PEOPLES' FRIENDSHIP UNIVERSITY OF RUSSIA

FACULTY OF PHYSICAL-MATHEMATICAL AND NATURAL SCIENCES

S.M. Nikol'skiĭ Mathematical Institute

Functional Analysis

V.I. Burenkov

2021

Contents

Functional Analysis

Functional analysis is an abstract branch of mathematics that originated from classical mathematical analysis. In functional analysis, from a general point of view, mappings (= functions = operators)

$$f : X \to Y$$

are studied acting from one general space (= set) X to another one Y. Spaces X and Y can consist of objects of arbitrary nature. However, in this course main attention is paid to the case in which X and Y are infinite-dimensional linear spaces of functions and f is a linear operator.

Functional analysis has many applications in various branches of mathematics: in the theory of differential equations, approximation theory, mathematical physics, theoretical physics, calculus of variations, optimal control and optimization, in probability theory, etc.

1. Metric spaces

1.1 Definition of a metric space

The notion of a distance between points is one of the basic notions of geometry. The notion of a metric is a generalization of this notion which can be applied to various, not necessarily geometric objects, say to sequences and functions.

Definition 1.1.1. *A metric space is a pair (X, d), where X is a set (whose elements are called points) and d is a metric on X (= distance function on X), that is, a function on $X \times X$, whose values are non-negative numbers,*[1] *briefly*

$$d : X \times X \to \mathbb{R}_+,$$

such that for all $x, y, z \in X$

1) $d(x, y) = 0$ *if and only if* $x = y$,

2) $d(x, y) = d(y, x)$ *(symmetry),*

3) $d(x, y) \leq d(x, z) + d(z, y)$ *(triangle inequality).*

Instead of (X, d) we shall often simply write X if there is no danger of confusion.
If instead of 1) we have only

1′) $d(x, y) = 0$ *if* $x = y$,

then the space (X, d) is said to be a semi-metric space.

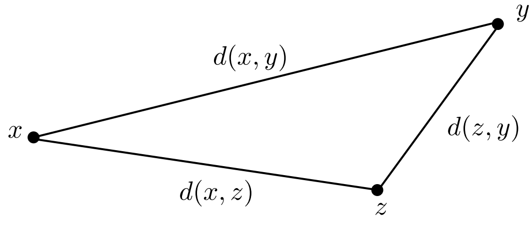


Figure 1.1 Triangle inequality in the plane

Definition 1.1.2. *Given a point $x \in X$ and $r > 0$, let*

$$\begin{aligned} B_d(x, r) &= \{y \in X : d(y, x) < r\} \quad \textit{(open ball)}, \\ \bar{B}_d(x, r) &= \{y \in X : d(y, x) \leq r\} \quad \textit{(closed ball)}, \\ S_d(x, r) &= \{y \in X : d(y, x) = r\} \quad \textit{(sphere)}. \end{aligned}$$

(x is the center and r is the radius.)

[1] $A \times B$ is the Cartesian product of sets A and B, that is, $A \times B$ is the set of all ordered pairs (a, b) where $a \in A,\ b \in B$. $\mathbb{R}$ is the set of all real numbers, $\mathbb{R}_+$ is the set of all non-negative real numbers.

Definition 1.1.3. *Given a sequence of points $x_k \in X,\ k \in \mathbb{N}$, and a point $x \in X$ it is said that*

$$\lim_{k\to\infty} x_k = x \quad \text{in} \quad X$$

if

$$\lim_{k\to\infty} d(x_k, x) = 0.$$

1.2 Examples

Example 1.1. $X = \mathbb{R}^2,\ x, y \in X,\ x = (x_1, x_2),\ y = (y_1, y_2),\ x_1, x_2, y_1, y_2 \in \mathbb{R}$,

$$d_2(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$$

(standard distance between the points x and y in the plane). $(\mathbb{R}^2, d_2)$ is a metric space.

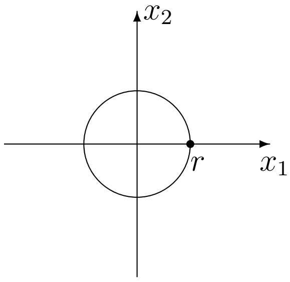


Figure 1.2 $B_{d_2}(0, r)$

Convergence. Let $x_k = (x_{1k}, x_{2k}),\ k \in \mathbb{N},\ a = (a_1, a_2)$. Then,

$$\lim_{k\to\infty} x_k = a \quad \text{in} \quad (\mathbb{R}^2, d_2) \iff \lim_{k\to\infty} x_{1k} = a_1, \quad \lim_{k\to\infty} x_{2k} = a_2.$$

Let $X = \mathbb{R}^2,\ a, b > 0$ and

$$d(x, y) = \sqrt{\frac{(x_1 - y_1)^2}{a^2} + \frac{(x_2 - y_2)^2}{b^2}}.$$

Exercise. Prove that $d(x, y)$ is a metric.

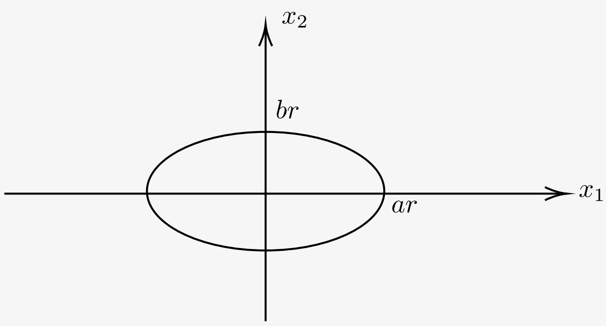


Figure 1.3 $B_d(0, r)$

Example 1.2. $X = \mathbb{R}^2$,

$$d_1(x, y) = |x_1 - y_1| + |x_2 - y_2|.$$

$(\mathbb{R}^2, d_1)$ is a metric space.

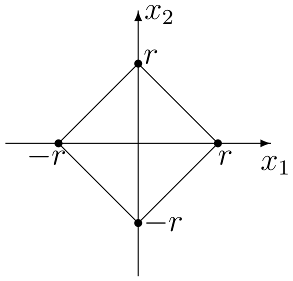


Figure 1.4 $B_{d_1}(0, r)$

Example 1.3. $X = \mathbb{R}^2$,

$$d_\infty(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\}.$$

$(\mathbb{R}^2, d_\infty)$ is a metric space.

Properties of a metric space.

1) Obvious. 2) Obvious. 3) For $z = (z_1, z_2)$,

$$\begin{aligned} d_\infty(x, y) = \max\{|x_1 - y_1|, |x_2 - y_2|\} &= \max\{|x_1 - z_1 + z_1 - y_1|, |x_2 - z_2 + z_2 - y_2|\} \\ &\le \max\{|x_1 - z_1| + |z_1 - y_1|, |x_2 - z_2| + |z_2 - y_2|\} \\ &\le \max\{|x_1 - z_1|, |x_2 - z_2|\} + \max\{|z_1 - y_1|, |z_2 - y_2|\} \\ &= d_\infty(x, z) + d_\infty(z, y). \end{aligned}$$

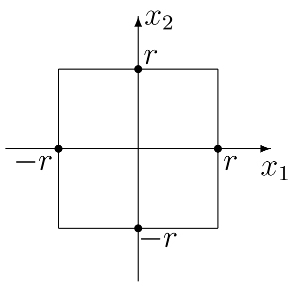


Figure 1.5 $B_{d_\infty}(0, r)$

Convergence.

$$\lim_{k\to\infty} x_k = a \iff \lim_{k\to\infty} x_{1k} = a_1, \quad \lim_{k\to\infty} x_{2k} = a_2.$$

Example 1.4a. $X = \mathbb{C}^n,\ 1 \le p \le \infty,\ x = (x_1, \dots, x_n),\ y = (y_1, \dots, y_n),\ x_i, y_i \in \mathbb{C}$.

$$d_p(x, y) = \begin{cases} \left(\sum\limits_{i=1}^{n} |x_i - y_i|^p \right)^{\frac{1}{p}} & \text{if } 1 \le p < \infty, \\ \max\limits_{i=1,\dots,n} |x_i - y_i| & \text{if } p = \infty. \end{cases}$$

$(\mathbb{C}^n, d_p)$ is a metric space.

Properties of a metric space.

1) Obvious. 2) Obvious. 3) For $1 \le p < \infty$,

$$d_p(x,y) = \Big(\sum_{i=1}^{n} |x_i - y_i|^p\Big)^{\frac{1}{p}} \le \Big(\sum_{i=1}^{n} (|x_i - z_i| + |z_i - y_i|)^p\Big)^{\frac{1}{p}}$$
$$\le \{\text{Minkowski's inequality for finite sequences}\} \le \Big(\sum_{i=1}^{n} |x_i - z_i|^p\Big)^{\frac{1}{p}} + \Big(\sum_{i=1}^{n} |z_i - y_i|^p\Big)^{\frac{1}{p}}$$
$$= d_p(x,z) + d_p(z,y).$$

Convergence.

Let $x_k = (x_{1k}, \ldots, x_{nk}),\ a = (a_1, \ldots, a_n)$. Then,

$$\lim_{k\to\infty} x_k = a \quad \text{in} \quad (\mathbb{C}^n, d_p) \iff \lim_{k\to\infty} d_p(x_k, a) = 0$$
$$\iff \lim_{k\to\infty} \Big(\sum_{i=1}^{n} |x_{ik} - a_i|^p\Big)^{\frac{1}{p}} = 0 \iff \lim_{k\to\infty} x_{ik} = a_i, \quad i = 1, \ldots, n.$$

Indeed,

$$\Big(\sum_{i=1}^{n} |x_{ik} - a_i|^p\Big)^{\frac{1}{p}} \ge |x_{ik} - a_i| \ge 0$$
$$\Longrightarrow \lim_{k\to\infty} |x_{ik} - a_i| = \lim_{k\to\infty} \Big(\sum_{i=1}^{n} |x_{ik} - a_i|^p\Big)^{\frac{1}{p}} = 0$$

and

$$\lim_{k\to\infty} \Big(\sum_{i=1}^{n} |x_{ik} - a_i|^p\Big)^{\frac{1}{p}} = \Big(\sum_{i=1}^{n} \Big(\lim_{k\to\infty} |x_{ik} - a_i|\Big)^p\Big)^{\frac{1}{p}} = 0.$$

Exercise. Prove that $\lim\limits_{p\to+\infty} d_p(x,y) = d_\infty(x,y)$.

Example 1.4b. $X = \mathbb{C}^n,\ 1 \le p \le \infty,\ 1 \le m < n.$

$$d_p(x,y) = \begin{cases} \Big(\sum\limits_{i=1}^{m} |x_i - y_i|^p\Big)^{\frac{1}{p}} & \text{if } 1 \le p < \infty, \\ \max\limits_{i=1,\ldots,m} |x_i - y_i| & \text{if } p = \infty. \end{cases}$$

$(\mathbb{C}^n, d_p)$ is a semi-metric space.

Properties of a semi-metric space.

1′) Obvious. 2) Obvious. 3) As in Example 1.4a.

1) Fails. $d_p(x,y) = 0 \iff x_i = y_i,\ i = 1, \ldots, m$ but it could be that $x_i \ne y_i,\ i = m+1, \ldots, n$, say, for $x = (0, \ldots, 0,)\ y = (\underbrace{0, \ldots, 0}_{m}, \underbrace{1, \ldots, 1}_{n-m}),\ d_p(x,y) = 0.$

Example 1.5a. $X = \mathbb{C}^n,\ 0 < p < 1$,

$$d_p(x,y) = \Big(\sum_{i=1}^{n} |x_i - y_i|^p\Big)^{\frac{1}{p}}.$$

d_p is not a metric.

Properties of a metric space.

1) Obvious. 2) Obvious. 3) Fails. Counter-example: for $x = (1,0,\dots,0),\ y = (0,1,0,\dots,0),\ z = (0,\dots,0)$,

$$d_p(x,y) = 2^{\frac{1}{p}} > d_p(x,z) + d_p(z,y) = 2.$$

Example 1.5b. $X = \mathbb{C}^n,\ 0 < p < 1$,

$$\tilde{d}_p(x,y) = \sum_{i=1}^{n} |x_i - y_i|^p.$$

$\tilde{d}_p$ is a metric.

Properties of a metric space.

1) Obvious. 2) Obvious.

$$3)\ \tilde{d}_p(x,y) = \sum_{i=1}^{n} |x_i - y_i|^p \le \sum_{i=1}^{n} (|x_i - z_i| + |z_i - y_i|)^p$$
$$\le \sum_{i=1}^{n} (|x_i - z_i|^p + |z_i - y_i|^p) = \tilde{d}_p(x,z) + \tilde{d}_p(z,y).$$

The following numerical inequality

$$(a+b)^p \le a^p + b^p, \quad a, b \ge 0,\ 0 < p < 1$$

was used.

Exercise. Prove this inequality. Hint. For $b > 0$, divide by b and set $x = \frac{a}{b}$.

