if

1. The sequence $\{\|T_n\|\}$ is bounded.
2. The sequence $\{\|T_n\xi\|\}$ converges for $\xi$ in a dense set in $V$.

*Hint:* Use the first part of previous exercise.

**Exercise 16.**

1. Show by an example that in Baire's category theorem the completeness condition cannot be omitted.
2. Show by an example that in Baire's category theorem condition of the countability of the decomposition (see Eq. (2.126)) cannot be omitted.

*Hint:* Do not look for complicated metric spaces.

**Exercise 17. (Weak convergence)** Let $V$ be a normed vector space. We say that a sequence $\{\xi_n\}$ *converges weakly* to $\xi$ if

$$\lim_n \lambda(\xi_n) \;=\; \lambda(\xi) \tag{2.182}$$

for all $\lambda \in V'$. Show the following properties of weak convergence

1. The weak limit of $\{\xi_n\}$, if it exists, is unique.
2. If $\xi_n$ converges weakly to $\xi$ then $\|\xi_n\|$ is bounded. *Hint:* Use the uniform boundedness theorem.
3. If $\xi_n$ converges to $\xi$ then $\xi_n$ converges to $\xi$ weakly but the converse is not necessarily true. *Hint:* Try a separable Hilbert space or $l^p$....
4. Show that in finite dimensional spaces weak convergence and strong convergence are equivalent.
5. Show that if $\{\xi_n\}$ is a sequence such that (i) $\sup_n \|\xi_n\| \leq c < \infty$ and (ii) $\lim_n \lambda(\xi_n) = \lambda(\xi)$ for a dense set of $\lambda$ in $V'$ then $\xi_n$ converges weakly to $\xi$.
6. Show that if $V = l^p$, $p > 1$, then $\xi_n$ converges weakly to $\xi$ if and only if the sequence $\{\|\xi_n\|\}$ is bounded and $\lim_n x_k^{(n)} = x_k$ for all $k$. (Here we have denoted $\xi_n = (x_1^{n)}, x_2^{(n)}, \cdots ...)$)
7. Show that in $l^1$ weak convergence and strong convergence are equivalent.

**Exercise 18.** In this problem we call a map *an open mapping* if for any open set $O \subset V$ the image $T(O)$ is an open set of $T(V)$.

1. Suppose $N$ is a closed subspace of a Banach space $V$ and consider the quotient space quotient $V/N$. Show that $V/N$ can be made into a Banach space with the norm
$$\|\widehat{\xi}\| \;=\; \inf_{\eta\in\widehat{\xi}} \|\eta\|\,. \tag{2.183}$$
2. Show that if $V$ is a Banach space then any $\lambda \in V'$ is an open mapping. Show also that if $T : V \to W$ has finite-dimensional range then $T$ is an open mapping. *Hint: Use part 1.*
3. Find an example of a bounded linear map which is not an open mapping.

**Exercise 19.** Let $V$ and $W$ be Banach spaces and $T : V \to W$ a bounded linear maps such that the range of $T$, $R(T)$ is finite-codimensional subspace of $W$. Show that $R(T)$ is closed.
*Hint:* Use the closed graph theorem. Extend $T$ to $V \oplus Z$ such that the range of $T$ is all of $W$.

**Exercise 20.** Suppose $V$ is a Banach space, $Y$ and $Z$ closed complementary subspaces of $V$ such that $V = Y \oplus Z$. Let $P_Y$ be the projection on $Y$ along $Z$ and $P_Z$ be the projection on $Z$ along $Y$. Show that $P_Y$ and $P_Z$ are continuous.
*Hint:* Use the closed graph theorem

# Chapter 3

# Spectral theory

## 3.1 Spectrum and resolvent

Let us first recall the spectral theory in finite-dimensional space (linear algebra). For $T \in \mathcal{L}(\mathbb{C}^n)$, $\lambda \in \mathbb{C}$ is called an *eigenvalue* of $T$ if $T - \mu\mathbf{1}$ is singular, i.e., if $\det(T - \mu\mathbf{1}) = 0$. The set of eigenvalues of $T$ is called the *spectrum* of $T$. Since $\det(T - \mu\mathbf{1})$ is a polynomial of order $n$ then the spectrum of $T$ contains at least one point and at most $n$ points. If $\mu$ is an eigenvalue, then the *eigenvalue equation* $T\xi = \mu\xi$ has at least one non-trivial solution. Such solutions are called *eigenvectors* for the eigenvalue $\lambda$. If $\mu$ is not an eigenvalue then $T - \mu\mathbf{1}$ is regular and so $(T - \mu\mathbf{1})^{-1}$ exists.

In infinite dimensional vector spaces is the spectral analysis hugely more complicated, but also much more interesting than in finite-dimensional spaces. From a practical point of view understanding the spectrum of an operator is essential part of understanding the operator itself!

**Convention/notation:** $V$ is a *complex* vector space and $T \in \mathcal{L}(V)$. For $\mu \in \mathbb{C}$ we set

$$T_\mu = T - \mu\mathbf{1} \tag{3.1}$$

**Definition 3.1.** Let $T \in \mathcal{L}(V)$

1. $\mu$ is a *regular value* of $T$ if $T_\mu$ is bijective. (Hence $T_\mu^{-1} \in \mathcal{L}(V)$ by the inverse mapping theorem).

2. The *resolvent set of T*, denoted by $\rho(T)$ is the set of regular values of $T$,

$$\rho(T) = \{\mu \in \mathbb{C} : \mu \text{ regular value of } T\}\,, \tag{3.2}$$

and the *resolvent* of $T$ is

$$R_\mu(T) = (T - \mu\mathbf{1})^{-1} \tag{3.3}$$

3. The *spectrum* of $T$ is

$$\sigma(T) = \mathbb{C} \setminus \rho(T)\,, \tag{3.4}$$

and $\lambda \in \sigma(T)$ is called a *spectral value*.

4. $\mu \in \sigma(T)$ is called an *eigenvalue* of $T$ if the equation

$$T\xi \,=\, \mu\xi \tag{3.5}$$

has a non-trivial solution $\xi \in V$. Such a solution $\xi$ is called an *eigenvector* for the eigenvalue $\mu$ and the subspace

$$E_\mu(T) \,=\, \{\xi\,;\, T\xi \,=\, \mu\xi\} \tag{3.6}$$

is called the *eigenspace* of $T$ for the eigenvalue $\mu$. The *point spectrum* of $T$ is

$$\sigma_p(T) \,=\, \{\mu\,;\, \mu \text{ eigenvalue of } T\} \tag{3.7}$$

5. The *continuous spectrum* of $T$ is

$$\sigma_c(T) \,=\, \{\mu \in \sigma(T) \setminus \sigma_p(T)\,;\, D_\mu = T_\mu(V) \text{ is dense in } V \text{ and } T_\mu^{-1} \text{ exists, but is unbounded}\} \tag{3.8}$$

6. The *residual spectrum* of $T$ is

$$\sigma_c(T) \,=\, \{\mu \in \sigma(T) \setminus \sigma_p(T)\,;\, D_\mu = T_\mu(V) \text{ is } \textbf{not} \text{ dense in } V\} \tag{3.9}$$

We clearly have

**Lemma 3.2.** *The sets $\sigma_p(T)$, $\sigma_c(T)$, and $\sigma_r(T)$ are mutually disjoint and*

$$\mathbb{C} \,=\, \rho(T) \cup \sigma(T) \,=\, \rho(T) \cup \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T) \tag{3.10}$$

It is not immediately obvious that $\sigma_r(T)$ is not empty.

**Example 3.3.** Let $T$ be the right shift operator on $l^2$, i.e for $\xi = (x_1, x_2, \cdots)$ we have

$$T\xi \,=\, (0, x_1, x_2, \cdots)\,. \tag{3.11}$$

Then $0$ is a spectral value since $T(l^2) \,=\, \{\xi\,;\, x_1 \,=\, 0\}$ is not dense in $l^2$. On $T(l^2)$ the inverse if the left shit $S\xi \,=\, (\xi_2, \xi_3, \cdots)$. On the other hand $0$ is not en egenvalue since the equation $T\xi = 0$ only the trivial solution $\xi = 0$. So $0 \in \sigma_r(T)$. ■

**Example 3.4.** Let $T$ be the multiplication operator on $L^2[0,1]$ given by

$$Tf(t) \,=\, tf(t)\,. \tag{3.12}$$

We have $\|Tf\|^2 \,=\, \int t^2|f(t)|^2\,dt \,\leq\, \|f\|^2$ showing that $\|T\| \,\leq\, 1$. it is left to the reader to prove that $\|T\| = 1$. note that

- If $\mu \notin [0,1]$ then $\mu \in \rho(T)$. Indeed we have

$$R_\mu f(t) \,=\, \frac{1}{t-\mu} f(t) \tag{3.13}$$

and $\|R_\mu\| \leq \operatorname{dist}(\mu, [0,1])^{-1}$

- If $\mu \in [0,1]$, $\mu$ is not eigenvalue since $tf(t) = \mu f(t)$ has only the solution $f(t) = 0$ a.e.
- If $\mu \in [0,1]$, then $T_\mu(L^2[0,1])$ is dense since it contains all $L^2$ functions which vanish in some neighborhood of $t_0 = \mu$. Then $T_\mu^{-1} f(t) = (t-\mu)^{-1} f(t)$ is defined on a dense set but unbounded.

Thus we have

$$\sigma(T) = \sigma_c(T) = [0,1], \quad \sigma_p(T) = \sigma_r(T) = \emptyset \tag{3.14}$$

■

The map

$$\begin{array}{rcl} R.(T) \,:\, \rho(T) & \longrightarrow & \mathcal{L}(V) \\ \mu & \longmapsto & R_\mu(T) \end{array}$$

is an *operator-valued* function on the resolvent set $\rho(T)$. We first investigate in this function can be understood as an "analytic" function in some way. Since $\mathcal{L}(V)$ is a Banach space we need to develop a bit the theory of *analytic Banach-spaced valued functions*.

**Definition 3.5.** Let $\Omega \subset \mathbb{C}$ be an open set, $V$ a Banach space and

$$\xi(\cdot) \,:\, \Omega \longrightarrow V \tag{3.15}$$

a map with valued in the Banach space $V$.

1. The map $\xi(z)$ is called *strongly differentiable* at $z_0 \in \Omega$ if

$$\xi'(z_0) \;=\; \lim_{h\to 0} \frac{1}{h}\left(\xi(z_0+h) - \xi(z_0)\right) \tag{3.16}$$

   exists. The map $\xi(z)$ is called *strongly analytic* in $\Omega$ if it is strongly differentiable at any $z \in \Omega$.