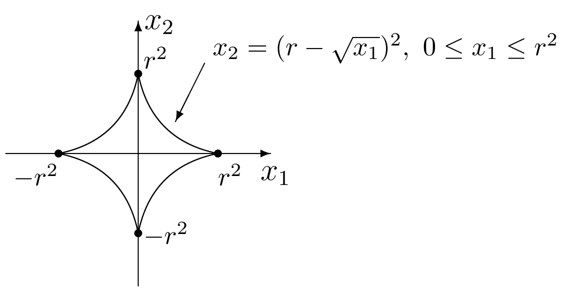


Figure 1.6 $B_{d_{\frac{1}{2}}}(0,r)$

As in Example 1.4a,

$$\lim_{k\to\infty} x_k = a \quad \text{in} \quad (\mathbb{C}^n, \tilde{d}_p) \iff \lim_{k\to\infty} x_{ik} = a_i, \quad i = 1,\dots,n.$$

So the convergence in $(\mathbb{R}^n, d_p)$ for any $0 < p \le \infty$ is the same, namely, it is the *co-ordinate convergence.*

Example 1.6 (Space ℓ_p). Let $0 < p \leq \infty$ and $x = (x_1, \ldots, x_i, \ldots)$, where $x_i \in \mathbb{C}$. Then,

$$x \in \ell_p \implies \begin{cases} \sum\limits_{i=1}^{\infty} |x_i|^p < \infty & \text{if } p < \infty \\ \sup\limits_{i \in \mathbb{N}} |x_i| < \infty & \text{if } p = \infty \text{ (bounded sequence).} \end{cases}$$

Let for $x, y \in \ell_p$

$$d_p(x,y) = \begin{cases} \sum\limits_{i=1}^{\infty} |x_i - y_i|^p & \text{if } 0 < p < 1 \\ \Big(\sum\limits_{i=1}^{\infty} |x_i - y_i|^p \Big)^{\frac{1}{p}} & \text{if } 1 \leq p < \infty \\ \sup\limits_{i \in \mathbb{N}} |x_i = y_i| & \text{if } p = \infty. \end{cases}$$

For all $0 < p \leq \infty$, $\ell_p \equiv (\ell_p, d_p)$ is a metric space. This is proved as in Examples 4, 5.
Convergence.
Let $x_k = (x_{1k}, \ldots, x_{ik}, \ldots)$, $a = (a_1, \ldots, a_i, \ldots) \in \ell_p$. Then,

$$\lim_{k \to \infty} x_k = a \quad \text{in} \quad \ell_p \implies \lim_{k \to \infty} x_{ik} = a_i, \quad i \in \mathbb{N}.$$

However, $\lim\limits_{k \to \infty} x_{ik} = a_i$, $i \in \mathbb{N} \;\not\Longrightarrow\; \lim\limits_{k \to \infty} x_k = a$ in ℓ_p.
Counter-example: Let $x_k = (0, \ldots, 0, k, 0, \ldots)$. Then, $\lim\limits_{k \to \infty} x_{ik} = 0$, $i \in \mathbb{N}$, but

$$d_p(x_k, 0) = \begin{cases} k^p & \text{if } 0 < p < 1 \\ k & \text{if } 1 \leq p \leq \infty, \end{cases}$$

hence,

$$\lim_{k \to \infty} d_p(x_k, 0) = \infty$$

and the equality $\lim\limits_{k \to \infty} x_k = 0$ does not hold in ℓ_p.

Example 1.7 (Space $C[a,b]$). Let $-\infty < a < b < \infty$. Then the space $C[a,b]$ is the space of all continuous complex-valued functions defined on $[a,b]$.
Let for $f, g \in C[a,b]$

$$d(f,g) = \max_{a \leq x \leq b} |f(x) - g(x)|.$$

Then, $C[a,b] \equiv (C[a,b], d)$ is a metric space.

Properties of a metric space.

1) Obvious. 2) Obvious.

$$\begin{aligned} \text{3) } d(f,g) &= \max_{a \leq x \leq b} |f(x) - g(x)| \leq \max_{a \leq x \leq b} (|f(x) - h(x)| + |h(x) - g(x)|) \\ &\leq \max_{a \leq x \leq b} |f(x) - h(x)| + \max_{a \leq x \leq b} |h(x) - g(x)|) = d(f,h) + d(h,g). \end{aligned}$$

Convergence.
Let $f_k \in C[a,b]$, $f \in C[a,b]$. Then,

$$\lim_{k \to \infty} f_k = f \quad \text{in} \quad C[a,b] \iff \lim_{k \to \infty} d(f_k, f) = 0 \iff \lim_{k \to \infty} \max_{a \leq x \leq b} |f_k(x) - f(x)| = 0$$

$$\iff \forall \varepsilon > 0 \; \exists k_0 \in \mathbb{N} : \forall k \geq k_0 \quad \max_{a \leq x \leq b} |f_k(x) - f(x)| < \varepsilon$$

$$\iff |f_k(x) - f(x)| < \varepsilon \quad \forall x \in [a,b].$$

So, the convergence in $C[a,b]$ is the uniform convergence.

Example 1.8 (Space $C_b(\Omega)$). Let $\Omega \subset \mathbb{R}^n$. A point $x \in \Omega$ is called an *isolated point* if

$$\exists\, r > 0 : \quad B(x, r) \cap \Omega = \{x\}. \tag{1.1}$$

A point $x \in \mathbb{R}^n$ is called a *limit point* of Ω if

$$\forall\, r > 0, \quad B(x, r) \cap \Omega \setminus \{x\} \neq \emptyset.$$

Here, $B(x, r)$ is the standard ball in $\mathbb{R}^n$, that is, with respect to the metric

$$d(x, y) = \Big(\sum_{i=1}^{n} |x_i - y_i|^2\Big)^{\frac{1}{2}}.$$

A function $f : \Omega \to \mathbb{C}$ is said to be continuous at a point $x \in \Omega$ if $\forall\, \varepsilon > 0\ \exists\, \delta > 0$ such that

$$|f(y) - f(x)| < \varepsilon \quad \forall\, y \in B(x, \delta) \cap \Omega. \tag{1.2}$$

If x is an isolated point of Ω, then any function $f : \Omega \to \mathbb{C}$ is continuous at x. Indeed, if r is as in (1.1) and $0 < \delta \leq r$, then the only $y \in B(x, \delta)$ is $y = x$ and $|f(x) - f(x)| = 0 < \varepsilon$.
If $x \in \Omega$ is a limit point, then (1.2) means that

$$\lim_{\substack{y \to x \\ y \in \Omega}} f(y) = f(x).$$

A function $f : \Omega \to \mathbb{C}$ is said to be continuous on Ω if it is continuous at any point $x \in \Omega$. The space of all functions $f : \Omega \to \mathbb{C}$ continuous on Ω is denoted by $C(\Omega)$.

The set of all functions $f : \Omega \to \mathbb{C}$ which are continuous and bounded on Ω is denoted by $C_b(\Omega)$. It is a metric space with the metric d, defined by

$$d(f, g) = \sup_{x \in \Omega} |f(x) - g(x)|.$$

As in Example 1.7, the convergence in $C_b(\Omega)$ is the uniform convergence.

Example 1.9 (Lebesgue space $L_p(\Omega)$). Let $\Omega \subset \mathbb{R}^n$ be a Lebesgue measurable set, $0 < p < \infty$ and $f : \Omega \to \mathbb{C}$. Then, $f \in L_p(\Omega)$ if

1) f is Lebesgue measurable on Ω,

2) $\int\limits_{\Omega} |f(x)|^p \, dx < \infty$.

The space $L_p(\Omega)$ is a semi-metric space with respect to the metric

$$d_p(f, g) = \begin{cases} \int\limits_{\Omega} |f(x) - g(x)|^p \, dx & \text{if } 0 < p < 1, \\ \Big(\int\limits_{\Omega} |f(x) - g(x)|^p \, dx\Big)^{\frac{1}{p}} & \text{if } 1 \leq p < \infty. \end{cases}$$

<u>Properties of a semi-metric space.</u>

$1'$) Obvious. 1) Fails. $d_p(f, g) = 0 \iff f \sim g$ on $\Omega \iff \operatorname{meas}\{x \in \Omega : f(x) \neq g(x)\} = 0$.

2) Obvious. 3) If $0 < p < 1$ then $\forall\, f, g, h \in L_p(\Omega)$,

$$d_p(f, g) = \int_{\Omega} |f(x) - g(x)|^p \, dx = \int_{\Omega} |f(x) - h(x) + h(x) - g(x)|^p \, dx$$
$$\leq \int_{\Omega} (|f(x) - h(x)| + |h(x) - g(x)|)^p \, dx \leq \int_{\Omega} |f(x) - h(x)|^p \, dx + \int_{\Omega} |h(x) - g(x)|^p \, dx$$
$$= d_p(f, h) + d_p(h, g).$$

If $1 \le p < \infty$, then by Minkowski's inequality,

$$d_p(f,g) = \Big(\int_\Omega |f(x)-g(x)|^p\,dx\Big)^{\frac{1}{p}} = \Big(\int_\Omega |f(x)-h(x)+h(x)-g(x)|^p\,dx\Big)^{\frac{1}{p}}$$
$$\le \Big(\int_\Omega |f(x)-h(x)|^p\,dx\Big)^{\frac{1}{p}} + \Big(\int_\Omega |h(x)-g(x)|^p\,dx\Big)^{\frac{1}{p}} = d_p(f,h)+d_p(h,g).$$

Let $n=1,\ \Omega=(0,1)$,

$$f_{\alpha,k}(x) = \begin{cases} k^\alpha & \text{if } 0 < x \le \frac{1}{k} \\ 0 & \text{if } \frac{1}{k} < x < 1 \end{cases}$$

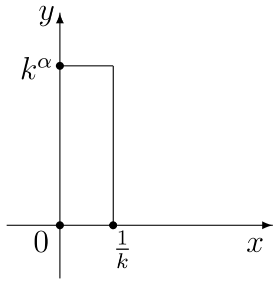


Figure 1.7

Clearly $\lim\limits_{k\to\infty} f_{\alpha,k}(x) = 0 \quad \forall\, x \in (0,1)$, but

$$\lim_{k\to\infty} f_{\alpha,k}(x) = 0 \quad \text{in} \quad L_p(0,1) \iff \alpha < \frac{1}{p}$$

since

$$d_p(f_{\alpha,k},0) = \Big(\int_0^1 |f_{\alpha,k}(x)-0|^p\,dx\Big)^{\frac{1}{p}} = \Big(\int_0^{\frac{1}{k}} k^{\alpha p}\,dx\Big)^{\frac{1}{p}} = k^{\alpha-\frac{1}{p}} \to 0 \iff \alpha < \frac{1}{p}.$$

So, if $\alpha \ge \frac{1}{p}$, then $\lim\limits_{k\to\infty} f_{\alpha,k}(x) = 0 \quad \forall\, x\in(0,1)$, but $f_{\alpha,k}$ does not converge to 0 on $L_p(0,1)$.

Note also that the function $f_{\frac{1}{p},k}$ converges to 0 in $L_q(0,1)$ for any $q < p$, but does not converge to 0 for any $q \ge p$.

Example 1.10 (Space $L_\infty(\Omega)$). Let $\Omega \subset \mathbb{R}^n$ be a measurable set and $f : \Omega \to \mathbb{C}$. Then $f \in L_\infty(\Omega)$: if

1) f is Lebesgue measurable on Ω.

2) $\rho(f) = \operatorname*{ess\,sup}\limits_{x\in\Omega} |f(x)| \equiv \inf\limits_{\substack{\omega\subset\Omega \\ \operatorname{meas}\omega=0}} \sup\limits_{x\in\Omega\setminus\omega} |f(x)| < \infty.$

Exercise. Prove that $L_\infty(\Omega)$ is a semi-metric space with the metric $d(f,g) = \rho(f-g)$.

Remark 1.2.1. If Ω is a Lebesgue measurable set and $\operatorname{meas}\Omega < \infty$, then $\forall\, f \in L_\infty(\Omega)$

$$\lim_{p\to+\infty} \|f\|_{L_p(\Omega)} = \|f\|_{L_\infty(\Omega)}.$$

If $\operatorname{meas}\Omega = \infty$, this equality holds not for all $f \in L_\infty(\Omega)$.

If, for example, $f \equiv 1$ on Ω, then $\|f\|_{L_\infty(\Omega)} = 1$, but $\lim\limits_{p\to+\infty} \|f\|_{L_p(\Omega)} = +\infty$.

Example 1.11. Let $\Omega \subset \mathbb{R}^n$ be a Lebesgue measurable set. $M(\Omega)$ is the space of all functions $f : \Omega \to \mathbb{C}$ measurable on Ω. $M(\Omega)$ is a semi-metric space with the metric

$$d(f, g) = \int_\Omega \frac{|f(x) - g(x)|}{1 + |f(x) - g(x)|}\, dx.$$

Properties of a semi-metric space.

1′) Obvious. 1) Fails. $d_p(f, g) = 0 \iff f \sim g$ on Ω. 2) Obvious.

3) Note that $\forall\, a, b \in \mathbb{C}$

$$\frac{|a + b|}{1 + |a + b|} \le \frac{|a| + |b|}{1 + |a| + |b|}$$

because $|a + b| \le |a| + |b|$ and the function $f(t) = \frac{t}{1+t} = 1 - \frac{1}{1+t}$ is increasing. Next

$$\frac{|a| + |b|}{1 + |a| + |b|} \le \frac{|a|}{1 + |a| + |b|} + \frac{|b|}{1 + |a| + |b|} \le \frac{|a|}{1 + |a|} + \frac{|b|}{1 + |b|}.$$

Therefore, $\forall\, f, g, h \in M(\Omega)$

$$d(f, g) = \int_\Omega \frac{|f(x) - h(x) + h(x) - g(x)|}{1 + |f(x) - h(x) + h(x) - g(x)|}\, dx$$
$$\le \int_\Omega \frac{|f(x) - h(x)|}{1 + |f(x) - h(x)|}\, dx + \int_\Omega \frac{|h(x) - g(x)|}{1 + |h(x) - g(x)|}\, dx = d(f, h) + d(h, g).$$

Remark 1.2.2. Following the tradition we shall say that the spaces $L_p(\Omega)$, $0 < p \le \infty$ and $M(\Omega)$ are metric spaces, since $d(f, g) = 0 \iff f \sim g$ on $\Omega \iff \operatorname{meas}\{x \in \Omega : f(x) \ne g(x) = 0\}$, and the distinction between f and g is negligible.

Example 1.12 (Discrete metric). Let X be a set and $\forall\, x, y \in X$

$$d(x, y) = \begin{cases} 0 & \text{if } y = x, \\ 1 & \text{if } y \ne x. \end{cases}$$

Clearly, (X, d) is a metric space (though not much useful). In this space

$$B(x, r) = \begin{cases} \{x\} & \text{if } r \le 1, \\ X & \text{if } r > 1. \end{cases}$$

Convergence.

$$\lim_{k\to\infty} d(x_k, x) = 0 \iff \forall\, \varepsilon > 0\ \exists\, k_0 = k_0(\varepsilon) : \forall\, k > k_0 \quad x_k = x.$$

So $\{x_k\}_{k\in\mathbb{N}} = \{x_1, x_2, \dots, x_{k_0}, x, \dots, x, \dots\}$. Such sequences are called *stabilizing sequences.*

1.3 Open and closed sets

Definition 1.3.1. *A subset M of a metric space (X, d) is said to be **open** if*

$$\forall\, x \in M,\ \exists\, r > 0 : B_d(x, r) \subset M.$$

Definition 1.3.2. *A subset M of a metric space (X,d) is said to be **closed** if its complement in X, ${}^{c}M = X \setminus M$, is open.*

Definition 1.3.3. *A point x of a metric space (X,d) is called a **limit (or cluster) point** of a set $M \subset X$ if*

$$\forall\, r > 0, \quad B_d(x,r) \cap (M \setminus \{x\}) \neq \oslash.$$

Lemma 1.3.1. *A subset M of a metric space (X,d) is closed if and only if it contains all its limit points (if there are limit points).*

Proof. 1. Let M be closed and $x \in X$ be its limit point. Assume that $x \notin M$. Then $x \in X \setminus M$ which is an open set, hence $\exists\, r > 0$ such that $B_d(x,r) \subset X \setminus M$. Consequently, $B_d(x,r) \cap M = \oslash$. Contradiction. So $x \in M$.