2. The map $\xi(z)$ is called *weakly differentiable* at $z_0 \in \Omega$ if for any linear functional $\lambda \in V'$ the complex valued function

$$z \mapsto \lambda(\xi(z)) \tag{3.17}$$

   is differentiable at $z_0$. The map $\xi(z)$ is called *weakly analytic* in $\Omega$ if it is weakly differentiable at any $z \in \Omega$.

We have

**Theorem 3.6.** *A Banach-space valued map $\xi(z)$ is strongly analytic if and only if it is weakly analytic.*

As a warm-up we have

**Lemma 3.7.** *Let $\{\xi_n\}$ be a sequence in the Banach space $V$. Then the sequence $\{\xi_n\}$ is Cauchy if and only if the sequence $\{\lambda(\xi_n)\}$ is Cauchy, uniformly for all $\lambda \in V'$, $\|\lambda\| \leq 1$.*

*Proof.* On one hand we have $\|\lambda(\xi_n) - \lambda(\xi_m)\| \leq \|\lambda\|\|\xi_n - \xi_m\| = \|\xi_n - \xi_m\|$ for $\|\lambda\| \leq 1$. On the other hand we have $\|\xi_n - \xi_m\| = \sup_{\|\lambda\|=1} |\lambda(\xi_n) - \lambda(\xi_m)|$ by Theorem 2.42. ∎

*Proof of Theorem 3.6.* Clearly strongly analytic implies weakly analytic. So let us assume that $\xi$ is weakly analytic in $\Omega \subset \mathbb{C}$, let $z_0 \in \Omega$ and $\Gamma$ a circle around $z_0$ contained in $\Omega$. For any $\lambda \in V'$ we have by Cauchy Theorem

$$\lambda\left(\frac{\xi(z_0+h)-\xi(z_0)}{h}\right) - \frac{d}{dz}\lambda(\xi(z_0)) = \frac{1}{2\pi i}\int_\Gamma \left[\frac{1}{h}\left(\frac{1}{z-(z_0+h)} - \frac{1}{z-z_0}\right) - \frac{1}{(z-z_0)^2}\right]\lambda(\xi(z)\,dz\,. \tag{3.18}$$

Since $\lambda(\xi(z))$ is continuous on $\Gamma$ there exists a constant $c_\lambda$ such that

$$\sup_{z\ in \Gamma} |\lambda(\xi(z))| \leq c_\lambda \tag{3.19}$$

So the family of

$$\begin{array}{rcl} \xi(z)\,;\, V' & \longrightarrow & \mathbb{C} \\ \lambda & \longmapsto & \lambda(\xi(z)) \end{array} \tag{3.20}$$

for $z \in \Gamma$ is a point wise bounded family of linear maps (use the isomorphism $V \to V''$). From the uniform boundedness theorem there exists a $c < \infty$ such that

$$\sup_{z\in\Gamma} \|\xi(z)\| = c \tag{3.21}$$

So we can bound (3.18) by

$$|(3.18)| \leq \frac{1}{2\pi} c\|\lambda\| \int_\Gamma \left|\left[\frac{1}{h}\left(\frac{1}{z-(z_0+h)} - \frac{1}{z-z_0}\right) - \frac{1}{(z-z_0)^2}\right]\right| d|z \leq \mathrm{Const}|h|\|\lambda\| \tag{3.22}$$

Therefore $\frac{1}{h}\lambda(\xi(z_0+h) - \xi(z_0))$ converges uniformly in $\lambda$ for all $\lambda$ with $\|\lambda\| \leq 1$. By the previous lemma this implies that $\frac{1}{h}(\xi(z_0+h) - \xi(z_0))$ is Cauchy for $|h| \to 0$. This concludes the proof. ∎

The theorem just proved is very useful since it allows us to speak simply of analytic Banach-space valued functions. All the theorem from analytic function theory can be used since they apply to the ordinary analytic function $\lambda(\xi(z))$ and so we can "lift" these results to the strongly analytic function $\xi(z)$.

**Theorem 3.8.** *Let $V$ be a Banach space and $T \in \mathcal{L}(V)$.*

1. *$\rho(T)$ is an open set in $\mathbb{C}$.*
2. *The resolvent $\mu \mapsto R_\mu(T)$ is analytic in $\rho(T)$.*

*Proof.* We use the Neumann series (see Theorem 2.23). if $\mu_0 \in \rho(T)$ consider the open ball

$$\Omega_0 = \left\{\mu \,:\, |\mu - \mu_0| < \|R_{\mu_0}\|^{-1}\right\}\,. \tag{3.23}$$

We have

$$\begin{aligned} T_\mu &= T_{\mu_0} - (\mu - \mu_0)\mathbf{1} \\ &= T_{\mu_0}\left[\mathbf{1} - (\mu - \mu_0)R_{\mu_0}\right] . \end{aligned} \tag{3.24}$$

Since $\|(\mu - \mu_0)R_{\mu_0}\| < 1$ we can invert $T_\mu$ and using Neumann series we obtain

$$R_\mu \;=\; \left[\mathbf{1} - (\mu - \mu_0)R_{\mu_0}\right]^{-1} R_{\mu_0} \;=\; \sum_{k=0}^{\infty} (\mu - \mu_0)^k R_{\mu_0}^{k+1} . \tag{3.25}$$

This is a convergent series in $\mathcal{L}(V)$ and so $R_\mu$ exist for $\mu \in \Omega_0$ and so $\Omega_0 \subset \rho(T)$. The power series also shows that $R_\lambda$ is is analytic. ∎

**Theorem 3.9.** *Let $V$ be a Banach space and $T \in \mathcal{L}(V)$. Then the spectrum $\sigma(T)$ is closed and non-empty.*

*Proof.* As the complement of $\rho(T)$, the spectrum $\sigma(T)$ is closed. By contradiction let us assume that $\sigma(T)$ is empty. Then $\rho(T) = \mathbb{C}$ and $R_\mu$ is an entire function. For $|\mu| \geq \|T\|$ we have the convergent series

$$R_\mu(T) \;=\; (T - \mu\mathbf{1})^{-1} \;=\; -\mu^{-1}(\mathbf{1} - \mu^{-1}T)^{-1} \;=\; -\mu^{-1}\sum_{k=0}^{\infty}(\mu^{-1}T)^k \tag{3.26}$$

and

$$\|R_\mu(T)\| \;\leq\; |\mu|^{-1}\sum_{k=0}^{\infty}\|\mu^{-1}T\|^k \;=\; \frac{\mu^{-1}}{1 - \|\mu^{-1}T\|} . \tag{3.27}$$

and thus $\|R_\mu(T)\| \longrightarrow 0$ as $|\mu| \to \infty$. Hence $R_\mu(T)$ is is a bounded entire function and by Liouville Theorem it as to be constant and identically null. But this is impossible and so $\sigma(T)$ is not empty. ∎

Eq. (3.26) also shows that

**Corollary 3.10.** *We have*

$$\sigma(T) \subset \{\mu \,:\, |\mu| \leq \|T\|\} \tag{3.28}$$

This can be strengthened into the following theorem which justifies the terminology spectral radius for $r(T)$.

**Theorem 3.11.** *Let $V$ be a Banach space and $T \in \mathcal{L}(V)$. Then we have*

$$r(T) \;=\; \sup_{\mu\in\sigma(T)} |\mu| \;=\; \max_{\mu\in\sigma(T)} |\mu| \tag{3.29}$$

*Proof.* The series (3.26) is nothing but the Laurent series for $R_\mu$ (expansion in power of $1/\mu$ around $\infty$). From complex analysis we know for a power series $f(z) = \sum_{n\geq 0} a_n z^n$ with convergence radius $r$ converges absolutely for $|z| < r$ but is not analytic in $\{|z| < (r + \epsilon)\}$. For the power series (3.26) this means that the series converges exactly outside the circle of radius $\sup_{\nu\in\sigma(T} |\nu|$.

We also know that the convergence radius of a power series is given by the formula

$$r^{-1} = \limsup_{n\to\infty} |a_n|^{1/n}, \tag{3.30}$$

that is in our case

$$r^{-1} = \limsup_{n\to\infty} \|T^n\|^{1/n} = \lim_{n\to\infty} \|T^n\|^{1/n} = r(T) \tag{3.31}$$

Finally note that since the spectra is closed we can replace the $\sup$ by a $\max$. ∎

Note that $\mathcal{L}(V)$ in addition of being a Banach space is also an algebra with the property that $\|TS\| \leq \|T\|\|S\|$. So we can certainly define *polynomial* of an element of $\mathcal{L}(V)$

$$p(T) = \sum_{k=1}^{N} a_k T^k \tag{3.32}$$

or more generally for *entire function* $f(z) = \sum_n a_n z^n$ we can define

$$f(T) = \sum_{k=1}^{\infty} a_n T^n \tag{3.33}$$

Note also that the formula (3.33) makes sense if $f$ has a convergence radius larger than the spectral radius $r(T)$.

In order to be more general we do not use a power series but instead use Cauchy integral formula.

**Definition 3.12.** Let $T \in \mathcal{L}(V)$, $f(z)$ a function analytic in a domain $\Omega$ containing $\sigma(T)$. Let $C$ be a contour in $\Omega \cap \rho(T)$ such that $C$ winds once around any point in $\sigma(T)$ and winds zero time around any point in $\Omega^C$. Set

$$f(T) = \frac{1}{2\pi i}\int_C (z-T)^{-1} f(z)dz = -\frac{1}{2\pi i}\int_C R_z(T) f(z)dz\,. \tag{3.34}$$

Note that the definition does not depend on the choice of $C$.

As a preparation for the next theorem we prove

**Lemma 3.13.** *(***Resolvent formula***)*

$$R_\mu(T)R_\nu(T) = \frac{R_\mu(T) - R_\nu(T)}{\mu - \nu} \tag{3.35}$$

*Proof.* We have

$$(T-\mu\mathbf{1}) - (T-\nu\mathbf{1}) = (\nu-\mu)\mathbf{1} \tag{3.36}$$

and so multiplying by $R_\mu R_\nu$ we have

$$R_\nu - R_\mu = (\nu-\mu)R_\mu R_\nu\,. \tag{3.37}$$

∎

The next theorem provides the basis for the functional calculus.

**Theorem 3.14.** *(***Functional calculus and spectral mapping theorem)** *Let $T \in \mathcal{L}(V)$.*

1. *If $f$ is a polynomial or analytic in a disk of radius greater than $\sigma(T)$ then the definition (3.34) coincides with the formulas (3.32) and (3.33).*

2. *The map (3.34) from the algebra of analytic functions on an open set containing $\sigma(T)$ into $\mathcal{L}(V)$ is a homomorphism.*

3. $$\sigma(f(T)) = f(\sigma(T)) \tag{3.38}$$

4. *If $f$ is analytic in an open set containing $\sigma(T)$ and $g$ is analytic in an open set containing $f(\sigma(T))$. If $h = g \circ f$ then*
$$h(T) = f(g(T)) \tag{3.39}$$

*Proof.* 1. Using the representation $R_\mu(T) = -\sum_{k=0}^{\infty} \mu^{-k-1}T^k$ we obtain by Cauchy integral formula

$$T^n = -\frac{1}{2\pi i}\int_C (T - z\mathbf{1})^{-1} z^n dz \tag{3.40}$$

and this shows that the formulas coincide.

For 2. we note that the mapping $f \mapsto f(T)$ is obviously linear. To show it is multiplicative we use the resolvent formula of Lemma 3.13. If $f$ and $g$ are analytic in a domain containing $\sigma(T)$ we pick two contours $C$ and $D$ as in definition 3.12 with $D$ lying inside $C$. Then we have

$$\begin{aligned} f(T)g(T) &= \left(\frac{1}{2\pi i}\right)^2 \int_C \int_D R_z(T)R_w(T)f(z)g(w)dzdw \\ &= \left(\frac{1}{2\pi i}\right)^2 \int_C \int_D \frac{R_z(T) - R_w(T)}{z - w} f(z)g(w)dzdw \\ &= \left(\frac{1}{2\pi i}\right)^2 \int_C \left[\int_D (z-w)^{-1} g(w)dw\right] R_z(T)f(z)dz \\ &\quad - \left(\frac{1}{2\pi i}\right)^2 \int_D \left[\int_C (z-w)^{-1} f(z)dz\right] R_w(T)g(w)dw \end{aligned} \tag{3.41}$$

Since $D$ lies inside $C$ we have $\int_D (z-w)^{-1}g(w)dw = 0$ while $\int_C (z-w)^{-1}f(z) = 2\pi i f(w)$ and thus

$$f(T)g(T) = -\int_D R_w(T)f(w)g(w)dw\,. \tag{3.42}$$

and thus $f(T)g(T) = h(T)$.

For 3. we have to show that $\mu$ belongs to the spectrum of $f(T)$ if and only if $\mu$ is of the form

$$\mu = f(\nu)\,, \quad \nu \in \sigma(T)\,. \tag{3.43}$$

If $\mu$ is not of the form (3.43), then $f(z) - \mu$ does not vanish on $\sigma(T)$. Therefore $g(z) \equiv (f(z) - \mu)^{-1}$ is analytic in an open set containing $\sigma(T)$. According to part 2. $(f(T) - \mu\mathbf{1})g(T) = h(T) = \mathbf{1}$. So $g(T)$ is the inverse of $(f(T) - \mu)$ and so $\mu \notin \sigma(f(T))$.

Conversely suppose that $\mu$ is of the form (3.43). Define $k(z)$ by

$$k(z) = \frac{f(z) - f(\nu)}{z - \nu}\,. \tag{3.44}$$

The function $k(z)$ is analytic in an open set containing $\sigma(T)$ so $k(T)$ can be defined using by (3.34). Since $(z-\nu)k(z) = f(z) - f(\nu)$ by part 2. we have

$$(T - \nu\mathbf{1})k(T) \;=\; f(T) - f(\nu)\mathbf{1}\,. \tag{3.45}$$

Since $\nu \in \sigma(T)$ the first factor is not invertible and so $f(T) - f(\nu)$ is not invertible either.

For 4. by assumption $g(w)$ is analytic on $f(\sigma(T))$. Since by 3. the spectrum of $f(T)$ is $f(\sigma(T))$ we can apply (3.34) to $g$ in place of $f$ and $f(T)$ in place of $T$:

$$g(f(T)) \;=\; \frac{1}{2\pi i}\int_D (w\mathbf{1} - f(T))^{-1} g(w)dw \tag{3.46}$$

If $w \in D$ then $w - f(z)$ is an analytic function on $\sigma(T)$, then applying (3.34) again

$$(w - f(T))^{-1} \;=\; \frac{1}{2\pi i}\int_C (z\mathbf{1} - T)^{-1}(w - f(z))^{-1} dz \tag{3.47}$$

provided $C$ does not wind around any point of $D$. Combining the tow formulas we find

$$g(f(T)) \;=\; \left(\frac{1}{2\pi i}\right)^2 \int_D \int_C (z - T)^{-1}(w - f(z))^{-1} g(w) dz dz\,. \tag{3.48}$$

We interchange the order to the integral and since $D$ winds around every point $z \in C$ we have by Cauchy integral formula

$$\frac{1}{2\pi i}\int_D (w - f(z))^{-1} g(w)dw \;=\; g(f(z)) = h(z)\,. \tag{3.49}$$

Setting this back in (3.48) we find that $g(f(T)) = h(T)$. ∎

Suppose $\sigma(T)$ can be decomposed into $N$ *pairwise disjoint closed components*:

$$\sigma(T) \;=\; \sigma_1 \cup \cdots \cup \sigma_N\,, \quad \sigma_j \cap \sigma_k = \emptyset \tag{3.50}$$

Since the $\sigma_i$ are closed they are at positive distance from each other and we can pick contours $C_j$ that winds once around each point of $\sigma_j$ but not $\sigma_k$, $k \neq j$. We set

$$P_j \;=\; \frac{1}{2\pi i}\int_{C_j} (z\mathbf{1} - T)^{-1} dz \tag{3.51}$$

**Theorem 3.15.** *The $P_j$ are disjoint projections, i.e.*

$$P_j^2 = P_j \quad and\; P_j P_k = 0 \quad for\; j \neq k \tag{3.52}$$

*and we have*

$$\sum_j P_j \;=\; \mathbf{1}\,. \tag{3.53}$$

*Proof.* Pick open set $\Omega_i$ such that $\Omega_i$ contains $\sigma_i$ and the $\Omega_i$ are pairwise disjoint. Set $\Omega = \bigcup_{j=1}^N \Omega_j$. Then consider the function $f_i(z)$ which is equal to 1 in $\Omega_i$ and 0 in $\Omega \setminus \Omega_i$. These functions are analytic in $\Omega$ and satisfy $f_i(z)^2 = f_i(z)$ as well as $f_i f_j = 0$ for $i \neq j$. By functional calculus we obtain the theorem. ∎

**Theorem 3.16.** *Let $V$ be a Banch space and $T \in LV$ and let $T' \in \mathcal{L}(V')$ it adjoint. Then*

$$\sigma(T) = \sigma(T') \tag{3.54}$$

**Corollary 3.17.** *If $T$ is a bounded operator on a Hilbert space and $T^*$ its (Hilbert-)adjoint. Then*

$$\sigma(T^*) \;=\; \overline{\sigma(T)} \tag{3.55}$$

We need the following lemma. For a bounded linear operator $T$ we denote $N(T)$ the nullspace of $T$ and $R(T)$ the range of $T$. For a subspace $W \subset V$ we define the annihilator of $W$, denoted by $W^\perp$ as the set of linear functional which vanish on $W$. For a subset $W'$ of $V'$ we define the annihilator of $W'$ (denoted by $W' \perp$) as the set of all vectors in $V$ annihilated by all functional in $W'$.

**Lemma 3.18.** *We have*

$$N(T') = R(T)^\perp\,, \quad N(T) = R(T')^\perp \tag{3.56}$$

*Proof.* Use the duality relation $\lambda(T\xi) = T'\lambda(\xi)$. The details are left as an exercise.

The theorem 3.16 is an immediate consequence of

**Theorem 3.19.** *$T \in \mathcal{L}(V)$ is invertible if and only if $T' \in \mathcal{L}(V')$ is invertible.*

*Proof.* IF $T$ is invertible with inverse $S$ then

$$TS = ST = \mathbf{1}_V\,. \tag{3.57}$$

Taking the adjoint gives

$$S'T' = T'S' \;=\; \mathbf{1}_{V'} \tag{3.58}$$

which shows that $S'$ is the inverse of $T'$. If $V$ is a reflexive space then the relation between $T$ and $T'$ is symmetric so the proof is complete. In case $V$ is not reflexive an additional argument is needed. If $T'$ is invertible then by taking adjoint

$$T''S'' = S''T'' = \mathbf{1}_{V''} \tag{3.59}$$

Since $T''$ and $I_{V''}$ restricted to $V$ are equal to $T$ and $I_V$ respectively it follows that the nullspace of $T$ is trivial and so $T$ is one- to-one. So $S''$ restricted to the range of $T$ is inverse to $T$. Suppose now that the range of $T$ is not all of $V$. Then we can find by Hahn- Banach a non-zero functional $\lambda \in V'$ with $\lambda = 0$ on the range of $T$. According to the previous lemma such a functional belongs to the nullpsace of $T'$. Since $T'$ is invertible this is impossible. ∎

**Example 3.20.** We return to the shift $T : l^2 \to l^2$ given by $T\xi = (0, x_1, x_2, \cdots)$ with adjoint $S\xi = (x_2, x_3, \cdots)$. We claim that

$$\sigma(T) = \sigma(S) \;=\; \{z \in \mathbb{C} : |z| \le 1\} \tag{3.60}$$

It is easy to check that $\|L\| = 1$ and similarly that $\|L^n\| = 1$ for all $n$ and thus $r(T) = 1$. So no number $z$ with $|z| \ge 1$ belongs to $\sigma(S)$. Next let us try to find eigenvalues for $L$: $(x_2, x_3, cdots) = \mu(x_1, x_2, cdots)$ which implies

$$x_n = \mu^{n-1} x_1 \tag{3.61}$$

We have then $\sum_n |x_n|^2 < \infty$ if and only if $|\mu| < 1$. So any $\mu$ in the open disk $\{z : |z| < 1\}$ is an eigenvalue. Since the spectrum is closed then $\sigma(S)$is the close unit disk. Finally the spectrum of $T$ is the same as the spectrum of $S$. The reader should check whether $T$ has any eigenvalues.

**Example 3.21. Nilpotent and quasi-nilpotent operators**
An operator is called *nilpotent* if $N^k = 0$ for some $k > 1$. Then $\sigma(N^k) = \{0\|$ and by the spectral mapping theorem $\sigma(N) = \{0\}$.
The Voltera operator acting on $C[a, b]$

$$Kf(t) = \int_0^t K(s,t) f(t)\, dt \tag{3.62}$$

with continuous kernel $k(t, s)$ was shown to have spectral radius $r(T) = 0$ in the exercises. Therefore

$$\sigma(K) = \{0\}. \tag{3.63}$$

An operator $K$ with $\sigma(K) = \{0\}$ but $N^k \neq 0$ for all $k$ is called *quasi-nilpotent*.