2. Let all limit points of M belong to M and $x \in {}^{c}M = X \setminus M$. Assume that there does not exist $r > 0$ such that $B_d(x,r) \subset X \setminus M$. Then

$$\forall r > 0 \quad B_d(x,r) \not\subset X \setminus M$$
$$\Updownarrow$$
$$\forall r > 0\ \exists\, y \in B_d(x,r) : y \notin X \setminus M \implies y \in M.$$

Note that $y \neq x$ because $x \in X \setminus M$. So

$$B_d(x,r) \cap (M \setminus \{x\}) \ni y,$$

hence

$$B_d(x,r) \cap (M \setminus \{x\}) \neq \oslash.$$

This means that x is a limit point of M and $x \notin M$. Contradiction. So $\exists\, r > 0$ such that $B_d(x,r) \subset {}^{c}M$ and this happens for all $x \in {}^{c}M$. Thus, ${}^{c}M$ is open. $\square$

Examples. Let (X,d) be a metric space. Then

1) X is simultaneously open and closed,

2) the empty set is also simultaneously open and closed,

3) any open ball $B_d(x,r)$ is an open set,

4) any closed ball $\bar{B}_d(x,r)$ is a closed set,

5) any sphere $S_d(x,r)$ is a closed set.

(Exercise.)

Lemma 1.3.2. *Let (X,d) be a metric space, A be an arbitrary set and $\forall\, \alpha \in A$ $M_\alpha \subset X$ be an open set. Then $\bigcup\limits_{\alpha \in A} M_\alpha$ is an open set.*

Proof. Let $x \in \bigcup\limits_{\alpha \in A} M_\alpha$, then there exist $\alpha_x \in A$ such that $x \in M_{\alpha_x}$. Since M_{α_x} is open there exists $r_x > 0$ such that $B_d(x,r_x) \subset M_{\alpha_x} \subset \bigcup\limits_{\alpha \in A} M_\alpha$. $\square$

Lemma 1.3.3. *Let (X,d) be a metric space, $k \in \mathbb{N}$ and $M_1, \dots, M_k \subset X$ be open sets. Then $\bigcap\limits_{i=1}^{k} M_i$ is an open set.*

Proof. Let $x \in \bigcap\limits_{i=1}^{k} M_i$, then $x \in M_i$ for all $i = 1, \dots, k$. Since M_i is open there exists $r_i > 0$ such that $B_d(x, r_i) \subset M_i$. Let $r = \min\{r_1, \dots, r_k\}$, then $B_d(x, r) \subset B_d(x, r_i)$ for all $i = 1, \dots, k$. Hence, $B_d(x, r) \subset \bigcap\limits_{i=1}^{k} M_i$. So $\bigcap\limits_{i=1}^{k} M_i$ is open. □

Example. The intersection of infinitely many open sets is not necessarily open. Indeed, let $X = \mathbb{R}^n$ and d be the standard metric on $\mathbb{R}^n$, then

$$\bigcap_{k=1}^{\infty} B_d\Big(x, \frac{1}{k}\Big) = \{x\}$$

is not open.

Lemma 1.3.4. *Let (X, d) be a metric space, A be an arbitrary set and $\forall\, \alpha \in A$ $M_\alpha \subset X$ be a closed set. Then $\bigcap\limits_{\alpha \in A} M_\alpha$ is a closed set.*

Proof. Since the complements ${}^{c}M_\alpha$ are open sets for any $\alpha \in A$, by Lemma 1.3.2 the complement

$${}^{c}\Big(\bigcap_{\alpha \in A} M_\alpha\Big) = \bigcup_{\alpha \in A} {}^{c}M_\alpha$$

is open. Hence, the set $\bigcap\limits_{\alpha \in A} M_\alpha$ is closed. □

Lemma 1.3.5. *Let (X, d) be a metric space, $k \in \mathbb{N}$ and $M_1, \dots, M_k \subset X$ be closed sets. Then $\bigcup\limits_{i=1}^{k} M_i$ is a closed set.*

Proof. Since the complements ${}^{c}M_i$ are open sets, $i = 1, \dots, k, $, by Lemma 1.3.3 the complement

$${}^{c}\Big(\bigcup_{i=1}^{k} M_i\Big) = \bigcap_{i=1}^{k} {}^{c}M_i$$

is open. Hence, the set $\bigcup\limits_{i=1}^{k} M_i$ is closed. □

Example. The infinite union of closed sets is not necessarily closed. Indeed, let $X = \mathbb{R}^n$ and d be the standard metric on $\mathbb{R}^n$, then

$$\bigcup_{k=2}^{\infty} \bar{B}_d\Big(x, 1 - \frac{1}{k}\Big) = B_d(x, 1)$$

is not closed.

Remark. One-dimensional open sets have a simple description, namely, any open set $M \subset \mathbb{R}$ with respect to the standard metric has the form

$$M = \bigcup_{k=1}^{s} (a_k, b_k),$$

where $s \in \mathbb{N}$ or $s = \infty$ and (a_k, b_k) are non-intersecting intervals.

1.4 Continuous functions

Let (X, d) be a metric space and $M \subset X$. A point $x \in M$ is called an isolated point if

$$\exists\, r > 0 : B_d(x, r) \cap M = \{x\}.$$

A point $x \in X$ is called a limit point of M if

$$\forall\, r > 0, \quad B_d(x, r) \cap (M \setminus \{x\}) \neq \emptyset.$$

(We have considered these definitions for the case $X = \mathbb{R}^n$ with the standard metric.)

Definition 1.4.1. *Let (X, d_X) and (Y, d_Y) be metric spaces and $M \subset X$, $f : M \to Y$, $x \in X$, be a limit point of M, $y \in Y$. Then $\lim\limits_{\substack{\xi \to x \\ \xi \in M}} f(\xi) = y$ means that*

$$\begin{gathered} \forall\, \varepsilon > 0,\ \exists\, \delta > 0 : \forall\, \xi \in M : d_X(\xi, x) < \delta,\ \xi \neq x, \quad d_Y(f(\xi), y) < \varepsilon \\ \Updownarrow \\ \forall\, \varepsilon > 0,\ \exists\, \delta > 0 : \forall\, \xi \in (B_{d_X}(x, \delta) \cap M) \setminus \{x\}, \quad f(\xi) \in B_{d_Y}(y, \varepsilon). \end{gathered} \tag{1.3}$$

This definition is equivalent to the following one in terms of sequences: for any sequence $\{\xi_k\}_{k\in\mathbb{N}} \subset M$, $\xi_k \neq x$, $k \in \mathbb{N}$ such that $\lim\limits_{k\to\infty} \xi_k = x$ in X, $\lim\limits_{k\to\infty} f(\xi_k) = y$ in Y. The proof of the equivalence is essentially the same as for the case $X = \mathbb{R}$, $Y = \mathbb{R}$ with the standard metric, given in the course of mathematical analysis.

Definition 1.4.2. *Let (X, d_X) and (Y, d_Y) be metric spaces and $M \subset X$, $f : M \to Y$, $x \in X$. Then f is continuous at x if*

$$\begin{gathered} \forall\, \varepsilon > 0,\ \exists\, \delta > 0 : \forall\, \xi \in M : d_X(\xi, x) < \delta \implies d_Y(f(\xi), f(x)) < \varepsilon \\ \Updownarrow \\ \forall\, \varepsilon > 0,\ \exists\, \delta > 0 : \forall\, \xi \in B_{d_X}(x, \delta) \cap M, \quad f(\xi) \in B_{d_Y}(f(x), \varepsilon) \\ \Updownarrow \\ \forall\, \varepsilon > 0,\ \exists\, \delta > 0 : f(B_{d_X}(x, \delta) \cap M) \subset B_{d_Y}(f(x), \varepsilon). \end{gathered}$$

Let (X, d) be a metric space. A neighbourhood of $x \in X$ is any set $V \subset X$ such that there exists an open set U for which $x \in U \subset V$. Since $x \in U$ implies that there is exists $r > 0$ such that $B_d(x, r) \subset U$, this is equivalent to saying that V is a neighbourhood of x if there exists $r > 0$ such that $B(x, r) \subset V$.

Lemma 1.4.1. *Let (X, d_X) and (Y, d_Y) be metric spaces and $M \subset X$, $f : M \to Y$, $x \in X$. Then f is continuous at x if and only if for any neighbourhood $U_{f(x)}$ of $f(x)$ in Y there exists a neighbourhood V_x of x in X such that*

$$f(V_x \cap M) \subset U_{f(x)}.$$

Proof. 1. Let f be continuous at x. Given a neighbourhood $U_{f(x)}$ of $f(x)$ in Y there exists $\varepsilon > 0$ such that $B_{d_Y}(f(x), \varepsilon) \subset U_{f(x)}$. Hence, by the continuity of f at x there exists $\delta > 0$ such that $\forall\, \xi \in B_{d_X}(x, \delta) \cap M$ we have $f(\xi) \in B_{d_Y}(f(x), \varepsilon)$. Let $V_x = B_{d_X}(x, \delta)$, then $f(V_x \cap M) = \bigcup\limits_{\xi \in V_x \cap M} f(\xi) \subset B_{d_Y}(f(x), \varepsilon) \subset U_{f(x)}$.

2. Let for any neighbourhood $U_{f(x)}$ of $f(x)$ in Y there exist a neighbourhood V_x of x in X such that $f(V_x \cap M) \subset U_{f(x)}$. Take here $U_{f(x)} = B_{d_Y}(f(x), \varepsilon)$ for any $\varepsilon > 0$. Then there exists a neighbourhood W_x of x in X such that $f(W_x \cap M) \subset B_{d_Y}(f(x), \varepsilon)$. Next, there exists a $\delta > 0$ such that $B_{d_X}(x, \delta) \subset W_x$, hence $f(B_{d_X}(x, \delta) \cap M) \subset B_{d_Y}(f(x), \varepsilon)$. So f is continuous at x.

□

Definition 1.4.3. *Let (X, d_X) and (Y, d_Y) be metric spaces and $M \subset X$, $f : M \to Y$. Then f is continuous on M if it is continuous at any $x \in M$.*

Theorem 1.4.1. *Let (X, d_X) and (Y, d_Y) be metric spaces. Then a function $f : X \to Y$ is continuous on X if and only if for any open set $M \subset Y$ its preimage $f^{-1}(M) = \{x \in X : f(x) \in M\}$ is an open set.*

Proof. $\boxed{\text{Step 1.}}$ Let $f : X \to Y$ be continuous on X and $M \subset Y$ be an open set. Assume that $f^{-1}(M)$ is not open, then there exists $x \in f^{-1}(M)$ such that for any $k \in \mathbb{N}$ $B_{d_X}(x, \frac{1}{k}) \not\subset f^{-1}(M)$. So for any $k \in \mathbb{N}$ there exists $x_k \in B_{d_X}(x, \frac{1}{k})$ such that $x_k \notin f^{-1}(M) \iff f(x_k) \notin M$. Since $f(x) \in M$ is open, there exists $r > 0$ such that $B_{d_Y}(f(x), r) \subset M$. Therefore $f(x_k) \notin B_{d_Y}(f(x), r) \iff d_Y(f(x_k), f(x)) \geq r$.

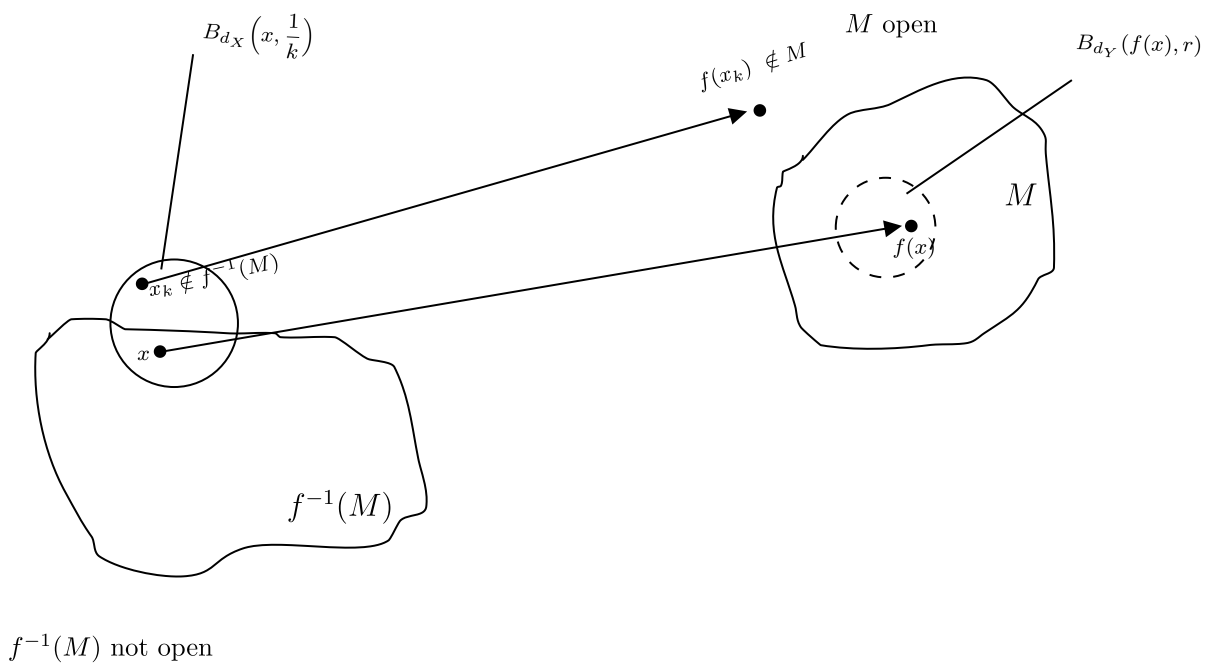


Figure 1.8

Clearly, $\lim\limits_{k\to\infty} x_k = x$ in X, hence by the continuity of f $\lim\limits_{k\to\infty} f(x_k) = f(x)$ in Y $\iff$ $\lim\limits_{k\to\infty} d_Y(f(x_k), f(x)) = 0$. Contradiction. So the set $f^{-1}(M)$ is open.