**Example 3.22. Fourier transform** The Fourier transform $T$ is defined by

$$\widehat{f}(k) = Tf(k) = \int_{\mathbb{R}} f(x) e^{-i2\pi xk} \tag{3.64}$$

As we have show in Math 623/624 $T$ is an invertible norm preserving map from $L^2(\mathbb{R})$ onto $L^2(\mathbb{R})$. The inverse is given by

$$f(x) = \int_{\mathbb{R}} \widehat{f}(k) e^{i2\pi xk}\, dk\,. \tag{3.65}$$

Consider the mapping $R$ given by

$$Rf(x) = f(-x)\,, \tag{3.66}$$

and obviously $R^2 = \mathbf{1}$. From (3.65) we have

$$f(-x) = \int_{\mathbb{R}} \widehat{f}(k) e^{-i2\pi xk}\, dk\,. \tag{3.67}$$

So we have $T^2 = R$ and thus

$$T^4 = \mathbf{1} \tag{3.68}$$

By the spectral mapping theorem we immediately obtain that

$$\sigma(T) \subset \{+1, -1, +i, -i\} \tag{3.69}$$

More details on the spectrum of $T$ will be given in the HWK.

## 3.2 Compact operators: Basic properties

One of the workhorse of analysis in finite-dimensional spaces is the Bolzano-Weierstrass theorem which asserts that that a set is compact if and only if it is closed and bounded. In infinite dimensional spaces this theorem fails.

**Theorem 3.23.** *The closed unit ball of a Banach space $V$ is compact if and only if $V$ is finite dimensional.*

Before we prove the theorem note that if $H$ is a separable Hilbert space with an orthonormal basis $\{\xi_n\}$ then we have

$$\|\xi_n - \xi_m\|^2 = (\xi_n - \xi_m, \xi_n - \xi_m) = \|\xi_n\|^2 + \|\xi_m\|^2 = 2\,. \tag{3.70}$$

and so the sequence $\xi_n$ satisfies $\|\xi_n\| = 1$ for all $n$ and $\|\xi_n - \xi_m\| = \sqrt{2}$ for $n \neq m$. Hence $\{\xi_n\}$ has no convergent subsequence and so the unit ball is not compact.

In order to adapt this argument to a general Banach space we need the following lemma.

**Lemma 3.24.** *Let $V$ be a Banach space and $W$ a proper closed subspace of $V$. Then there exists*

$$\zeta \in V \text{ with } \|\zeta\| = 1 \quad \text{and} \quad \|\zeta - \eta\| > 1/2 \text{ for all } \eta \in W\,. \tag{3.71}$$

*Proof.* Since $W$ is a proper there exists $\xi \in V$ with $\xi \notin W$ and since $W$ is closed

$$\inf_{\eta \in W} \|\xi - \eta\| = d > 0\,. \tag{3.72}$$

We pick $\eta_0 \in W$ such that $\|\xi - \eta_0\| \leq 2d$ and set $\zeta' = \xi - \eta_0$. Then we have

$$\|\zeta'\| \leq 2d \text{ and } \|\zeta' - \eta\| = \|\xi - (\eta + \eta_0)\| \geq d \text{ for all } \eta \in W. \tag{3.73}$$

Finally we set $\zeta = \zeta'/\|\zeta'\|$ so that $\|\zeta\| = 1$ and

$$\|\zeta - \eta\| > 1/2 \text{ for all } \eta \in W. \tag{3.74}$$

■

Equipped with this lemma we obtain easily

*Proof of Theorem 3.23.* Let $\{\xi_n\}$ be the sequence constructed inductively as follows. Pick an arbitrary $\xi_1$ with $\|\xi_1\| = 1$. Then if $V_n$ is the subspace spanned $\xi_1, \cdots, \xi_n$, it is a closed proper subspace since it is finite-dimensional and thus by Lemma 3.24 there exists $\xi_{n+1}$ with $\|\xi_{n+1}\| = 1$ and $\|\xi_{n+1} - \xi_l\| > 1/2$ for $l = 1, \cdots, n$. The sequence $\{\xi_n\}$ does not have a convergent subsequence. ■

A bounded operator has the property to maps bounded sets into bounded sets. A operator will be called compact if it transforms bounded sets into sets whose closure compact. To make this precise we recall

**Definition 3.25.** Let $X$ be a complete metric space. The set $S$ is called *precompact* if its closure $\overline{S}$ is compact.

- $S$ is precompact if and only if any sequence $\{\xi_n\} \subset S$ contains a Cauchy subsequence.
- $S$ is precompact if and only if for any $\epsilon > 0$ $S$ can be covered by finitely many balls of radius less than $\epsilon$.

**Definition 3.26.** Let $V$ and $W$ be Banach spaces.

1. A linear map $T : V \to W$ is called *compact* if the image of the unit ball $T(B_1(0))$ is precompact. (This is implies that $T$ maps any bounded set into a precompact set.)
2. We denote by $\mathcal{C}(V, W)$ the set of all compact operators and set $\mathcal{C}(V) \equiv \mathcal{C}(V, V)$

Let us derive some of the elementary properties of compact operator.

**Theorem 3.27.** *Let $U, V$, and $W$ be Banach spaces.*

1. *If $T, S \in \mathcal{C}(V, W)$ then $T + S \in \mathcal{C}(V, W)$*
2. *If $T \in \mathcal{C}(V, W)$ and $\alpha \in \mathbb{K}$ then $\alpha T \in \mathcal{C}(V, W)$.*
3. *If $T \in \mathcal{C}(U, V)$ and $S \in \mathcal{L}(V, W)$ then $ST \in \mathcal{C}(U, W)$*
4. *If $T \in \mathcal{L}(U, V)$ and $S \in \mathcal{C}(V, W)$ then $ST \in \mathcal{C}(U, W)$*
5. *If $\{T_n\}$ is a sequence in $\mathcal{C}(V, W)$ with $\|T_n - T\| \to 0$ then $T \in \mathcal{C}(V, W)$.*

This means that the set of compact operators is a *closed subspace* of the vector space of bounded operators and that $\mathcal{C}(V)$ is a *closed two-sided ideal* in the algebra $\mathcal{L}(V)$.

*Proof.* For 1. let $\{\xi_n\}$ be a sequence in $B_1(0)$ then since $T$ is compact there exists a subsequence $\{\xi_{n_k}\}$ such that $\{T\xi_{n_k}\}$ converges. Since $S$ is compact there exists a subsubsequence $\{\xi_{n_{k_j}}\}$ such that $\{S\xi_{n_{k_j}}\}$ and $\{(T + S)\xi_{n_{k_j}}\}$ converges too.

2. is a special case of 3. and 3. itself follows from the fact that a bounded map maps precompact set into precompact sets. In turn 4. is obvious.

For 5. given $\epsilon > 0$ we pick $n$ such that $\|T_n - T\| \le \epsilon/2$ and since $T_n$ is compact then $T_n(B_1(0))$ can be covered by finitely many balls of radius $\epsilon/2$. Then $T_(B_1(0))$ can be covered by finitely many balls of radius $\epsilon$. ■

**Definition 3.28.** Let $V$ and $W$ be Banach spaces. A linear operator $T \in \mathcal{L}(V)$ is a *finite rank* operator if its range $R(T)$ is finite dimensional.

Clearly finite-rank operators are compact. A finite rank operator $T \in \mathcal{L}(V, W)$ can always be written as follows. Pick $\eta_1, \cdots, \eta_N \in W$ and $\lambda_1, \cdots, \lambda_N \in V'$ and set

$$T\xi = \sum_{j=1}^{N} \lambda_j(\xi)\eta_j \tag{3.75}$$

In general it is *not true* that the set of finite rank operators is dense in the set of compact operators even on separable Banach spaces. It is true in Hilbert spaces as well as in many standard Banach spaces. We will prove this property for separable Hilbert spaces in the next section. As we will see in examples one can sometimes prove that an operator is compact by proving that it can be approximated by finite-rank operators.

We conclude this section with a first series of examples of compact operators. In order to prove compactness we will use two classical results of analysis. The first one characterizes compact sets for spaces of continuous functions.

**Theorem 3.29. (Arzela-Ascoli)** *Let $K$ be a compact metric space and let $C(K)$ the Banach space of complex-values continuous functions with norm $\|f\| = \sup_{x\in K} |f(x)|$. Let $\{f_\alpha\}_{\alpha\in I}$ be a collection of function such that*

1. *$\{f_\alpha\}$ is uniformly bounded: there exists $c < \infty$ such that $\sup_\alpha \|f_\alpha\| \le c$.*

2. *$\{f_\alpha\}$ is equicontinuous: given $\epsilon > 0$ there exists $\delta > 0$ such that*

$$d(x,y) \le \delta \Longrightarrow \|f_\alpha(x) - f_\alpha(y)\| \le \epsilon \text{ for all } \alpha \in I \tag{3.76}$$

*Then the family $\{f_\alpha\}$ is a precompact subset of $(C(K)$.*

A related theorem gives a compactness criterion for $L^2(Q)$ where $Q$ is a bounded domain.

**Theorem 3.30. (Rellich)** *Let $Q$ be open and bounded with smooth boundary. Let $\{f_\alpha\}_{\alpha\in I}$ be a family of functions in $L^2(Q)$ such that*

1. *The family $\{f_\alpha\}$ is uniformly bounded: there exists $c < \infty$ such that $\sup_\alpha \|f_\alpha\| \le c$.*
2. *The derivative of $\{f_\alpha\}$ are uniformly bounded: there exists $c < \infty$ such that $\sup_\alpha \|\partial_{x_i} f_\alpha\| \le c$ for $i = 1, 2, \cdots, d$.*

*Then the family $\{f_\alpha\}$ is a precompact subsets of $L^2(Q)$.*

**Example 3.31. (Integral operator I)** Consider the integral operator

$$Tf(t) \,=\, \int_a^b k(t,s) f(s)\, ds\,. \tag{3.77}$$

If $k(t,s) \in C[a,b] \times [a,b]$ then we show that $T : C[a,b] \to C[a,b]$ is compact. Let $\{f_n\}$ be a bounded sequence, e.g. $\|f_n\| \le 1$ for all $n$, then we show that the sequence $\{Tf_n\}$ satisfies the conditions of Arzela-Ascoli theorem. We have

$$\begin{aligned} |Tf(t) - Tf(t')| &\le \int_a^b |k(t,s) - k(t',s)||f(s)|\, ds \\ &\le \underbrace{\|f_n\|}_{\le 1} \int_a^b |k(t,s) - k(t',s)|\, ds \end{aligned} \tag{3.78}$$

By the continuity of $k(t,s)$ the right hand side of (3.78) goes to 0 as $t \to t'$ uniformly in $n$ and thus $\{Tf_n\}$ is an equicontinuous family. ■

**Example 3.32. (Integral operator II)** Consider an integral operator on some separable Hilbert space $L^2(X,\mu)$

$$Tf(t) \,=\, \int k(t,s) f(s)\, d\mu\,. \tag{3.79}$$

with $k(t,s) \in L^2(X \times X, \mu \times \mu)$. It was proved in the exercise that $T$ is bounded with

$$\|T\| \,\le\, \left(\int\int |k(t,s)|^2 d\mu(t) d\mu(s)\right)^{1/2}\,. \tag{3.80}$$

We show that $T$ is compact by showing that is it the limit of a sequence of finite-rank operators. Note that a finite rank integral operators has the form

$$Tf(t) \,=\, \sum_{j=1}^N f_j(s) \int g_j(s) f(s)\, ds \tag{3.81}$$

for some $f_j, g_j \in L^2$. We now expand the $k(t,s)$ (for fixed $s$) using an o.n.b. $\{f_j\}$ of $L^2$ we have

$$k(t,s) \;=\; \sum_{j=1}^{\infty} f_j(t)k_j(s)\,, \qquad k_j(s) \;=\; \int k(t,s)\overline{f_j(t)}\,d\mu(t) \tag{3.82}$$

By Parseval equality we have

$$\int |k(t,s)|^2 d\mu(t) \;=\; \sum_{j=1}^{\infty} |k_j(s)|^2 \tag{3.83}$$

and

$$\int\int |k(t,s)|^2 d\mu(s) d\mu(t) \;=\; \sum_{j=1}^{\infty} \int |k_j(s)|^2 d\mu(s)\,. \tag{3.84}$$

Let us define

$$k_N(t,s) \;=\; \sum_{j=1}^{N} f_j(t)k_j(s) \tag{3.85}$$

and define $T_N f(t) \;=\; \int k_N(t,s)f(s)d\mu(s)$. This is a finite rank operator and we have by the same calculation has for $\|T\|$

$$\|T - T_N\|^2 \;\leq\; \int\int |k_N(t,s)|^2 d\mu(t)d\mu(s) \;=\; \sum_{j=N}^{\infty} \int |k_j(s)|^2 d\mu(s) \tag{3.86}$$

Because of (3.83), the right side of (3.86) goes to 0 as $N \to \infty$. Hence $T$ as the limit of a sequence of finite-rank operators is compact. ∎

**Example 3.33.** (**Laplace equation**) It is well known that the boundary value problem

$$\Delta u = f \text{ in } Q \qquad u = 0 \text{ on } \partial Q \tag{3.87}$$

has a unique solution $u$ for every $f \in C^\infty(Q)$. Let us denote by $S$ the linear map

$$u = Sf \tag{3.88}$$

which gives the solution of (3.87). We claim that $S$ defines a compact map from $L^2(Q)$ into $L^2(Q)$.

We have

**Lemma 3.34.** *Let $f$ be compactly supported in $Q$. Then we have*

$$\int |f|^2 dx \;\leq\; C\int \sum_j |f_j|^2 dx \quad \text{ with } f_j \equiv \partial_{x_j} f\,. \tag{3.89}$$

*Proof.* Since $f$ vanishes on the boundary of $Q$ we have, at any point $x \in Q$

$$f(x) \;=\; \int_{x_b}^{x} f_1 dx_1 \tag{3.90}$$

where $x_b$ is a point on the boundary of $Q$ with the same $x_2, \cdots, x_d$ coordinates as $x$. By Cauchy-Schwartz we have

$$f(x)^2 \leq d \int |f_1|^2 dx\,, \tag{3.91}$$

and integrating over $Q$ gives the result. ■

Let us denote $\|f\|_0 = \|f\|_{L^2}$ and $\|f\|_1 = (\int \sum_j |f_j|^2 dx)^{1/2}$. Then the lemma asserts that $\|f\|_0 \leq c\|f\|_1$ for smooth. Now multiply (3.87) by $u$, integrate over $Q$ and integrate by parts:

$$-\int \sum_j |u_j|^2\, dx = \int_Q f u\, dx \tag{3.92}$$

Using Cauchy-Schwartz and the lemma this yields

$$\|u\|_1^2 \leq \|f\|_0 \|u\|_0 \leq , c\|f\|_0 \|u\|_0\,. \tag{3.93}$$

from which we obtain

$$\|u\|_1 \leq c\|f\|_0\,, \quad \text{and} \quad \|u\|_0 \leq , c^2\|f\|_0\,. \tag{3.94}$$

This shows that the image of the unit ball $\{\|f\|_0 \leq 1\}$ is mapped by $S$ to solutions $u$ which satisfies the conditions of Theorem 3.30. Hence $S$ is a compact map. ■

**Example 3.35.** (**Heat equation**) Let us consider the heat equation (initial value problem)

$$u_t = \Delta u\,, \quad u(x,0) = u(x) \tag{3.95}$$

for function $u(x,t)$, with $x \in Q$ a bounded open domain with smooth boundary and $t > 0$. It is well-known that the initial value problem where $u(x,0) = u(x)$ is given has a unique solution for all $t > 0$. Let us denote by $S_T$ the operator mapping the initial condition to the solution at time $T$

$$S_T(u) = u(x,T) \tag{3.96}$$

We have

**Theorem 3.36.** *The map $S_T : L^2(Q) \to L^2(Q)$ is compact for any $T > 0$.*

*Proof.* Multiply (3.95) by $u$ and integrate with respect to $x \in Q$ and $t \in [0,T]$. After integrating by parts one gets

$$\begin{aligned} \int_0^T \int_Q u u_t dx dt &= \int_Q \frac{1}{2}|u|^2(x,T)\, dx - \int_Q \frac{1}{2}|u|^2(x,0)\, dx \\ &= \int_0^T \int_Q u\Delta u = -\int_0^T \int_Q \sum_j |u_j|^2\, dx dt \leq 0 \end{aligned} \tag{3.97}$$

and thus

$$\int_Q |u|^2(x,T)\, dx \leq \int_Q |u|^2(x,0)\, dx \tag{3.98}$$

i.e. the $L^2$ norm of $u$ is decreasing in $t$ or

$$\|S_T\| \le 1\,. \tag{3.99}$$

Next let us multiply (3.95) by $t\Delta u$ and integrate with respect to $x$ and $t$. We get

$$\int_0^T \int_Q t u_t \Delta u \, dxdt \;=\; \int_0^T \int_Q t(\Delta u)^2 \, dxdt \;\ge\; 0 \tag{3.100}$$

Integrating by parts the left side of (3.100) with respect to $x$ and then $t$ we find

$$\begin{aligned}
\int_0^T \int_Q t u_t \Delta u \, dxdt &= -\int_0^T \int_Q t \sum_j u_{tj} u_j \, dxdt \\
&= -\frac{1}{2}\int_0^T \int_Q t \frac{d}{dt} \sum_j |u_j|^2 \, dxdt \\
&= \frac{1}{2}\int_0^T \int_Q \frac{d}{dt} \sum_j |u_j|^2 \, dxdt - \frac{T}{2}\int_Q |u_j(x,T)|^2 \, dx
\end{aligned} \tag{3.101}$$

Combining with (3.97) we find

$$\begin{aligned}
\frac{T}{2}\int_Q |u_j(x,T)|^2 \, dx &\le \frac{1}{2}\int_0^T \int_Q \frac{d}{dt} \sum_j |u_j|^2 \, dxdt \\
&= \frac{1}{2}\int_Q |u(x,0)|^2 \, dx - \frac{1}{2}\int_Q |u(x,T)|^2 \, dx \\
&\le \frac{1}{2}\int_Q |u(x,0)|^2 \, dx
\end{aligned} \tag{3.102}$$

Combining (3.98) and (3.102) show that

$$\|S_T u\|_0 \le \|u\|_0 \qquad \|S_T u\|_1 \le \frac{1}{2T}\|u\|_0 \tag{3.103}$$

which implies that $S_T$ is compact by Rellich Theorem. ∎

## 3.3 Spectral theory of compact operators

The general result is the following

**Theorem 3.37.** *Let $V$ be a Banach space and $T \in \mathcal{L}(V)$ a compact operator.*

*1.* **Spectrum:**

   *(a) If $\lambda \neq 0$ is in $\sigma(T)$ then $\lambda \in \sigma_p(T)$.*

   *(b) $\sigma_p(T)$ is countable and the only possible accumulation point is $0$ which may ar may not belong to $\sigma_p(T)$.*

   *(c) For $\lambda \neq 0$, the eigenspace $E_\lambda = \{\xi : T\xi = \lambda\xi\}$ is finite dimensional.*

(d) *The adjoint $T'$ has the same eigenvalues and the eigenspaces have the same dimensions.*

2. **Fredholm Alternative:** *For $T_\mu = T - \mu\mathbf{1}$ we have either (a) or (b)*

   (a) *$T_\mu \xi = \eta$ and $T'_\mu \alpha = \beta$ have unique solutions $\xi$ and $\alpha$ for every $\eta \in V$, $\beta \in V'$. In particular if $\eta = 0$, $\beta = 0$ then $\xi = 0$ and $\alpha = 0$.*

   (b) *$T_\mu \xi = 0$ and $T'_\mu \alpha = 0$ have $n$ linearly independent solutions $\xi_1, \xi_n$ (resp. $\alpha_1, \cdots, \alpha_n$) and $T_\xi = \eta$ and $T'\alpha = \beta$ have solution if and only if*

$$\alpha_k(\eta) = 0\,, k = 1, \cdots, n \qquad \text{resp.} \quad \beta(\xi_k) = 0 k = 1, \cdots, n \tag{3.104}$$

We shall not prove this theorem here (see e.g. Lax for a proof) in full generality but we are going to concentrate on the case of *separable Hilbert spaces*. We show first that finite rank operators are dense in separable Hilbert spaces. This remains true for non-separable Hilbert spaces but maybe surprisingly this in general not true even for separable Hilbert spaces, although this holds for many of the usual Banach spaces.