$\boxed{\text{Step 2.}}$ Let for any open set $M \subset Y$ the set $f^{-1}(M)$ be open. Then for any open ball $B_{d_Y}(f(x), \varepsilon)$ $(\varepsilon > 0)$ its preimage $f^{-1}(B_{d_Y}(f(x), \varepsilon))$ is open. Note that $x \in f^{-1}(B_{d_Y}(f(x), \varepsilon))$ because $f(x) \in B_{d_Y}(f(x), \varepsilon)$. Therefore, there exists $\delta > 0$ such that $B_{d_X}(x, \delta) \subset f^{-1}(B_{d_Y}(f(x), \varepsilon))$. Hence

$$\begin{aligned} f(B_{d_X}(x, \delta)) &\subset f(f^{-1}(B_{d_Y}(f(x), \varepsilon))) = \{f(\xi) : \xi \in f^{-1}(B_{d_Y}(f(x), \varepsilon))\} \\ &= \{f(\xi) : f(\xi) \in B_{d_Y}(f(x), \varepsilon)\} = B_{d_Y}(f(x), \varepsilon). \end{aligned}$$

□

Definition 1.4.4. *Let (X, d_X) and (Y, d_Y) be metric spaces and $M \subset X$, $f : M \to Y$. Then f is uniformly continuous on M if*

$$\forall\, \varepsilon > 0,\ \exists\, \delta > 0 : \forall\, \xi, x \in M : d_X(\xi, x) < \delta \implies d_Y(f(\xi), f(x)) < \varepsilon.$$

Lemma 1.4.2. *Let (X, d_X) be a metric space. Then for any $x_1, y_1, x_2, y_2 \in X$*

$$|d(x_1, y_1) - d(x_2, y_2)| \le d(x_1, x_2) + d(y_1, y_2).$$

Proof. By applying the triangle inequality twice we get

$$d(x_1, y_1) \le d(x_1, x_2) + d(x_2, y_1) \le d(x_1, x_2) + d(x_2, y_2) + d(y_2, y_1).$$

For similar reasons

$$d(x_2, y_2) \le d(x_2, x_1) + d(x_1, y_2) \le d(x_1, x_2) + d(x_1, y_1) + d(y_1, y_2).$$

Hence

$$-(d(x_1, x_2) + d(y_1, y_2)) \le d(x_1, y_1) + d(x_2, y_2) \le d(x_1, x_2) + d(y_1, y_2)$$

and the desired inequality follows. □

Corollary 1.4.1. *Let (X, d_X) be a metric space, $x_k, y_k \in X$, $k \in \mathbb{N}$, $x, y \in X$ and*

$$\lim_{k\to\infty} x_k = x, \quad \lim_{k\to\infty} y_k = y \quad \text{in} \quad X.$$

Then

$$\lim_{k\to\infty} d(x_k, y_k) = d(x, y).$$

Proof. By Lemma 1.4.2

$$|d(x_k, y_k) - d(x, y)| \le d(x_k, x) + d(y_k, y).$$

By passing to the limit as $k \to \infty$ we get the desired equality. □

For a metric space (X, d) the metric d is a function defined on the direct product $X \times X$. In order to discuss the continuity of d we need to define a metric on $X \times X$.

Lemma 1.4.3. *Let (X, d_X) and (Y, d_Y) be metric spaces. Then $(X \times Y, d_{X\times Y})$, where for all $x = (x_1, x_2)$, $y = (y_1, y_2) \in X \times Y$*

$$d_{X\times Y}(x, y) \equiv d_{X\times Y}((x_1, x_2), (y_1, y_2)) = d_X(x_1, y_1) + d_Y(x_2, y_2),$$

is a metric space.

Proof. Exercise. □

Lemma 1.4.4. *Let (X, d_X) be a metric space. Then the function $d_X : X \times X \to \mathbb{R}_+$ is uniformly continuous on $(X \times X, d_{X\times X})$. ($\mathbb{R}_+$ is endowed with the standard metric $d_{\mathbb{R}_+}(x, y) = |x - y|$ for all $x, y \in \mathbb{R}_+$.)*

Proof. By Lemma 1.4.2 for all $x = (x_1, x_2)$, $y = (y_1, y_2)$ where $x_1, x_2, y_1, y_2 \in X$

$$|d_X(x_1, x_2) - d_X(y_1, y_2)| \le d_X(x_1, y_1) + d_X(x_2, y_2).$$

This means that

$$d_{\mathbb{R}_+}(d_X(x), d_X(y)) \le d_{X\times X}(x, y).$$

Hence $\forall\, \varepsilon > 0$, $\exists\, \delta > 0$ (namely, $\delta = \varepsilon$) such that for all $x, y \in X$ for which

$$d_{X\times X}(x, y) < \delta = \varepsilon,$$

we have

$$d_{\mathbb{R}_+}(d_X(x), d_X(y)) < \varepsilon.$$

□

1.5 Compact sets

Definition 1.5.1. *Let (X, d) be a metric space. A set $K \subset X$ is said to be compact if any open cover of K, that is, a collection of open sets $\{O_\alpha\}_{\alpha \in A}$ such that*

$$K \subset \bigcup_{\alpha \in A} O_\alpha$$

(A is an arbitrary set) contains a finite subcover, that is, a finite collection $O_{\alpha_1}, \dots, O_{\alpha_k},\ k \in \mathbb{N}$ such that

$$K \subset \bigcup_{i=1}^{k} O_{\alpha_i}.$$

This definition is equivalent to the following one.

Definition 1.5.2. *Let (X, d) be a metric space. A set $K \subset X$ is said to be sequentially compact if every sequence $\{x_k\}_{k \in \mathbb{N}} \subset K$ contains a subsequence $\{x_{k_m}\}_{m \in \mathbb{N}}$ convergent to an element $x \in K$.*

Without proof. (In the case of more general topological spaces, these definitions, in general, are not equivalent.)

Example 1.13. If $X = \mathbb{R}^n$, then a set $K \subset \mathbb{R}^n$ is compact if and only if it is closed and bounded.

Theorem 1.5.1. *Let (X, d) be a metric space, $K \subset X$ be a compact set, and $f : K \to \mathbb{R}$ be a continuous function. Then*

1) *f is bounded on K,*
2) *f attains its supremum and infimum on K.*

Proof. 1) Suppose that f is not bounded. Then $\forall\, k \in \mathbb{N},\ \exists\, x_k \in K$ such that $|f(x_k)| \geq k$. Since $\{x_k\}_{k \in \mathbb{N}} \subset K$, there exists a subsequence $\{x_{k_m}\}_{m \in \mathbb{N}}$ and $x \in K$ such that $\lim\limits_{m \to \infty} x_{k_m} = x$. Since f is continuous on K, $\lim\limits_{m \to \infty} |f(x_{k_m})| = |f(x)|$. On the other hand $\lim\limits_{m \to \infty} |f(x_{k_m})| = \infty$. Contradiction. So f is bounded.

2) Let $M = \sup\limits_{x \in K} f(x)$, then $\forall\, k \in \mathbb{N},\ \exists\, x_k \in K$ such that $M \geq f(x_k) > M - \frac{1}{k}$. By Definition 1.5.2 there exists a subsequence $\{x_{k_m}\}_{m \in \mathbb{N}}$ and $x \in K$ such that $\lim\limits_{m \to \infty} x_{k_m} = x$. Since f is continuous on K $\lim\limits_{m \to \infty} f(x_{k_m}) = f(x)$ and

$$M \geq f(x) = \lim_{m \to \infty} f(x_{k_m}) \geq \lim_{m \to \infty} \left(M - \frac{1}{k_m} \right) = M.$$

So $f(x) = M$.

Similarly, f attains its infimum on K.

□

Definition 1.5.3. *Let (X, d) be a metric space. A sequence $\{x_k\}_{k \in \mathbb{N}} \subset X$ is said to be Cauchy (= fundamental) if $\forall\, \varepsilon > 0,\ \exists\, k_0 \in \mathbb{N} : \forall\, k, m > k_0$*

$$d(x_k, x_m) < \varepsilon$$

$$\Updownarrow$$

$$\lim_{k,m \to \infty} d(x_k, x_m) = 0.$$

Lemma 1.5.1. *Any convergent sequence in a metric space is a Cauchy sequence.*

Proof. Let $\{x_k\}_{k\in\mathbb{N}} \subset X,\ x \in X$ and $\lim\limits_{k\to\infty} x_k = x \iff \lim\limits_{k\to\infty} d(x_k, x) = 0$. Then

$$0 \le d(x_k, x_m) \le d(x_k, x) + d(x, x_m).$$

Hence $\lim\limits_{k,m\to\infty} d(x_k, x_m) = 0$. □

Definition 1.5.4. *A subset K of a metric space (X, d) is bounded if $\exists\, R > 0\ \exists\, x_0 \in X : \forall\, x \in K \quad d(x, x_0) < R\ (\iff K \subset B_d(x_0, R))$.*

Lemma 1.5.2. *Any compact subset K of a metric space (X, d) is closed and bounded.*

Proof. Let $K \subset X$ be a compact set.

1. Let $\{x_k\}_{k\in\mathbb{N}} \subset K$ and $x \in X$ such that $\lim\limits_{k\to\infty} x_k = x$ in X. Since K is compact there exists a subsequence $\{x_{k_m}\}_{m\in\mathbb{N}}$ and $y \in K$ such that $\lim\limits_{m\to\infty} x_{k_m} = y$ in X. By the uniqueness of the limit $x = y$, hence $x \in K$. This implies that K is closed.
2. Assume that K is unbounded. Then $\forall\, x \in X\ \forall\, R > 0, \quad K \not\subset B_d(x, R)$, that is, $\exists\, y \in K : y \notin B_d(x, R) \iff d(y, x) \ge R$.

   Let $x_1 \in K$. Choose $x_2 \in K : d(x_1, x_2) \ge 1$. Next choose $x_3 \in K : d(x_1, x_3) \ge 1 + d(x_1, x_2)$. Then $d(x_1, x_3) \ge 1$ and $d(x_2, x_3) \ge d(x_1, x_3) - d(x_1, x_2) \ge 1$ and so on.

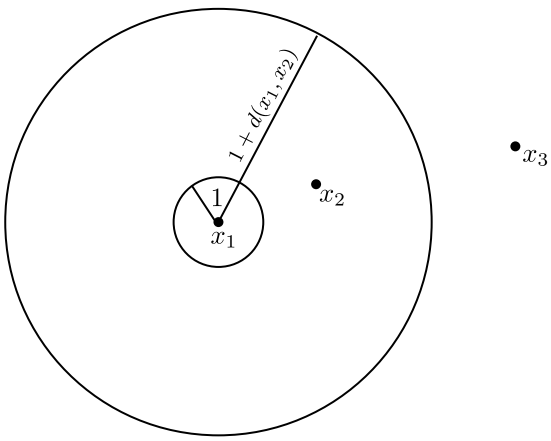


Figure 1.9

   Then $\{x_k\}_{k\in\mathbb{N}} \subset K$ and $d(x_k, x_m) \ge 1$ for any $k, m \in \mathbb{N}$. Hence for any subsequence $\{x_{k_s}\}_{s\in\mathbb{N}}$ we have

$$d(x_{k_s}, x_{k_\sigma}) \ge 1 \quad \forall\, s, \sigma \in \mathbb{N}.$$

   Therefore $\{x_{k_s}\}_{s\in\mathbb{N}}$ is not a Cauchy sequence, and hence by Lemma 1.5.1, not convergent.

□

Example. The closed unit ball in $C[0, 1]$

$$B = \{f \in C[0, 1] : d(0, f) = \max_{0\le x\le 1} |f(x)| \le 1\}$$

is a closed bounded set but it is not compact.

Proof. 1. B is bounded by its definition.

2. Let $g \in [0,1]$ be a limit point of B. Then there exist $f_k \in B,\ k \in \mathbb{N}$, such that $\lim\limits_{k\to\infty} f_k = g$ in $C[0,1]$. Since $f_k \in B$ for any $x \in [0,1]$
$$|f_k(x)| \le 1.$$
By passing to the limit as $k \to \infty$ we get that for any $x \in [0,1]$
$$|g(x)| \le 1 \implies g \in B.$$

3. Consider the sequence $\{x^k\}_{k\in\mathbb{N}} \subset B$ and any of its subsequences $\{x^{k_m}\}_{m\in\mathbb{N}}$. Assume that there exists $f \in B$ such that $x^{k_m} \to f(x)$ in $C[0,1]$. Then, in particular, for any $x \in [0,1]$
$$\lim_{m\to\infty} x^{k_m} = f(x)$$
which implies that
$$f(x) = \begin{cases} 0 & 0 \le x < 1 \\ 1 & x = 1. \end{cases}$$
Hence $f \notin C[0,1]$. Contradiction. So there are no subsequences of $\{x^k\}_{k\in\mathbb{N}}$ which are convergent to an element of B.

□

Lemma 1.5.3. *Any Cauchy sequence in a metric space is bounded.*

Proof. Let (X,d) be a metric space and $\{x^k\}_{k\in\mathbb{N}} \subset X$ be a Cauchy sequence $\iff \lim\limits_{k,m\to\infty} d(x_k, x_m) = 0$. Then there exists $n \in \mathbb{N}$ such that $\forall\, k, m \ge n \quad d(x_k, x_m) < 1$, hence
$$\forall\, k \ge n \quad d(x_k, x_n) < 1 \quad \text{and}$$
$$\forall\, k \in \mathbb{N} \quad d(x_k, x_n) < 1 + \max_{1\le s\le n} d(x_s, x_n).$$

□

Definition 1.5.5. *A set K in a metric space (X,d) is said to be precompact if its closure $\bar{K}$ is compact.*

By the definition of a compact set it follows that a set $K \subset X$ is precompact if and only if every sequence $\{x^k\}_{k\in\mathbb{N}} \subset K$ contains a subsequence convergent to an element $x \in X$.

Lemma 1.5.4. *A precompact subset K of a metric space (X,d) is bounded.*

Proof. Follows from Lemma 1.5.3 because $K \subset \bar{K}$ and $\bar{K}$ is compact. □

Theorem 1.5.2. *Let (X,d) be a metric space, $K \subset X$ be precompact and $f : K \to \mathbb{R}$ be a uniformly continuous function on K. Then f is bounded.*

Proof. Suppose that f is not bounded. Then $\forall\, k \in \mathbb{N},\ \exists\, x_k \in K : |f(x_k)| \ge k$. The set $\{x_k\}_{k\in\mathbb{N}}$ contains a convergent subsequence $\{x_{k_m}\}_{m\in\mathbb{N}}$ with $|f(x_{k_m})| \ge k_m$. On the other hand, by uniform continuity $\forall\, \varepsilon > 0\ \exists\, \delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ for $d(x,y) < \delta$. Since the sequence $\{x_{k_m}\}_{m\in\mathbb{N}}$ converges, by Lemma 1.5.1
$$\lim_{s,m\to\infty} d(x_{k_m}, x_{k_s}) = 0.$$
Hence there exists $n_0 : \forall\, m, s > n_0$
$$d(x_{k_m}, x_{k_s}) < \delta$$

which implies that

$$|f(x_{k_m}) - f(x_{k_s})| < \varepsilon$$

and for any fixed $s > n_0$ for all $m > n_0$

$$|f(x_{k_m})| \le |f(x_{k_s})| + |f(x_{k_m}) - f(x_{k_s})| \le |f(x_{k_s})| + \varepsilon.$$

So the sequence $\{f(x_{k_m})\}$ is bounded.
Contradiction. □

Theorem 1.5.3 (Arzelà–Ascoli). *Let $-\infty < a < b < +\infty$. A set $K \subset C[a,b]$ is precompact if and only if functions in K are:*

1) *uniformly bounded, that is, $\exists\, M > 0 : \forall\, f \in K,\ \forall\, x \in [a,b], \quad |f(x)| \le M$,*
2) *equicontinuous, that is, $\forall\, \varepsilon > 0,\ \exists\, \delta > 0 : \forall\, f \in K,\ \forall\, x, y \in [a,b],\ |x-y| < \delta \implies |f(x) - f(y)| < \varepsilon$.*

Without proof.

Remark. Condition 1) means that the set K is bounded in $C[a,b]$. The necessity of this condition follows from Lemma 1.5.4

1.6 Complete metric spaces

Definition 1.6.1. *A metric space (X,d) is said to be complete if any Cauchy sequence in X is convergent in X*

$$\Updownarrow$$

$$\{x_k\}_{k\in\mathbb{N}} \subset X : \lim_{k,m\to\infty} d(x_k, x_m) = 0 \implies \exists\, x \in X : \lim_{k\to\infty} d(x_k, x) = 0.$$

Example. $\mathbb{R}^n, \mathbb{C}^n,\ n \in \mathbb{N}$, are complete metric spaces endowed with the standard metric. $\mathbb{Q}$ (the set of all rational numbers) is not complete.

Theorem 1.6.1. *Let $K \subset \mathbb{R}^n$ be a non-empty set and $C_b(K)$ be the metric of all functions $f : K \to \mathbb{C}$ continuous and bounded on K with the metric*

$$d(f,g) = \sup_{x\in K} |f(x) - g(x)| \quad \forall\, f, g \in C_b(K).$$

Then the space $C_b(K)$ is complete.

Proof. Let $\{f_k\}_{k\in\mathbb{N}} \subset C_b(K)$ be a Cauchy sequence in $C_b(K)$, that is, $\forall\, \varepsilon > 0,\ \exists\, k_0 \in \mathbb{N} : \forall\, k, m \in \mathbb{N},\ k, m > k_0$

$$d(f_k, f_m) = \sup_{x\in K} |f_k(x) - f_m(x)| < \varepsilon.$$

Then

$$\forall\, x \in K \quad |f_k(x) - f_m(x)| < \varepsilon, \tag{1.4}$$

hence $\forall\, x \in K$ the sequence $\{f_k(x)\}_{k\in\mathbb{N}} \subset \mathbb{C}$ is a Cauchy sequence in $\mathbb{C}$. Since $\mathbb{C}$ is complete $\forall\, k \in K\ \exists\, y \in \mathbb{C} : \lim\limits_{k\to\infty} f_k(x) = y$. Let $\forall\, k \in K\ f(x) = y$. By passing to the limit in (1.4) as $m \to \infty$ we get

$$\forall\, k \in \mathbb{N},\ k > k_0,\ \forall\, x \in K \quad |f_k(x) - f(x)| \le \varepsilon$$

$$\Downarrow$$

$$\forall\, k \in \mathbb{N},\ k > k_0 \quad d(f_k, f) \le \varepsilon,$$

hence

$$\lim_{k\to\infty} f_k = f \quad \text{in} \quad C_b(K).$$

Since the convergence in $C_b(K)$ is the uniform convergence, it follows that $f \in C_b(K)$. □

Exercise. Prove this.

Theorem 1.6.2 (Riesz–Fischer). *Let $K \subset \mathbb{R}^n$ be a Lebesgue measurable set, $0 < p \leq \infty$. Then the space $L_p(K)$ with the metric*

$$d_p(f,g) = \begin{cases} \int\limits_K |f(x) - g(x)|^p \, dx & \text{if } 0 < p < 1, \\ \Big(\int\limits_K |f(x) - g(x)|^p \, dx \Big)^{\frac{1}{p}} & \text{if } 1 \leq p < \infty, \\ \operatorname*{ess\,sup}_{x \in K} |f(x) - g(x)| & \text{if } p = \infty, \end{cases}$$

is complete.

Without proof.

1.7 Dense sets and separability

Definition 1.7.1. *Let (X, d) be a metric space. A subset $K \subset X$ is said to be dense in X if $\forall\, x \in X,\ \forall\, \varepsilon > 0,\ \exists\, y \in K : d(y, x) < \varepsilon \iff \bar{K} = X$.*

Definition 1.7.2. *A metric space (X, d) is said to be separable if there exists a countable set $K \subset X$ dense in X.*

Example 1.14. The set $\mathbb{Q}$ of all rational numbers and $\mathbb{I}$ of all irrational numbers are dense in $\mathbb{R}$.

Example 1.15. $\mathbb{R}$ is separable because the countable set $\mathbb{Q}$ is dense in $\mathbb{R}$.

Lemma 1.7.1. *Let $-\infty < a < b < \infty$. The space $C[a, b]$ is separable.*

Proof. By the Weierstrass approximation theorem $\forall\, f \in C[a, b],\ \forall\, \varepsilon > 0\ \exists$ a polynomial

$$p_n(x) = \sum_{k=0}^{n} a_k x^k, \quad a_k \in \mathbb{R}$$

such that

$$d(f, p_n) = \max_{a \leq x \leq b} |f(x) - p_n(x)| < \frac{\varepsilon}{2}.$$

Note that for all $k = 0, 1, \ldots, n$ and $x \in [a, b]$

$$|x|^k \leq 1 + |x|^n \leq \max_{a \leq x \leq b} (1 + |x|^n) \leq 1 + R^n \equiv c_n,$$

where $R = \max\{|a|, |b|\}$. Next for each $k = 0, 1, \ldots, n$ there exists $r_k \in \mathbb{Q}$ such that

$$|a_k - r_k| < \frac{\varepsilon}{2c_n(n+1)}.$$

Let

$$q_n(x) = \sum_{k=0}^{n} r_k x^k.$$

Then

$$\begin{aligned} d(f, q_n) &\le d(f, p_n) + d(p_n, q_n) = \max_{a\le x\le b} |f(x) - p_n(x)| + \max_{a\le x\le b} |p_n(x) - q_n(x)| \\ &< \frac{\varepsilon}{2} + \max_{a\le x\le b} \sum_{k=0}^{n} |a_k - r_k| \cdot |x|^k \le \frac{\varepsilon}{2} + \sum_{k=0}^{n} |a_k - r_k| \cdot \max_{a\le x\le b} |x|^k \\ &\le \frac{\varepsilon}{2} + c_n \sum_{k=0}^{n} \frac{\varepsilon}{2c_n(n+1)} = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

So the set of all polynomials with rational coefficients is dense in $C[a, b]$. This set is countable. Hence, the space $C[a, b]$ is separable. □

Lemma 1.7.2. *Let (X, d) be a metric space, $c > 0$, A be an uncountable set and let for any $\alpha, \beta \in A$ such that $\alpha \ne \beta$ there exist $x_\alpha, x_\beta \in X$ such that*

$$d(x_\alpha, x_\beta) \ge c.$$

Then the space X is not separable.

Proof. Assume to the contrary that X is separable. Then there exists a countable dense subset $G = \{y_k\}_{k\in\mathbb{N}}$ of X.

Hence, $\forall\, \alpha \in A,\ \exists\, y_{k_\alpha} \in G$ ($k_\alpha \in \mathbb{N}$) such that

$$d(x_\alpha, y_{k_\alpha}) < \frac{c}{2}.$$

Therefore,

$$F \equiv \bigcup_{\alpha\in A} \{x_\alpha\} \subset \bigcup_{k=1}^{\infty} B\Big(y_k, \frac{c}{2}\Big).$$

The set F is uncountable because there is a one-to-one correspondence between A and F, namely, $\varphi : A \to F$ defined by $\varphi(\alpha) = x_\alpha\ \forall\, \alpha \in A$. (Recall that $x_\alpha \ne x_\beta\ \forall\, \alpha, \beta \in A,\ \alpha \ne \beta$.)

Consequently, there exist $\alpha, \beta \in A,\ \alpha \ne \beta$ such that both x_α, x_β belong to the same ball $B(y_{k_0}, \frac{c}{2})$. (Otherwise, each $x_\alpha \in F$ is contained in only one ball $B(z_{k_\alpha}, \frac{c}{2})$, hence F is countable.) Therefore,

$$d(x_\alpha, x_\beta) \le d(x_\alpha, y_{k_0}) + d(y_{k_0}, x_\beta) < \frac{c}{2} + \frac{c}{2} = c.$$

Contradiction. So X is not separable. □

Lemma 1.7.3. *The space $C_b[0, \infty)$ is not separable.*

Proof. Let $\Omega_k = [k - 1, k),\ k \in \mathbb{N}$. Then

$$[0, \infty) = \bigcup_{k=1}^{\infty} \Omega_k, \quad \Omega_k \cap \Omega_m = \oslash,\ k \ne m.$$

Let $\varphi(x) = 0$ if $x \le 0$ or $x \ge 1$, $\varphi \in C[0, 1]$ and $\|\varphi\|_{C[0,1]} = 1$, say $\varphi(x) = 1 - 2[x - \frac{1}{2}],\ x \in [0, 1]$.

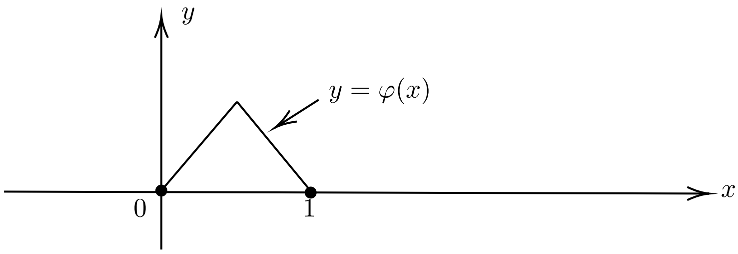


Figure 1.10

Let A be the set of all sequences $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k, \ldots)$, where $\alpha_k = 0$ or $\alpha_k = 1$. Recall that the set A is uncountable.
Finally, let

$$f_\alpha(x) = \sum_{\substack{k=1 \\ k:\alpha_k=1}}^{\infty} \varphi(x - k + 1).$$

Say, for $\alpha = (1, 0, 1, 1, 0, \ldots, 0, \ldots)$

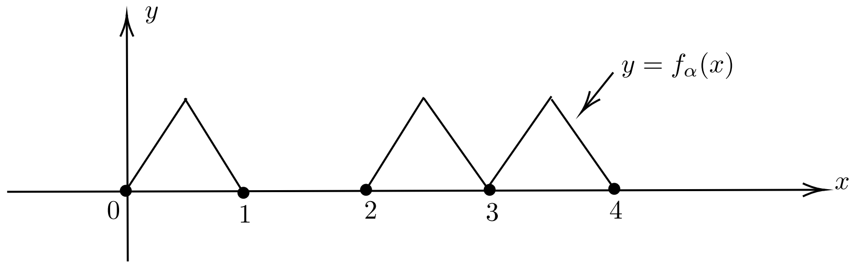


Figure 1.11

If $\alpha, \beta \in A,\ \alpha \neq \beta$, then for some $k_0 \in \mathbb{N}$ $\alpha_{k_0} \neq \beta_{k_0}$, say $\alpha_{k_0} = 1$ and $\beta_{k_0} = 0$. Hence

$$\begin{aligned} d(f_\alpha, f_\beta) &= \sup_{k\in\mathbb{N}} \max_{k-1\le x\le k} |f_\alpha(x) - f_\beta(x)| \\ &\ge \max_{k_0-1\le x\le k_0} |\varphi(x - k_0 + 1) - 0| = 1. \end{aligned}$$

(In fact $d(f_\alpha, f_\beta) = 1$.) Hence, by Lemma 1.7.2 the space $C_b[0, \infty)$ is not separable. □

Remark. In a similar way one can prove that the spaces $C_b[a, \infty),\ a > -\infty,\ C_b(-\infty, b],\ b < \infty$ and $C_b(-\infty, \infty)$ are not separable.

Let $-\infty \le a < b \le \infty$ and $B[a, b]$ be the space of all bounded functions on $[a, b]$.