**Theorem 3.38.** *Let $H$ be a separable Hilbert space and $T \in \mathcal{C}(H)$. Then there exists a sequence of finite rank operator $\{T_n\}$ such that $\|T_n - T\| \to 0$.*

*Proof.* Pick an o.n.b $\{\xi_i\}$ of $H$ and let us set

$$P_n \xi \,=\, \sum_{k=1}^n (\xi\,,\,\xi_k)\xi_k\,, \quad T_n \equiv T P_n\,. \tag{3.105}$$

Then we have

$$\begin{aligned} \|T - T_n\| &= \sup_{\|\xi\|=1} \|(T - T_n)\xi\| \\ &= \sup_{\|\xi\|=1} \|T(\mathbf{1} - P_n)\xi\| \\ &= \sup_{\|\xi\|=1, P_n\xi=0} \|T\xi\| \equiv \lambda_n \end{aligned} \tag{3.106}$$

Since $P_{n+1}\xi = 0$ implies $P_n \xi = 0$ we have

$$\lambda_{n+1} \le \lambda_n, , \quad \lim_{n\to\infty} \lambda_n = \lambda \text{ exists}\,. \tag{3.107}$$

We need to show $\lambda = 0$. For any $n$ let us pick $\eta_n$ with $\|\eta_n\| = 1$, $P_n \eta_n = 0$ and $\|T\eta_n\| \ge \lambda_n/2$. We have

$$\lim_n (\eta_n\,,\,\xi) \,=\, 0 \text{ for all } \xi \in H\,, \tag{3.108}$$

and thus

$$\lim_n (T\eta_n\,,\,\xi) \,=\, (\eta_n, T\xi) \,=\, 0 \text{ for all } \xi \in H\,, \tag{3.109}$$

Suppose that $\lim_n \|T\eta_n\| \neq 0$, then for any $\epsilon > 0$ there exists a subsequence $\eta_{n_k}$ such that

$$\|T\xi_{n_k}\| > \epsilon\,, \quad k = 1, 2, \cdots \tag{3.110}$$

Since $T$ is compact there exists a subsubsequence $\{\eta_{n_{k_j}}\}$such that $\{T\eta_{n_{k_j}}\}$ is convergent, i.e.

$$\lim_j T\eta_{n_{k_j}} \,=\, \eta \neq 0 \tag{3.111}$$

But then

$$\lim_j (T\eta_{n_{k_j}}\,,\,\eta) \;=\; \|\eta\|^2 \neq 0 \tag{3.112}$$

which contradicts (3.109). Therefore we have $\lim_n \|T\eta_n\| \;\neq\; 0$ and so $\lambda = 0$. ∎

As preparation we have

**Theorem 3.39.** **(Analytic Fredholm theory)** *Let $\Omega \subset \mathbb{C}$ be a connected domain and $T(\cdot) : \Omega \to \mathcal{C}(V)$ an analytic operator-valued map taking value in the compact operators for all $z \in \Omega$. Then we have one of the two following alternatives*

1. *$1$ is in the spectrum of $T$ for all $z \in \Omega$.*
2. *$1 \in \rho(T(z))$ for all $z \in \Omega \setminus \Sigma$ where $\Sigma$ is a discrete set without accumulation point. The resolvent $(T(z) - \mathbf{1})^{-1}$ is meromorph in $\Omega$, analytic in $\Omega \setminus \Sigma$. The residue by the poles are finite rank operators and for $z \in \Sigma$, $T(z)\xi = \xi$ has a nontrivial solution space of finite dimension.*

The proof of this theorem as well as the next few ones uses the following idea and a basic construction which we now explain. The idea is to reduce the solvability of $T(z)\xi = \xi$ to the solvability of $f(z) = 0$ for some analytic function. Now we have either $f \equiv 0$ or $f$ vanish on a discrete set. In order to do this we will write $f(z)$ has the determinant of some matrix obtained from a finite rank operator $S(z)$ and the solvability of $T(z)\xi = \xi$ is equivalent to the solvability of $T(z)\xi = \xi$.

The basic construction is as follows: Given $z_0 \in \Omega$ and $z \in \Omega_0 \equiv \{z : \|T(z) - T(z_0)\| < 1/2\}$ we pick $\widehat{T}$ an operator of finite rank such that

$$\|T(z_0) - \widehat{T}\| < 1/2 \tag{3.113}$$

Then

$$\|T(z) - \widehat{T}\| < 1\,, \tag{3.114}$$

and

$$A(z) \;=\; (\mathbf{1} - (T(z) - \widehat{T})) \tag{3.115}$$

is invertible for $z \in \Omega_0$. We now consider the *finite rank operator*

$$S(z) \;=\; \widehat{T}A(z)^{-1} \tag{3.116}$$

and note that

$$T(z) - \mathbf{1} = (S(z) - \mathbf{1})A(z) \tag{3.117}$$

*Proof of Theorem 3.39.* With $S(z)$ as in (3.117) we note that

$$T(z) - \mathbf{1} \text{ invertible } \Longleftrightarrow S(z) - \mathbf{1} \text{ invertible.} \tag{3.118}$$

and

$$T(z)\xi = \xi \text{ has nontrivial solutions} \Longleftrightarrow S(z)\xi = \xi \text{ has nontrivial solutions} \tag{3.119}$$

Note that the range of $S(z) = \widehat{T}A(z)^{-1}$ is always contained in $H_0 = R(\widehat{T})$ which is independent of $z$ and $\dim(H_0) = n < \infty$. We decompose

$$H = H_0 \oplus H_1 \quad \text{with } \; H_1 = H_0^{\perp} \tag{3.120}$$

We have a unique decomposition $\xi = \xi_0 + \xi_1$ for any $\xi \in H$ and if we write

$$\xi = \begin{bmatrix} \xi_0 \\ \xi_1 \end{bmatrix} \tag{3.121}$$

and we have the matrix-like representation

$$A = \begin{bmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{bmatrix}, \quad A_{ij} \in \mathcal{L}(H_j, H_i)\,. \tag{3.122}$$

Since $R(S(z)) = H_0$ we have (with $S_{ij} \equiv S_{ij}(z)$)

$$S(z) = \begin{bmatrix} S_{00} & S_{01} \\ 0 & 0 \end{bmatrix} \tag{3.123}$$

and so

$$(S(z) - \mathbf{1})\xi \;=\; \begin{bmatrix} S_{00} & S_{01} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \xi_0 \\ \xi_1 \end{bmatrix} \;=\; \begin{bmatrix} (S_{00} - \mathbf{1})\xi_0 + S_{01}\xi_1 \\ -\xi_1 \end{bmatrix} \tag{3.124}$$

Therefore $(S(z) - \mathbf{1})\xi = 0$ implies $\xi_1 = 0$ and $(S_{00} - \mathbf{1})\xi_0 = 0$ and thus

$$\dim\,\{\xi\, : S(z)\xi = \xi\} \;=\; \dim\,\{\xi_0\, : S_{00}(z)\xi_0 = \xi_0\} \;=\; n < \infty \tag{3.125}$$

In particular $S(z)\xi = \xi$ has a nontrivial solution if and only if

$$f(z) \;\equiv\; \det(S_{00}(z) - \mathbf{1}) = 0\,. \tag{3.126}$$

Since $S_{00}(z)$ can be represented as a $n \times n$ matrix and is analytic then $f(z)$ is $\Omega_0$ and

$$\Sigma_0 \;=\; \{z \in \Omega_0\,;\, f(z) = 0\} \tag{3.127}$$

is either $\Omega_0$ or is a discrete set.

In conclusion if $f(z) \equiv 0$ in $\Omega_0$ then $S(z)\xi = \xi$ has nontrivial solution for all $z \in \Omega_0$ and so does $T(z)\xi = \xi$ and 1 belong to the spectrum of $T(z)$ for all $z \in \Omega_0$.

If $f(z) \not\equiv 0$ then $\Sigma_0 \;=\; \{z \in \Omega_0\,;\, f(z) = 0\}$ is a discrete set without accumulation points and if $z \notin \Sigma_0$ we have

$$\begin{aligned} (T(z) - \mathbf{1})^{-1} \;&=\; A(z)^{-1}(S(z) - \mathbf{1})^{-1} \\ &=\; A(z)^{-1} \begin{bmatrix} (S_{00} - \mathbf{1})^{-1} & (S_{00} - \mathbf{1})^{-1}S_{01} \\ 0 & -\mathbf{1} \end{bmatrix}. \end{aligned} \tag{3.128}$$

Since $(S_{00} - \mathbf{1})^{-1} = f(z)^{-1}C$ for some analytic matrix $C$, $(S_{00} - \mathbf{1})^{-1}$ is meromorph in $\Omega_0$ with poles on the zeros of $f(z)$ and residuum of the form

$$\mathrm{Res}(f(z)^{-1}) \begin{bmatrix} C & CS_{01} \\ 0 & 0 \end{bmatrix} \tag{3.129}$$

which is of finite rank. The same holds for $T(z)$. ∎

**Lemma 3.40.** *Let $T \in \mathcal{L}(H)$ be compact. Then for $\mu \neq 0$ the range $R(T - \mu\mathbf{1})$ is closed.*

*Proof.* Replacing $T$ by $z^{-1}T$ we need to show that $R(T - \mathbf{1})$ is closed. By (3.117) we have

$$(T - \mathbf{1}) = (S - \mathbf{1})A \tag{3.130}$$

and so $R(T - \mathbf{1}) = R(S - \mathbf{1})$. We also have

$$(S(z) - \mathbf{1})^{-1} = \begin{bmatrix} (S_{00} - \mathbf{1})^{-1} & (S_{00} - \mathbf{1})^{-1} S_{01} \\ 0 & -\mathbf{1} \end{bmatrix} \tag{3.131}$$

Let $\{\eta_n\}$ be a Cauchy sequence in $R(S - \mathbf{1})$, i.e.

$$\begin{bmatrix} \eta_{k0} \\ \eta_{k1} \end{bmatrix} = \begin{bmatrix} S_{00} - \mathbf{1} & S_{01} \\ 0 & -\mathbf{1} \end{bmatrix} \begin{bmatrix} \xi_{k0} \\ \xi_{k1} \end{bmatrix} = \begin{bmatrix} (S_{00} - \mathbf{1})\xi_{k0} + +S_{01}\xi_{k1} \\ -\xi_{k1} \end{bmatrix} \tag{3.132}$$

Therefore $\xi_{k1}$ is a Cauchy sequence and $\lim_k \xi_{k1} = \xi_1$ and $\lim_k S_{01}\xi_{k1} = S_{01}\xi_1$. Therefore $(S_{00} - \mathbf{1})\xi_{k0}$ is a Cauchy sequence. Since $\dim(H_0) < \infty$ then the range of $(S_{00} - \mathbf{1})$ is closed. ∎