Lemma 1.7.4. *The space $B[a, b]$ for any $-\infty \le a < b \le \infty$ is not separable.*

Proof. For the case of an infinite interval the statement follows by the above argument.
Let $[a, b]$ be a finite closed interval and

$$\Omega_k = \left(a + \frac{b-a}{k+1}, a + \frac{b-a}{k}\right], \quad k \in \mathbb{N}.$$

Then

$$(a, b] = \bigcup_{k=1}^{\infty} \Omega_k, \quad \Omega_k \cap \Omega_m = \varnothing,\ k \neq m.$$

Let as in the proof of Lemma 1.7.3 A be an uncountable set of all sequences $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k, \ldots)$, where $\alpha_k = 0$ or $\alpha_k = 1$, and let $\forall\, \alpha \in A$

$$f_\alpha(x) = \begin{cases} 1 & \text{if } x \in \Omega_k \text{ with } \alpha_k = 1,\ k \in \mathbb{N}, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\forall\, \alpha, \beta \in A,\ \alpha \neq \beta$ there is $k_0 \in \mathbb{N}$ such that $\alpha_{k_0} \neq \beta_{k_0}$, say $\alpha_{k_0} = 1,\ \beta_{k_0} = 0$, and

$$\begin{aligned} d(f_\alpha, f_\beta) &= \sup_{a\le x\le b} |f_\alpha(x) - f_\beta(x)| \\ &\ge \sup_{x\in\Omega_{k_0}} |f_{\alpha_{k_0}}(x) - f_{\beta_{k_0}}(x)| = 1. \end{aligned}$$

(In fact $d(f_\alpha, f_\beta) = 1$.) Hence, by Lemma 1.7.2 the space $B[a, b]$ is not separable. □

Theorem 1.7.1. *Let $\Omega \subset \mathbb{R}^n$ be a measurable set,* $\operatorname{meas}\Omega > 0$, $0 < p \leq \infty$.

1) *If $0 < p < \infty$, then the space $L_p(\Omega)$ is separable.*

2) *If $p = \infty$, then the space $L_\infty(\Omega)$ is not separable.*

Proof. 1) The case $0 < p < \infty$ is without proof.

> **Exercise.** 2) By using the idea of the proof of Lemma 1.7.4, prove that the space $L_\infty(\Omega)$ is not separable.

□

1.8 Contraction mapping principle. Successive approximations

Let (X, d) be a metric space, and
$$T : X \to X.$$
Consider the equation
$$Tx = x, \quad x \in X.$$
Solutions to this equation are called fixed points of the mapping T (because the point x does not change after application of the mapping T to it).

How to prove the existence of x satisfying this equation? Let $x_0 \in X$ and let
$$x_{k+1} = Tx_k, \quad k \in \mathbb{N}_0.$$
Assume that
$$\lim_{k\to\infty} x_k = x \quad \text{in} \quad X$$
and that
$$\lim_{k\to\infty} Tx_k = Tx \quad \text{in} \quad X.$$
Then $x = Tx$.

The elements x_k are called successive approximations of x. Note that
$$x_k = T^k x_0, \quad k \in \mathbb{N}.$$
Indeed,
$$x_k = Tx_{k-1} = T(Tx_{k-2}) = T^2 x_{k-2} = \ldots = T^k x_0.$$
Our aim is to formulate the assumptions ensuring that this procedure works.

Definition 1.8.1. *Let (X, d) be a metric space. A mapping $T : X \to X$ is called a contraction mapping (or a contraction) if there exists $0 < \alpha < 1$ such that*

$$d(Tx, Ty) \leq \alpha d(x, y) \quad \forall\, x, y \in X. \tag{1.5}$$

Theorem 1.8.1 (Banach). *Let (X, d) be a complete metric space and $T : X \to X$ be a contraction mapping. Then T has a unique fixed point, that is, the equation $Tx = x$ has a unique solution.*

Moreover, for any $x_0 \in X$ the successive approximations $x_k = T^k x_0$, $k \in \mathbb{N}$, converge to the unique fixed point x :
$$\lim_{k\to\infty} x_k = x \quad \text{in} \quad X.$$

Proof. $\boxed{\text{Step 1.}}$ *Existence.* Let $x_0 \in X,\ x_{k+1} = Tx_k,\ k \in \mathbb{N}_0$. Then by (1.5)

$$\begin{aligned} d(x_{k+1}, x_k) &= d(Tx_k, Tx_{k-1}) \le \alpha d(x_k, x_{k-1}) \\ &= \alpha d(Tx_{k-1}, Tx_{k-2}) \le \alpha^2 d(x_{k-1}, x_{k-2}) \le \cdots \le \alpha^k d(x_1, x_0). \end{aligned} \tag{1.6}$$

If $m < n$, then by the triangle inequality

$$\begin{aligned} d(x_m, x_n) &\le d(x_m, x_{m+1}) + d(x_{m+1}, x_m) \\ &\le d(x_m, x_{m+1}) + d(x_{m+1}, x_{m+2}) + \cdots + d(x_{n-1}, x_n) \\ &\le (\alpha^m + \cdots + \alpha^{n-1}) d(x_1, x_0) \le \frac{\alpha^m}{1-\alpha} d(x_1, x_0). \end{aligned}$$

If $m > n$, then

$$d(x_m, x_n) = d(x_n, x_m) \le \frac{\alpha^n}{1-\alpha} d(x_1, x_0).$$

Hence, for all $m, n \in \mathbb{N}$

$$d(x_m, x_n) \le \frac{\max\{\alpha^m, \alpha^n\}}{1-\alpha} d(x_1, x_0)$$

and

$$\lim_{m,n\to\infty} d(x_m, x_n) = 0.$$

So $\{x_k\}_{k\in\mathbb{N}}$ is a Cauchy sequence. By the completeness of X there exists $x \in X$ such that

$$\lim_{k\to\infty} x_k = x \quad \text{in} \quad X \iff \lim_{k\to\infty} d(x_k, x) = 0.$$

Since by (1.5)

$$d(x_k, Tx) = d(Tx_{k-1}, Tx) \le \alpha d(x_{k-1}, x),$$

we have

$$\lim_{k\to\infty} d(x_k, Tx) = 0.$$

Next

$$0 \le d(x, Tx) \le d(x, x_k) + d(x_k, Tx) \to 0 \quad \text{as} \quad k \to \infty.$$

Hence

$$d(x, Tx) = 0 \iff Tx = x.$$

$\boxed{\text{Step 2.}}$ *Uniqueness.* If $Tx = x$ and $Ty = y$, then

$$d(x, y) = d(Tx, Ty) \le \alpha d(x, y).$$

Since $0 < \alpha < 1$, this is possible only if

$$d(x, y) = 0 \iff x = y.$$

□

Remark (Rate of convergence). By passing to the limit in (1.6) as $n \to \infty$, taking into account the continuity of a metric, we get

$$\begin{aligned} d(x_m, x) = \lim_{n\to\infty} d(x_m, x_n) &\le \frac{\alpha^m}{1-\alpha} d(x, x_0) = \frac{\alpha^m}{1-\alpha} d(Tx_0, x_0) \\ \le \frac{\alpha^m}{1-\alpha} (d(Tx_0, Tx) + d(Tx, x_0)) &\le \frac{\alpha^m}{1-\alpha} (\alpha d(x, x_0) + d(x, x_0)) \\ &\le \frac{2\alpha^m}{1-\alpha} d(x, x_0). \end{aligned}$$

So the smaller the value of $\alpha \in (0, 1)$ the quicker the convergence of the successive approximations x_m to the fixed point x of T. Also, it is desirable to choose the initial approximation x_0 to be reasonably close to the fixed point x.

Remark. The assumption $\alpha < 1$ is essential. For example, if $X = \mathbb{R},\ Tx = x \quad \forall\, x \in \mathbb{R}$, then

$$d(Tx, Ty) = |Tx - Ty| = |x - y|.$$

So (1.5) is satisfied with $\alpha = 1$. However, in this case the equation

$$Tx = x \iff x = x$$

has infinitely many solutions. If $Tx = x + 1 \quad \forall\, x \in \mathbb{R}$, then

$$d(Tx, Ty) = |Tx - Ty| = |x + 1 - (y + 1)| = |x - y|,$$

so (1.5) is again satisfied with $\alpha = 1$. However, in this case the equation

$$Tx = x \iff x + 1 = x$$

has no solutions.

Remark. Condition (1.5) cannot be replaced by the following weaker condition

$$d(Tx, Ty) < d(x, y) \quad \forall\, x, y \in X,\ x \neq y.$$

For example, if $X = \mathbb{R}_+,\ Tx = x + \frac{1}{x}$. Since $(Tx)' = 1 - \frac{1}{x^2} < 1$

$$\left| x + \frac{1}{x} - \left(y - \frac{1}{y} \right) \right| < |x - y|,$$

but the equation

$$Tx = x \iff x + \frac{1}{x} = x$$

has no solutions.

Remark. The assumption in Theorem 1.8.1 that a metric space X is complete cannot be omitted. Let, for example, $X = \mathbb{Q} \cap [1, \infty)$ with the standard metric and

$$Tx = \frac{x}{2} + \frac{1}{x}, \quad x \in X.$$

Then $T : X \to X,\ T$ is a contraction (with $\alpha = \frac{1}{2}$), but T has no fixed points in X.

Exercise. Prove this statement.

Remark. Let (X, d) be a complete linear metric space and $S : X \to X$. Consider the equation

$$Sx = 0. \tag{1.7}$$

It is equivalent to the equation

$$(I - S)x = x - Sx = x.$$

So, if $I - S$ is a contraction mapping, then (1.7) has a unique solution.

Example. Let $0 < \lambda_1 \leq \lambda_2,\ X = \mathbb{R}^2$ and

$$d(x, y) = \sqrt{|x_1 - y_1|} + \sqrt{|x_2 - y_2|}, \quad x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R}^2.$$

Let the mapping $T : \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$Tx = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \iff Tx = (\lambda_1 x_1, \lambda_2 x_2).$$

Prove that T is a contraction if and only if $\lambda_2 < 1$.

Exercise. Prove this statement.

1.9 Topological spaces

In Sections 1.1–1.8 we emphasized that the main definitions in the theory of metric spaces, namely, convergence, continuity and compactness, can be formulated by using only the notion of an open set. Moreover, basic facts related to these notions can be proved using only the properties of open sets related to their unions and intersections. This leads to the idea of considering a far-reaching generalization of the notion of a metric space, which uses only the notion of an open set.

Definition 1.9.1. *A topological space is a pair $(X, \mathcal{T})$ where X is a set and $\mathcal{T}$ is a topology, that is, a family of subsets of X satisfying the following properties:*

1) $\varnothing, X \in \mathcal{T}$,

2) *if* $O_\alpha \in \mathcal{T},\ \alpha \in A$, *where A is an arbitrary set, then* $\bigcup\limits_{\alpha \in A} O_\alpha \in \mathcal{T}$,

3) *if* $k \in \mathbb{N}$ *and* $O_1, \ldots, O_k \in \mathcal{T}$, *then* $\bigcap\limits_{i=1}^{k} O_i \in \mathcal{T}$.

The elements of $\mathcal{T}$ are called open sets. A set $A \subset X$ is said to be closed if its complement is open.

Example. Clearly, any metric space (X, d) is a topological space with the topology $\mathcal{T}$ consisting of all open subsets of X.

A topological space $(X, \mathcal{T})$ is said to be *metrizable* if there exists a metric d on X such that $\mathcal{T}$ is the family of all open subsets of X with respect to the metric d.

Let $(X, \mathcal{T})$ be a topological space. A neighbourhood of $x \in X$ is any set $V \subset X$ such that there exists an open set U for which $x \in U \subset V$.

Definition 1.9.2 (Convergence). *Let $(X, \mathcal{T})$ be a topological space. Given a sequence $\{x_k\}_{k \in \mathbb{N}} \subset X$ and a point $x \in X$, it is said that*

$$\lim_{k \to \infty} x_k = x \quad \text{in} \quad X$$

if for any neighbourhood of x there exists $k_0 \in \mathbb{N}$ such that $\forall\, k \in \mathbb{N},\ k > k_0$

$$x_k \in U.$$

Definition 1.9.3 (Continuity). *Let $(X, \mathcal{T}_X), (Y, \mathcal{T}_Y)$ be topological spaces, $M \subset X,\ f : M \to Y$ and $x \in M$. Then f is continuous at x if for any neighbourhood $U_{f(x)}$ of $f(x)$ in Y there exists a neighbourhood V_x of x in X such that*

$$f(V_x \cap M) \subset U_{f(x)}.$$

(Lemma 1.4.1 *for metric spaces was taken into account.*) *Moreover f is continuous on M if it is continuous at any point $x \in M$.*

Similar to the case of metric spaces it can be proved that a function $f : X \to Y$ is continuous on X if and only if $\forall\, A \in \mathcal{T}_Y$

$$f^{-1}(A) = \{x \in X : f(x) \in M\} \in \mathcal{T}_X.$$

Definition 1.9.4 (Compactness). *Let $(X, \mathfrak{T})$ be a topological space. A set $K \subset X$ is said to be compact if any open cover of K, that is, a collection $\{O_\alpha\}_{\alpha \in A}$ such that*

$$K \subset \bigcup_{\alpha \in A} O_\alpha$$

(A is an arbitrary set) contains a finite subcover, that is, a finite collection $O_{\alpha_1}, \dots, O_{\alpha_k},\ k \in \mathbb{N}$ such that

$$K \subset \bigcup_{i=1}^{k} O_{\alpha_i}.$$

(*Compare with* Definition 1.5.1 *for metric spaces.*)

Definition 1.9.5. *A topological space is said to be **normal** if all singletons are closed and any two disjoint closed sets are separable, that is, there exist two disjoint open sets containing these closed sets.*

Definition 1.9.6. *A collection of open sets is said to be a **base** of a topology if any open set can be represented as a union of open sets belonging to the base.*

Example. $\mathbb{R}^n$ with the standard topology has a countable base, namely, the set of all open balls $B(x, r)$, where $x = (x_1, \dots, x_n),\ x_k \in \mathbb{Q},\ k = 1, \dots, n$ and $r \in \mathbb{Q},\ r > 0$.

Theorem 1.9.1 (Urysohn). *A topological space with a countable base is metrizable if and only if it is normal.*

Without proof.

2. Normed spaces

2.1 Definition of a normed space

The notion of a norm is a generalization of the notion of the length of a vector in geometry, which can be applied to various, not necessarily geometric objects, say to sequences and functions.

Definition 2.1.1. *A normed space is a pair $(X, \|\cdot\|)$, where X is a linear (= vector) space and $\|\cdot\|$ is a norm on X, that is, a function on X, whose values are non-negative numbers, briefly*

$$\|\cdot\| : X \to \mathbb{R}_+,$$

such that for any $x, y \in X$ and for all scalars α, that is, for all $\alpha \in \mathbb{R}$ if X is a linear space with multiplication by real numbers, and for all $\alpha \in \mathbb{C}$ if X is a linear space with multiplication by complex numbers,

1) $\|x\| = 0$ *if and only if* $x = 0$,

2) $\|\alpha x\| = |\alpha| \cdot \|x\|$,

3) $\|x + y\| \le \|x\| + \|y\|$ *(triangle inequality).*

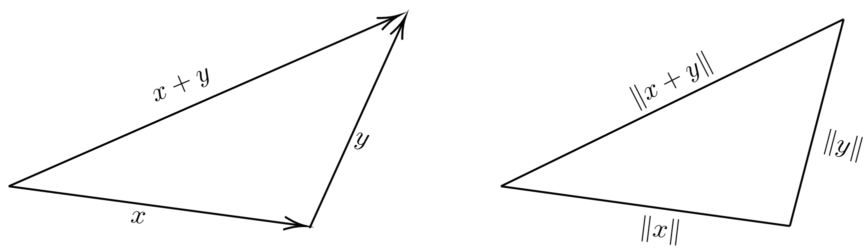


Figure 2.1 Triangle inequality

Instead of $(X, \|\cdot\|)$ we shall often write just X if there is no danger of confusion.
If instead of 1) we only have

1′) $\|0\| = 0$

then the space $(X, \|\cdot\|)$ is said to be a *semi-normed* space.
If instead of 3) we have 3′:

3′) $\exists\, c \ge 1 : \ \forall\, x, y \in X \quad \|x + y\| \le c(\|x\| + \|y\|)$,

then the space $(X, \|\cdot\|)$ is said to be a *quasi-normed* space.

Lemma 2.1.1. *A normed space $(X, \|\cdot\|)$ is a metric space (X, d) with the metric d defined by*

(2.1) $$d(x, y) = \|x - y\| \quad \forall\, x, y \in X.$$

Proof. Indeed by 1)

a) $d(x, y) = \|x - y\| = 0 \iff x - y = 0 \iff x = y,$

by 2)

b) $d(y, x) = \|y - x\| = \|(-1)(x - y)\| = \|x - y\| = d(x, y),$

and by 3)

c) $d(x, y) = \|x - y\| = \|x - z + z - y\| \leq \|x - z\| + \|z - y\| = d(x, z) + d(z, y).$

□

The metric d defined by (2.1) is said to be the metric induced by the norm $\|\cdot\|$.

The main definitions of the theory of metric spaces are transferred to normed spaces by using the metric induced by the norm.

Definition 2.1.2. *Given a point $x \in X$ and $r > 0$, we define three types of sets:*

$$\begin{aligned} B(x, r) &= \{y \in X : \|y - x\| < r\} \quad (\textit{open ball}), \\ \bar{B}(x, r) &= \{y \in X : \|y - x\| \leq r\} \quad (\textit{closed ball}), \\ S(x, r) &= \{y \in X : \|y - x\| = r\} \quad (\textit{sphere}). \end{aligned}$$

(*In all three cases, x is the center and r is the radius.*)

Lemma 2.1.2. *All balls in a normed space are convex.*

Proof. Let us consider, for example, an open ball, $B(x, r)$. Then for any $y, z \in B(x, r)$ and $0 < \alpha < 1$

$$\begin{aligned} \|\alpha y + (1 - \alpha)z - x\| &= \|\alpha(y - x) + (1 - \alpha)(z - x)\| \leq \|\alpha(y - x)\| + \|(1 - \alpha)(z - x)\| \\ &= \alpha\|y - x\| + (1 - \alpha)\|z - x\| < \alpha r + (1 - \alpha)r = r, \end{aligned}$$

hence $\alpha y + (1 - \alpha)z \in B(x, r)$. □

Remark. Recall that in a metric space a ball can be non-convex, see Section 1.1.

Definition 2.1.3. *Let $(X, \|\cdot\|)$ be a normed space, $\{x_k\}_{k \in \mathbb{N}} \subset X$. It is said that*

$$\lim_{k \to \infty} x_k = x \quad \text{in} \quad X$$

if

$$\lim_{k \to \infty} \|x_k - x\| = 0.$$

Example 2.1. Let $X = \mathbb{R}^n$ and for $x = (x_1, \ldots, x_n)$

$$\|x\|_2 = \Big(\sum_{i=1}^{n} |x_i|^2\Big)^{\frac{1}{2}}.$$

This is the standard norm in $\mathbb{R}^n$.

Example 2.2. Let $X = \mathbb{R}^n$ and for $x = (x_1, \ldots, x_n)$

$$\|x\|_1 = \sum_{i=1}^{n} |x_i|.$$

This is a norm on X.

Example 2.3. Let $X = \mathbb{R}^n$ and for $x = (x_1, \ldots, x_n)$

$$\|x\|_\infty = \max_{i=1,\ldots,n} |x_i|.$$

This is also a norm on X.

Note that the norms $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$ are equivalent on $\mathbb{R}^n$, namely, by Jensen's inequality

$$\|x\|_\infty \le \|x\|_2 \le \|x\|_1 \le n\|x\|_\infty$$

for all $x \in \mathbb{R}^n$, and the convergence of a sequence $\{x_k\}_{k\in\mathbb{N}} = \{x_{1k}, \ldots, x_{nk}\}_{k\in\mathbb{N}}$ to $x = (x_1, \ldots, x_n)$ is the same with respect to any of the norms $\|\cdot\|_1$, $\|\cdot\|_2$ and $\|\cdot\|_\infty$, namely,

$$\lim_{k\to\infty} x_{ik} = x_i, \quad i = 1, \ldots, n.$$

Exercise. Prove this.

Lemma 2.1.3 ("Normability" of a linear metric space). *A linear metric space (X, d) is a normed space $(X, \|\cdot\|)$ with the norm for which condition (2.1) is satisfied if and only if for all $x, y \in X$ and all scalars α*

$$1)\ d(x, y) = d(x - y, 0), \qquad 2)\ d(\alpha x, 0) = |\alpha| d(x, 0).$$

Proof. Step 1. If $(X, \|\cdot\|)$ is a normed space and the metric d is defined by (2.1), then

$$d(x, y) = \|x - y\| = \|(x - y) - 0\| = d(x - y, 0),$$
$$d(\alpha x, 0) = \|\alpha x - 0\| = \|\alpha x\| = |\alpha|\|x\| = |\alpha| d(x, 0).$$

So conditions 1) and 2) are satisfied.

Step 2. Let (X, d) be a metric space for which conditions 1) and 2) are satisfied. Put

$$\|x\| = d(x, 0) \quad \forall\, x \in X.$$

Then

1) $\|x\| = d(x, 0) = 0$ if and only if $x = 0$,
2) $\|\alpha x\| = |\alpha| d(x, 0) = |\alpha| \cdot \|x\|$,
3) $$\|x + y\| = d(x + y, 0) = d(x, -y) \le d(x, 0) + d(0, -y) = \|x\| + d(-y, 0)$$
$$= \|x\| + d(y, 0) = \|x\| + \|y\|.$$

□

Example 2.4 (Space ℓ_p with $1 \le p \le \infty$). Let $1 \le p \le \infty$ and $x = (x_1, \ldots, x_i, \ldots)$, where $x_i \in \mathbb{C}$. Recall that $x \in \ell_p$ if

$$\sum_{i=1}^{n} |x_i|^p < \infty \quad \text{if } p < \infty,$$
$$\sup_{i\in\mathbb{N}} |x_i| < \infty \quad \text{if } p = \infty.$$

For $1 \le p \le \infty$ ℓ_p is a normed space with the norm defined by

$$\|x\|_{\ell_p} = \begin{cases} \left(\sum\limits_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} & \text{for } 1 \le p < \infty, \\ \sup\limits_{i\in\mathbb{N}} |x_i| & \text{for } p = \infty. \end{cases}$$

The proof is similar to the proof of the fact that ℓ_p is a metric space (see Section 1.1).

Example 2.5 (Space ℓ_p with $0 < p < 1$). Let us try to define $\|x\|_{\ell_p}$ for $0 < p < 1$ similarly to the case $1 \le p < \infty$, namely,

$$\|x\|_{\ell_p} = \Big(\sum_{i=1}^{\infty} |x_i|^p\Big)^{\frac{1}{p}}.$$

However, this is not a norm. Indeed, if $x = (1, 0, \dots, 0, \dots),\ y = (0, 1, 0, \dots, 0, \dots)$, then

$$\|x + y\|_{\ell_p} = 2^{\frac{1}{p}} > \|x\|_{\ell_p} + \|y\|_{\ell_p} = 2.$$

So, the triangle inequality is not satisfied.

ℓ_p for $0 < p < 1$ is a quasi-normed space. Indeed, conditions 1) and 2) are clearly satisfied. Moreover, $\forall\, x, y \in \ell_p$

$$\begin{aligned}\|x + y\|_{\ell_p} &= \Big(\sum_{i=1}^{\infty} |x_i + y_i|^p\Big)^{\frac{1}{p}} \le \Big(\sum_{i=1}^{\infty} (|x_i| + |y_i|)^p\Big)^{\frac{1}{p}} \\ &\le \Big(\sum_{i=1}^{\infty} (|x_i|^p + |y_i|^p)\Big)^{\frac{1}{p}} = \Big(\sum_{i=1}^{\infty} |x_i|^p + \sum_{i=1}^{\infty} |y_i|^p\Big)^{\frac{1}{p}} \\ &\le 2^{\frac{1}{p}-1}\Big(\Big(\sum_{i=1}^{\infty} |x_i|^p\Big)^{\frac{1}{p}} + \Big(\sum_{i=1}^{\infty} |y_i|^p\Big)^{\frac{1}{p}}\Big) = 2^{\frac{1}{p}-1}(\|x\|_{\ell_p} + \|y\|_{\ell_p}). \qquad (2.2)\end{aligned}$$

So condition 3′) is satisfied with $c = 2^{\frac{1}{p}-1} > 1$. (We have applied the inequalities $(a + b)^p \le a^p + b^p$ for $0 < p < 1$ and $(a + b)^q \le 2^{q-1}(a^p + b^p)$ for $1 \le q < \infty,\ a, b \ge 0$.) Note that $2^{\frac{1}{p}-1}$ is a sharp multiple. It cannot be replaced by any smaller one. Indeed, if $x = (1, 0, \dots, 0, \dots),\ y = (0, 1, 0, \dots, 0, \dots)$, then there is equality in (2.2).

Example 2.6. Let $\Omega \subset \mathbb{R}^n$ be a non-empty set. The space $B(\Omega)$ of all bounded functions $f : \Omega \to \mathbb{C}$ is a normed space with the norm

$$\|f\|_{B(\Omega)} = \sup_{x \in \Omega} |f(x)|.$$

Properties 1) and 2) are obviously satisfied and $\forall\, f, g \in B(\Omega)$

$$\begin{aligned}\|f + g\|_{B(\Omega)} &= \sup_{x \in \Omega} |f(x) + g(x)| \le \sup_{x \in \Omega}(|f(x)| + |g(x)|) \\ &\le \sup_{x \in \Omega} |f(x)| + \sup_{x \in \Omega} |g(x)| = \|f\|_{B(\Omega)} + \|g\|_{B(\Omega)}\end{aligned}$$

which means that the triangle inequality is satisfied.

The space $C_b(\Omega)$ of all bounded continuous functions $f : \Omega \to \mathbb{C}$ is also a normed space witht the same norm.

Recall that if Ω is compact, then the space $C(\Omega)$ of all functions $f : \Omega \to \mathbb{C}$ coincides with the space $C_b(\Omega)$ and

$$\|f\|_{C(\Omega)} = \max_{x \in \Omega} |f(x)|.$$

Convergence in each of the three spaces is uniform, that is, $\lim\limits_{k \to \infty} f_k = f$ in any of these spaces means that $\forall\, \varepsilon > 0,\ \exists\, k_0 \in \mathbb{N}$ such that $\forall\, k \in \mathbb{N},\ k > k_0$

$$\forall\, x \in \Omega \quad |f_k(x) - f(x)| < \varepsilon.$$

Example 2.7 (Space L_p with $1 \le p \le \infty$). Let $\Omega \subset \mathbb{R}^n$ be a Lebesgue measurable set of positive measure. Recall that $L_p(\Omega)$ is the set of all Lebesgue measurable functions $f : \Omega \to \mathbb{C}$ for which

$$\int_\Omega |f(x)|^p \, dx < \infty \quad \text{for } 1 \le p < \infty$$

and

$$\operatorname*{ess\,sup}_{x\in\Omega} |f(x)| < \infty \quad \text{for } p = \infty.$$

Let

$$\|f\|_{L_p(\Omega)} = \Big(\int_\Omega |f(x)|^p \, dx \Big)^{\frac{1}{p}} \quad \text{for } 1 \le p < \infty$$

and

$$\|f\|_{L_\infty(\Omega)} = \operatorname*{ess\,sup}_{x\in\Omega} |f(x)| \quad \text{for } p = \infty.$$

Then $\|f\|_{L_p(\Omega)}$ is a semi-norm on $L_p(\Omega)$ and $L_p(\Omega)$ is a semi-normed space. The proof is similar to the proof of the fact that $L_p(\Omega)$ is a semi-metric space (see Section 1.1). Recall that following the tradition it is usually said that $L_p(\Omega)$ is a normed space, because $\|f\|_{L_p(\Omega)} = 0 \iff \text{meas}\{x \in \Omega : f(x) \ne 0\} = 0$ and the distinction between f and 0 is negligible.

Example 2.8 (Space L_p with $0 < p < 1$). Under the same assumptions on Ω as in Example 2.7 the expression

$$\|f\|_{L_p(\Omega)} = \Big(\int_\Omega |f(x)|^p \, dx \Big)^{\frac{1}{p}}, \quad 0 < p < 1$$

is a semi-quasi-norm.

Exercise. Prove this.

Remark. Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces, $M \subset X$, $f : M \to Y$, $x \in X$ be a limit point of M, $y \in Y$. The definition of $\lim\limits_{\substack{\xi\to y\\ \xi\in M}} f(\xi) = y$ is essentially the same as for the case of metric spaces. It is only required to replace in (1.3) $d_X(\xi, x)$ by $\|\xi - x\|_X$ and $d_Y(f(\xi), y)$ by $\|f(\xi) - y\|_Y$. The same refers to the definition of continuity of f at $x \in M$ and on M.