**Lemma 3.41.** *Under the same assumptions as in Theorem 3.39 for $z \in \Omega$ we have*

$$\dim N(T(z) - \mathbf{1}) = \dim N(T^*(z) - \mathbf{1}) \,. \tag{3.133}$$

*Proof.* Since $T(z) - \mathbf{1} = (S(z) - \mathbf{1})A(z)$ we have $N(T(z) - \mathbf{1}) = A(z)^{-1}N(S(z) - \mathbf{1})$. Also we have

$$(T(z) - \mathbf{1})^* = A(z)^*(S(z) - \mathbf{1})^* \tag{3.134}$$

we have

$$\dim N(T(z)^* - \mathbf{1}) = \dim N(S^*(z) - \mathbf{1}) \,. \tag{3.135}$$

So it is enough to prove the theorem for $S$ instead of $T$. Since

$$S = \begin{bmatrix} S_{00} & S_{01} \\ 0 & 0 \end{bmatrix} \tag{3.136}$$

we have $N(S - \mathbf{1}) = N(S_{00} - \mathbf{1})$. Since

$$S^* - \mathbf{1} = \begin{bmatrix} S_{00} - \mathbf{1} & 0 \\ S_{01}^* & -\mathbf{1} \end{bmatrix} \tag{3.137}$$

we have

$$0 = (S^* - \mathbf{1}) \begin{bmatrix} \xi_0 \\ \xi_1 \end{bmatrix} = \begin{bmatrix} (S_{00}^* - \mathbf{1})\xi_0 \\ S_{01}^*\xi_0 \end{bmatrix} \tag{3.138}$$

if and only if

$$\xi_0 \in N(S_{00}^* - \mathbf{1}) \quad \text{and } \xi_1 = S_{01}^*\xi_0 \,. \tag{3.139}$$

Thus we have

$$\dim N(S(z) - \mathbf{1}) = \dim N(S_{00} - \mathbf{1}) = \dim N(S_{00}^* - \mathbf{1}) = \dim N(S(z)^* - \mathbf{1}) \,. \tag{3.140}$$

∎

**Lemma 3.42.** *For $T \in \mathcal{L}(H)$ the following are equivalent.*

1. *$T$ is compact.*
2. *$T^*T$ is compact.*
3. *$TT^*$ is compact.*
4. *$T^*$ is compact.*

*Proof.* 1. If $T$ is compact, then $T^* \in \mathcal{L}(H)$ and so $T^*T$ and $TT^*$ are compact.

2. If $T^*T$ is compact then for any bounded sequence $\xi_n$ there exists a subsequence $\xi'_k = \xi_{n_k}$ such that $T^*T\xi'_k$ converges. Then we have

$$\begin{aligned} \|T(\xi'_k - \xi'_l)\|^2 &= (T(\xi'_k - \xi'_l)\,,\, T(\xi'_k - \xi'_l)) \;=\; ((\xi'_k - \xi'_l)\,,\, T^*T(\xi'_k - \xi'_l)) \\ &= \|\xi'_k - \xi'_l\|\|T^*T(\xi'_k - \xi'_l)\| \leq 2M\|T^*T(\xi'_k - \xi'_l)\|\,. \end{aligned} \tag{3.141}$$

and so $T\xi'_k$ converges and so $T$ is compact.

3. Finally we have $T^*T$ compact $\Rightarrow T$ compact $\Rightarrow TT^*$ compact $\Rightarrow T^*$ compact. ■

We are now in the position to prove

**Theorem 3.43. (Riesz-Schauder)** *Let $T \in \mathcal{L}(H)$ compact. Then we have*

1. *Every spectral point $\mu \in \sigma(T)$, $\mu \neq 0$ is an eigenvalue of $T$.*
2. *The point spectrum $\sigma_p(T)$ is countable, $\mu = 0$ is the only possible accumulation point of $\sigma_p(T)$.*

*Proof.* Let us set $T(z) = zT$. Then is $zT$ analytic in $\mathbb{C}$ with values $\mathcal{C}(H)$. The conditions of Theorem 3.39 are satisfied with $\Omega = \mathbb{C}$. The first case cannot occur since 1 would be a spectral value of $zT$ for all $z$, i.e. $1/z$ would be a spectral value of $T$ for all $z \in \mathbb{C}$ and thus $\sigma(T) = \mathbb{C}$ which is not possible for a bounded operator.

Therefore the set

$$\Sigma \;=\; \{z\,;\, 1\text{is not a singular value of } T(z)\} \tag{3.142}$$

is a discrete set without accumulation point. Furthermore we have

$$z \in \Sigma \Leftrightarrow zT - \mathbf{1} \text{ not bijective } \Leftrightarrow T - \frac{1}{z}\mathbf{1} \text{ not bijective} \tag{3.143}$$

So the set

$$\widehat{\sigma}(T) \;=\; \left\{\mu\,;\, \mu = \frac{1}{z} \text{ with } z \in \Sigma\right\} \tag{3.144}$$

is contained in $\sigma_p(T)$ (and is is equal to it if $0 \in \sigma_p(T)$). ■

**Theorem 3.44. (Fredholm alternative)** *Let $T \in \mathcal{L}(H)$ compact. Then we have for $T_\mu = T - \mu\mathbf{1}$, $\mu \neq 0$ of one the two options*

*1.*

$$\left.\begin{array}{l} T_\mu \xi = \eta \\ T^*_\mu \xi = \eta \end{array}\right\} \text{has a 1d dimensional solution space} \tag{3.145}$$

$$\left.\begin{array}{l} T_\mu \xi = 0 \\ T^*_\mu \xi = 0 \end{array}\right\} \text{has only the trivial solution} \tag{3.146}$$

*2.*

$$\begin{array}{ll} T_\mu \xi = 0 \text{ has a solution space } M_\mu & (\dim(M_\mu) < \infty) \\ T^*_\mu \xi = 0 \text{ has a solution space } M^*_\mu & (\dim(M^*_\mu) = \dim(M_\mu)) \end{array}$$

$$\begin{array}{l} T_\mu \xi = \eta \text{ has a solution if and only if } \eta \perp M^*_\mu \\ T^*_\mu \xi = \eta \text{ has a solution if and only if } \eta \perp M_\mu \end{array}$$

*Proof.* Let $\mu \neq 0$. If $\mu \notin \sigma_p(T)$ then 1. holds and $\xi = R_\mu(T)\eta$ is the unique solution of $T_\mu \xi = \eta$. The rest is obvious.

If $\mu \in \sigma_p(T)$ then $M_\mu = N(T_\mu)$ is the eigenspace of $T$ for the eigenvalue $\mu$ and so $\dim(M_\mu) < \infty$ and $\dim(M_\mu) = \dim(M^*_\mu)$ by Lemma 3.41. It remains to show that $T_\mu \xi = \eta$ has a solution if and only if $\eta \in N(T^*_\mu)^\perp$.

If $T_\mu \xi = \eta$ then for $\zeta \in N(T^*_\mu)$ we have

$$(\eta, \zeta) \;=\; (T_\mu \xi, \zeta) \;=\; (\xi, T^*_\mu \zeta) \;=\; 0 \tag{3.147}$$

and so $\eta \perp N(T^*_\mu)$.

Conversely we show that if $T\mu\xi = \eta$ has no solution then $\eta \notin N(T^*_\mu)^\perp$. By Lemma 3.40, $R(T_\mu) \equiv H_0$ is closed. If $T_\mu \xi = \eta$ has no solution then there exists $\delta > 0$ such that

$$\inf_{\beta \in H_0} \|\beta - \eta\| = \inf_{\xi \in H} \|T_\mu \xi - \eta\| \;=\; \delta \tag{3.148}$$

By Corollary chb3 to Hahn-Banach theorem the exists $\lambda \in H'$ such that $\lambda(\eta) = \delta$ and $\lambda(T_\mu \xi) = 0$ for all $\xi \in H$. By Riesz representation theorem there exists $\zeta \in H$ such that

$$(\eta, \zeta) = \delta \qquad 0 = (T_\mu \xi, \zeta) = (\xi, T^*_\mu \zeta) \text{ for all } \xi \in H\,. \tag{3.149}$$

So $T^*_\mu \zeta = 0$ and so $\zeta \in N(T^*_\mu)$. Since $(\eta, \zeta) = \delta > 0$ we have on one hand $\zeta \neq 0$ and, on the other hand, that $\eta \in N(T^*_\mu)^\perp$. ∎

Before we pursue we we recall a few facts and definitions form Hilbert space theory.

- If $T = T^*$ then the eigenvalues are of $T$ are real and the eigenvectors for distinct eigenvalues are orthogonal.
- $T$ is called a positive operator if the quadratic form $(T\xi, \xi) \geq 0$ (it is important to work in complex vector spaces here!) and for positive operator every eigenvalue is nonnegative.
- If $T = T^*$ from 2.56 we know that $\|T^*T\| = \|T\|^2$ and so $\|T^2\| = \|T\|^2$. By induction $\|T^{2n}\| = \|T\|^{2n}$ and thus $r(T) = \inf_n \|T^n\|^{1/n} \;=\; \|T\|$.

- $T$ is called *normal* if $T^*T = T^*$.

**Theorem 3.45. (Hilbert-Schmidt)** *Let $T \in \mathcal{L}(H)$ be compact and self-adjoint. Then there exists an orthonormal basis $\{\eta_i\}$ of $H$ such that $T\eta_i = \lambda_i \eta_i$ with $\lim_i \lambda_i = 0$. If $\mu_1, \mu_2, \cdots$ are the pairwise distinct eigenvalues of $T$ and $E_1, E_2, \cdots$ the corresponding eigenspace of $T$. Then we have*

$$H = E_1 \oplus E_2 \oplus \cdots \qquad T = \sum_i \lambda P_i \tag{3.150}$$

*where $P_i$ is the orthogonal projection on $H_i$.*

*Proof.* Eigenspaces for different eigenvalues are orthogonal, so we can choose in each eigenspace an orthonormal basis and this gives an orthonormal sequence $\{\eta_i\}$. Let $H_1$ be the closed linear span of $\{eta_i\}$. Clearly $T$ maps $H_1$ into itself. Let us write $H = H_1 \oplus H_1^\perp$. and let us assume $H_1^\perp$ is not trivial. Since $T = T^*$ we have $T$ maps $H_1^\perp$ into itself. Let $T = T|_{H_1}$ and $T_2 = T|_{H_1 \perp}$. The operator $T_2$ is self-adjoint and compact.

From Riesz-Schauder theorem if $\mu \in \sigma(T_2)$ with $\mu \neq 0$ then $\mu \in \sigma_p(T_2)$. But then $\mu$ is also an eigenvalue of $T$ which is impossible. Therefore $\sigma(T_2) = \{0\} = \|T_2\|$ and thus $T_2 = 0$. So $H_1^\perp$ is the eigenspace of $T$ for the eigenvalue $0$. But since all the eiegenspace for $T$ are already in $H_1$ we have $H_1^\perp = 0$. ■

We have previously developed a functional calculus to define $f(T)$ for general bounded operators and were able to use functions $f$ which are analytic in an open set containing $\sigma(T)$. For a compact self-adjoint operator $T$ we can define $f(T)$ for *every* function defined on the spectrum of $A$.

**Theorem 3.46.** *(Functional Calculus) Let $T$ be a compact self-adjoint operator and $f$ a bounded complex-valued function defined on $\sigma(T)$. To such an $f$ we can define $f(T)$ such that*

1. *If $f \equiv 1$ then $f(T) = 1$.*
2. *If $f(x) = x$ then $f(T) = T$.*
3. *The map $f \mapsto f(T)$ is an isomorphism of the ring of bounded functions on $\sigma(A)$ into the algebra $\mathcal{L}(H)$.*
4. *The isomorphism is isometric:*
$$\|f(T)\| = \sup_{\mu \in \sigma(T)} |f(\mu)\| \tag{3.151}$$
5. *If $f$ is real-valued, $f(T)$ is symmetric.*
6. *If $f$ is positive on $\sigma(T)$, then $f(T)$ is positive.*

*Proof.* The proof is shorter than the statement. From Hlbert-Schmidt theorem we can pick an o.n.b. basis $\{\eta_k\}$ of eigenvectors such that $T\eta_k = \mu_k \eta_k$. If $\xi \in H$ has the decompostion $\xi = \sum_{x_i} \eta_i$ then define $f(T)$ by

$$f(T)\xi = f(\mu_k) x_k \eta_k \,. \tag{3.152}$$

The theorem is now obvious. ■

We also have

**Theorem 3.47.** *Suppose $T_1, \cdots T_n$ are bounded operators in $\mathcal{L}(H)$ such that*

1. *$T_k$ is self-adjoint for $k = 1, \cdots, n$.*
2. *$T_1$ is compact.*
3. *$T_k T_l = T_k T_l$ for all $k, l$.*

*Then there exists an orthonormal basis $\{\xi_i\}$ of $H$ such that $\xi_i$ is an eigenvector for all $T_k$.*

*Proof.* Let $S_i$ be the eigenspaces of $T_1$ for the eigenvalue $\mu_i$. Then by Hilbert-Schmidt theorem we have

$$H = S_1 \oplus S_2 \oplus \cdots \tag{3.153}$$

Then the subspace $S_i$ is invariant under $T_k$ since if $\xi \in S_i$m, then $T\xi \,=\, \mu_i \xi$ then

$$T_1 T_k \xi \,=\, T_k T_1 \xi \,=\, \mu_i T_k \xi \tag{3.154}$$

and thus $T_k \xi \in S_i$. The restriction of $T_2$ on $S_n$ is finite-dimensional and thus we can choose a basis of $S_i$ such that the basis elements are eigenvectors for $T_1$ and $T_2$ and we can repeat the arguments for $T_3$, etc.... Since the $S_n$ are finite dimensional the procedure will stop after a finite number of steps. ∎

**Corollary 3.48.** *If $T$ is a normal operator then there exists an o.n.b $\{\xi_i\}$ of $H$ such that*

$$T\xi \,=\, \sum_{n=1}^{\infty} \mu_n (\xi, \xi_n) \xi_n \tag{3.155}$$

*Proof.* We write

$$T \,=\, \underbrace{\frac{T + T^*}{2}}_{\equiv R} + \underbrace{\frac{T - T^*}{2}}_{\equiv J} \,. \tag{3.156}$$

Then $T$ and $J$ are compact, $T$ is self-adjoint and $iJ$ is self-adjoint. ∎

Finally we consider non self-adjoint operators.

**Theorem 3.49.** *Let Let $T \in \mathcal{L}(H)$ compact. Then there exists 2 sequences $\{\eta_i\}_{i=1}^N$ and $\{\zeta_i\}_{i=1}^N$ in $H$ (with $N$ finite or $\infty$) and a sequence $\{\nu_i\}$ of non-negative numbers such that*

$$T \,=\, \sum_{i=1}^{N} \nu_i (\,\cdot\,,\, \alpha_i) \beta_i \tag{3.157}$$

*Proof.* If $T$ is compact then $T^*T$ is compact, self-adjoint, and positive. So there exists an orthonormal sequence $\{\eta_i\}$ such that

$$T^*T\eta_i \,=\, \mu_i \eta_i \tag{3.158}$$

with $\mu_i \geq 0$ and $\lim_i \mu_i = 0$. Moreover $T^*T$ restricted on the orthogonal complement to the span of the $\eta_i$ is equal to $0$. Now let

$$\nu_i = \sqrt{\mu_i} \tag{3.159}$$

and set

$$\zeta_i \;=\; \frac{1}{\nu_i} T\eta_i\,. \tag{3.160}$$

We have then

$$(\zeta_i, \zeta_j) \;=\; \frac{1}{\nu_i\nu_j}(T\eta_i, T\eta_j) \;=\; \frac{1}{\nu_i\nu_j}(\eta_i, T^*T\eta_j) \;=\; \frac{1}{\nu_i\nu_j}\mu_j(\eta_i, \eta_j) \;=\; \delta_{ij} \tag{3.161}$$

and so the $\beta_i$ are orthonormal. For any $\xi \in H$ we have

$$\xi \;=\; \sum_{i=1}^{N} (\xi, \eta_i)\eta_i + \xi^{\perp}\,. \tag{3.162}$$

where $T^*T\xi^{\perp} = 0$. This implies that

$$0 = (\xi^{\perp}, T^*T\xi^{\perp}) \;=\; (T\xi^{\perp}, T\xi^{\perp}) \;=\; \|T\xi^{\perp}\|^2 \tag{3.163}$$

and so $T\xi^{\perp} = 0$. So we have

$$T\xi \;=\; \sum_{i=1}^{N} (\xi, \eta_i) T\eta_i \;=\; \sum_{i=1}^{N} \nu_i (\xi, \eta_i)\zeta_i\,, \tag{3.164}$$

as claimed. ■

## 3.4 Applications

## 3.5 Exercises

**Exercise 21.** 1. Show that the Fourier transform $T$ maps the space of function of the form $p(x)e^{-\pi x^2}$ where $P$ is a polynomial of degree $\leq n$ into itself.

2. Find the eigenfunctions of $T$.
*Hint:* With our conventions the Fourier transform of $e^{-\pi x^2}$ is itself.

**Exercise 22.** Prove Lemma 3.18

**Exercise 23.** Show that if $T$ is a normal operator (i.e., $T^*T = TT^*$) then $r(T) \;=\; \|T\|$.

**Exercise 24.** For $T \in \mathcal{L}(V)$ let us define $\tau(T)$ by

$$\tau(T) \;=\; \max_{\mu \in \sigma(T)} \mathrm{Re}\,(\mu) \tag{3.165}$$

i.e. the spectrum of $T$ is contained in the half-plane $\{\mathrm{Re}\,(z) \leq \tau(T)\}$. Show that we have

$$\tau(T) \;=\; \lim_{n \to \infty} \frac{\ln \|e^{nT}\|}{n}\,. \tag{3.166}$$

**Exercise 25.** Suppose $T$ is bounded operator on a Hilbert space $H$ such that $T\xi_k = \lambda_k \xi_k$ with $\xi_k$ an orthonormal basis of $H$. Show that $T$ is compact if and only if $\lambda_k \to 0$.

**Exercise 26.** Show that the operator $Tf(t) = \int_0^t f(s)\,ds$ on $C[0,1]$ is compact and has no eigenvalues. Give an example of an operator $T$ on $l^p$ which is compact and has no eigenvalue.

**Exercise 27.** Consider the integral equation

$$f(t) - \mu \int_0^1 k(t,s) f(s)\,ds = g(t) \tag{3.167}$$

with a continuous $k(t,s)$ and denote $Tf(t) = \int_0^1 k(t,s)f(s)\,ds$.

1. Let $r(T)$ be the spectral radius of $T$ on $C[0,1]$. Show that if $|\mu| < \frac{1}{r(T)}$ then (3.167) has a unique solution which can be written in the form

   $$f(t) = g(t) + \int_0^1 r(t,s) g(s)\,ds \tag{3.168}$$

   (the kernel $(r(t,s)$ is called the resolvent kernel). ***Hint:*** Neumann series.

2. Compute the resolvent kernel and the solution for the integral equation

   $$f(t) = \frac{1}{2}\int_0^1 e^{t-s} f(s)\,ds + g(s)\,. \tag{3.169}$$

**Exercise 28.** Consider a kernel $k(t,s) = \sum_{j=1}^n a_j(t)b)j(s)$. Without loss of generality we may assume that the $a_j$ and $b_j$ are linearly independent. Show that if the equation

$$f(t) - \mu \int_0^1 k(t,s) f(s)\,ds = g(t) \tag{3.170}$$

has a solution then it must be of the form

$$f(t) = g(t) + \mu \sum_{j=1}^n c_j a_j(t),, \qquad c_j = \int_0^1 b_j(s) x_j(s)\,ds \tag{3.171}$$

and the constants $c_j$ must satisfy the linear equations

$$c_j - \mu \sum_{j=1}^n a_{jk} c_k = y_i\,, \tag{3.172}$$

with

$$a_{jk} = \int_0^1 b_j(s) a_k(s)\,ds\,, \quad y_j = \int_0^1 b_j(s) y(s)\,ds \tag{3.173}$$

**Exercise 29.** Consider the equation

$$f(t) - \mu \int_0^1 (s+t) x(t) dt = g(s) \tag{3.174}$$

1. Use the previous exercise to solve this equation if $\mu^2 + 12\mu - 12 \neq 0$.

2. Find the eigenvalues and eigenfunctions.

**Exercise 30.** Let $H$ be a Hilbert space and $T$ a self-adjoint compact operator with eigenvalues $\mu_k$ with eigenvectors $\xi_k$. Consider the equation

$$(\lambda \mathbf{1} - T)\xi \;=\; \eta \tag{3.175}$$

for a given $\lambda \neq 0$ and $\eta \in H$ given. Show that this equation has a solution if and only if $\eta$ is orthogonal to $N(\lambda \mathbf{1} - T)$ and that the set of all solution general solution is given by

$$\xi = \xi_0 + \zeta \tag{3.176}$$

where $\zeta \in N(\lambda \mathbf{1} - T)$ and

$$\xi_0 \;=\; \frac{1}{\lambda}\eta + \frac{1}{\lambda}\sum_{\mu_n \neq \lambda} \frac{\mu_n}{\mu_n - \lambda}(\eta\,,\,\xi_n)\xi_n \tag{3.177}$$

*Hint:* Consider separately the case where $\lambda$ is an eigenvalue or not. Do not forget the show that $\xi_0$ is well-defined.

# Bibliography


[1] Peter Lax: *Functional Analysis*, Wiley, ISBN: 978-0471556046

[2] Michael Reed and Barry Simon, *Methods of Modern Mathematical Physics, Vol I: Functional Analysis*, Academic Press, ISBN: 978-0125850506

[3] Martin Schechter *Principles of Functional Analysis*, Graduate Studies in Mathematics, American Mathematical Society, ISBN: 978-0821828953