Definition 2.1.4. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces, $M \subset X$, $f : M \to Y$. The function f is said to be continuous at $x \in M$ if*

$$\forall\, \varepsilon > 0,\ \exists\, \delta > 0 : \forall\, y \in M \ \|y - x\|_X < \delta \implies \|f(y) - f(x)\|_Y < \varepsilon.$$

f is continuous on M if it is continuous at any point $x \in M$. f is uniformly continuous on M if

$$\forall\, \varepsilon > 0,\ \exists\, \delta > 0 : \forall\, x, y \in M \ \|x - y\|_X < \delta \implies \|f(x) - f(y)\|_Y < \varepsilon.$$

Lemma 2.1.4 (Reverse triangle inequality). *Let $(X, \|\cdot\|)$ be a normed space. Then $\forall\, x, y \in X$*

$$\big|\|x\| - \|y\|\big| \le \|x - y\|.$$

Exercise. Prove this.

Lemma 2.1.5. *Let $(X, \|\cdot\|)$ be a normed space. Then the function $\|\cdot\| : X \to \mathbb{R}_+$ is uniformly continuous on X.*

Exercise. Prove this by using Lemma 2.1.3.

Definition 2.1.5. *A normed space $(X, \|\cdot\|)$ is said to be a Banach space if it is complete, that is,*

$$\{x_k\}_{k\in\mathbb{N}} \subset X, \quad \lim_{k,m\to\infty} \|x_k - x_m\| = 0$$
$$\Downarrow$$
$$\exists\, x \in X : \lim_{k\to\infty} \|x_k - x\| = 0.$$

Definition 2.1.6. *Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ in a linear space X are said to be equivalent if there exist $c_1, c_2 > 0$ such that $\forall\, x \in X$*

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

2.2 Basic properties of operators in normed spaces

Definition 2.2.1. *A set M in a normed space $(X, \|\cdot\|)$ is said to be bounded if there exists $R > 0$ such that $M \subset B(0, R)$ or, equivalently, if $\sup\limits_{x\in M} \|x\| < \infty$.*

Definition 2.2.2. *An operator $T : X \to Y$ where X and Y are normed spaces is said to be bounded if the image of any bounded set in X is a bounded set in Y.*

Definition 2.2.3. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces and $T : X \to Y$. The norm of T is defined by*

$$\|T\| \equiv \|T\|_{X\to Y} = \sup_{\substack{x\in X\\ x\neq 0}} \frac{\|Tx\|_Y}{\|x\|_X}. \tag{2.3}$$

Remark. Consider the inequality: for some $c > 0$

$$\|Tx\|_Y \le c\|x\|_X \quad \forall\, x \in X. \tag{2.4}$$

It implies that $T(0) = 0$ and that

$$\frac{\|Tx\|_Y}{\|x\|_X} \le c \quad \forall\, x \in X,\ x \neq 0,$$

hence

$$\|T\|_{X\to Y} = \sup_{x\neq 0} \frac{\|Tx\|_Y}{\|x\|_X} \le c.$$

On the other hand, if $\|T\|_{X\to Y} < \infty$, then

$$\frac{\|Tx\|_Y}{\|x\|_X} \le \|T\|_{X\to Y} \quad \forall\, x \in X,\ x \neq 0,$$

hence

$$\|Tx\|_Y \le \|T\|_{X\to Y}\|x\|_X \quad \forall\, x \in X,\ x \neq 0.$$

If $T(0) = 0$, then

$$\|Tx\|_Y \le \|T\|_{X\to Y}\|x\|_X \quad \forall\, x \in X.$$

So inequality (2.4) holds with $c = \|T\|$. This means that if $T(0) = 0$, then the minimal constant in inequality (2.4) is equal to $\|T\|_{X\to Y}$.

In particular, this property holds for any linear operator T.

Remark. Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces and $B(X,Y)$ be the set of all operators $T : X \to Y$ for which $\|T\|_{X\to Y} < \infty$. $B(X,Y)$ is a linear space with respect to addition, defined for all $T_1, T_2 \in B(X,Y)$ by

$$(T_1 + T_2)(x) = T_1(x) + T_2(x) \quad \forall\, x \in X,$$

and multiplication by scalars α defined for all $T \in B(X,Y)$ by

$$(\alpha T)(x) = \alpha T(x) \quad \forall\, x \in X.$$

Lemma 2.2.1. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces and $B_0(X,Y)$ be the subset of $B(X,Y)$ of all operators $T : X \to Y$ for which $T(0) = 0$. Then $B_0(X,Y)$ is a normed space with the norm defined by equality* (2.3).

Proof. This follows since

1) $\|T\|_{X\to Y} = 0 \iff (\|Tx\|_Y = 0 \iff Tx = 0 \quad$ for all $x \in X,\ x \neq 0$ and $T(0) = 0$ since $T \in B_0(X,Y)) \iff T = 0$;

2) for any scalar α and any $T \in B_0(X,Y)$

$$\begin{aligned}\|\alpha T\|_{X\to Y} &= \sup_{\substack{x\in X\\ x\neq 0}} \frac{\|(\alpha T)(x)\|_Y}{\|x\|_X} = \sup_{\substack{x\in X\\ x\neq 0}} \frac{\|\alpha T(x)\|_Y}{\|x\|_X} \\ &= |\alpha| \sup_{\substack{x\in X\\ x\neq 0}} \frac{\|Tx\|_Y}{\|x\|_X} = |\alpha| \|T\|_{X\to Y};\end{aligned}$$

3) for any $T_1, T_2 \in B_0(X,Y)$

$$\begin{aligned}\|T_1 + T_2\|_{X\to Y} &= \sup_{\substack{x\in X\\ x\neq 0}} \frac{\|(T_1+T_2)(x)\|_Y}{\|x\|_X} = \sup_{\substack{x\in X\\ x\neq 0}} \frac{\|T_1x + T_2x\|_Y}{\|x\|_X} \\ \le \sup_{\substack{x\in X\\ x\neq 0}} \left(\frac{\|T_1x\|_Y}{\|x\|_X} + \frac{\|T_2x\|_Y}{\|x\|_X} \right) &\le \sup_{\substack{x\in X\\ x\neq 0}} \frac{\|T_1x\|_Y}{\|x\|_X} + \sup_{\substack{x\in X\\ x\neq 0}} \frac{\|T_2x\|_Y}{\|x\|_X} \\ &= \|T_1\|_{X\to Y} + \|T_2\|_{X\to Y}.\end{aligned}$$

□

Remark. Let $L(X,Y)$ be the set of all bounded linear operators $T : X \to Y$. Then it is a normed space with the norm defined by equality (2.3).

Lemma 2.2.2. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces and $T : X \to Y$ be a linear operator. Then*

$$\|T\|_{X\to Y} = \sup_{\substack{x\in X\\ \|x\|_X \le 1}} \|Tx\|_Y = \sup_{\substack{x\in X\\ \|x\|_X = 1}} \|Tx\|_Y. \tag{2.5}$$

Proof. Let

$$\|T\|^{(1)}_{X\to Y} = \sup_{\substack{x\in X\\ \|x\|_X \le 1}} \|Tx\|_Y, \qquad \|T\|^{(2)}_{X\to Y} = \sup_{\substack{x\in X\\ \|x\|_X = 1}} \|Tx\|_Y.$$

Then,

$$\|T\|^{(2)}_{X\to Y} \le \|T\|^{(1)}_{X\to Y} \le \sup_{\substack{x\in X\\ \|x\|_X \neq 1}} \frac{\|Tx\|_Y}{\|x\|_X} \le \|T\|_{X\to Y}.$$

Also, by the linearity of T

$$\begin{aligned}\|T\|_{X\to Y} &= \sup_{\substack{x\in X\\ x\neq 0}} \frac{\|Tx\|_Y}{\|x\|_X} = \sup_{\substack{x\in X\\ x\neq 0}} \left\| \frac{1}{\|x\|_X} Tx \right\|_Y \\ &= \sup_{\substack{x\in X\\ x\neq 0}} \left\| T\left(\frac{x}{\|x\|_X}\right)\right\|_Y \le \sup_{\substack{y\in X\\ \|y\|_X=1}} \|Ty\|_Y = \|T\|^{(2)}_{X\to Y},\end{aligned}$$

and equality (2.5) follows. □

Lemma 2.2.3. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces. A linear operator $T : X \to Y$ is bounded if and only if $\|T\|_{X\to Y} < \infty$.*

Proof. Let $\|T\|_{X\to Y} < \infty$ and $M \subset X$ be a bounded set. Then

$$\|Tx\|_Y \le \|T\|_{X\to Y}\|x\|_X \quad \forall\, x \in X.$$

Hence

$$\sup_{y\in T(M)} \|y\|_Y = \sup_{x\in M} \|Tx\|_Y \le \|T\|_{X\to Y} \sup_{x\in M} \|x\|_X < \infty.$$

Next, assume that the images $T(M)$ of all bounded subsets $M \subset X$ are bounded in Y. Then, in particular, $T(\bar{B}_X(0,1))$ is bounded in Y. Hence, by Lemma 2.2.2

$$\|T\|_{X\to Y} = \|T\|^{(1)}_{X\to Y} = \sup_{\substack{x\in X\\ \|x\|_X\le 1}} \|Tx\|_Y = \sup_{x\in \bar{B}_X(0,1)} \|Tx\|_Y = \sup_{y\in T(\bar{B}_X(0,1))} \|y\|_Y < \infty.$$

□

Definition 2.2.4. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces and $T : X \to Y$. The operator T is called a Lipschitz mapping if for some $L > 0$*

$$\|Tx - Ty\|_Y \le L\|x - y\|_X \quad \forall\, x, y \in X.$$

Example. If T is a bounded linear operator, the T is a Lipschitz mapping with $L = \|T\|_{X\to Y}$. Indeed, by the linearity of T

$$\|Tx - Ty\|_Y = \|T(x - y)\|_Y \le \|T\|_{X\to Y}\|x - y\|_X \quad \forall\, x, y \in X. \tag{2.6}$$

Lemma 2.2.4. *Let $(X, \|\cdot\|)$ be a normed space and $T : X \to X$ be a linear operator. Then T is a contraction mapping if and only if*

$$\|T\| < 1. \tag{2.7}$$

Proof. If (2.7) is satisfied, then by (2.6) T is a contraction mapping.

If T is a contraction mapping, then for some $0 < \alpha < 1$

$$d(Tx, Ty) = \|Tx - Ty\| \le \alpha\|x - y\| \quad \forall\, x, y \in X.$$

In particular, if $y = 0$, then $T(0) = 0$ and this inequality implies that

$$\|Tx\| \le \alpha\|x\| \quad \forall\, x \in X,$$

hence

$$\|T\| = \sup_{\substack{x\in X\\ x\neq 0}} \frac{\|Tx\|}{\|x\|} \le \alpha < 1.$$

□

Remark. Let $(X, \|\cdot\|_X)$ be a normed space, $T : X \to X$ and $f \in X$. Consider the equation

$$Tx = f. \tag{2.8}$$

Let the operator $S : X \to X$ be defined by

$$Sx = (I - T)x + f \quad \forall\, x \in X,$$

where I is the identity operator. Then (2.8) is equivalent to the equation

$$Sx = x. \tag{2.9}$$

Note that

$$\begin{aligned} d(Sx - Sy) &= \|Sx - Sy\| = \|(I-T)x + f - ((I-T)y + f)\| \\ &= \|(I-T)(x-y)\| \le \|I - T\|\|x - y\|. \end{aligned}$$

So, by the Banach theorem, if

$$I - T < 1,$$

then (2.8) has a unique solution.

Theorem 2.2.1. *Let X and Y be a normed spaces and $T : X \to Y$ be a linear operator. Then the continuity of T on X is equivalent to the boundedness of T on X.*

We shall prove a more general statement containing the statement of Theorem 2.2.1.

Theorem 2.2.2. *Let X and Y be a normed spaces and $T : X \to Y$ be a linear operator. Then the following statements are equivalent:*

1) *T is continuous at 0,*
2) *T is continuous at some point $x \in X$,*
3) *T is continuous on X,*
4) *T is uniformly continuous on X,*
5) *T is a Lipschitz mapping,*
6) *T is bounded on X.*

Proof. We shall use the following scheme of proof:

$$6 \underset{\text{proved on page 37}}{\longrightarrow} 5 \longrightarrow 4 \underset{\text{obvious}}{\longrightarrow} 3 \underset{\text{obvious}}{\longrightarrow} 2 \longrightarrow 1 \longrightarrow 6$$

5 $\longrightarrow$ 4 Let for some $L > 0$

$$\|Tx - Ty\|_Y \le L\|x - y\|_X \quad \forall\, x, y \in X.$$

Given an arbitrary $\varepsilon > 0$ choose $\delta = \frac{\varepsilon}{L}$. Then, for any $x, y \in X$ for which $\|x - y\|_X < \delta$ we have $\|Tx - Ty\|_Y \le L \cdot \frac{\varepsilon}{L} = \varepsilon$. Hence, T is uniformly continuous on X.

2 $\longrightarrow$ 1 Let T be continuous at $x \in X$. Then, due to the linearity of T

$$\begin{aligned} \lim_{t \to 0} Ty &= (y = z - x) = \lim_{z \to x} T(z - x) = \lim_{z \to x}(Tz - Tx) \\ &= \lim_{z \to x} Tz - Tx = Tx - Tx = 0 = T(0), \end{aligned}$$

hence T is continuous at 0.

① $\longrightarrow$ ⑥ $\boxed{\text{First proof.}}$ Let T be continuous at 0. Then there exists $\delta > 0$ such that

$$\forall\, z \in X : \|z\|_X \le \delta \quad \text{we have} \quad \|Tz\|_Y = \|Tz - T(0)\|_Y \le 1.$$

Let $x \in X,\ x \neq 0$. Take $z = \frac{x}{\|x\|_X}\delta$. Then

$$\|z\|_X = \Big\| \frac{x}{\|x\|_X}\delta \Big\|_X = \frac{\delta}{\|x\|_X}\|x\|_X = \delta,$$

hence

$$\|Tz\|_Y = \Big\| T\Big(\frac{x\delta}{\|x\|_X}\Big) \Big\|_Y \le 1.$$

By the linearity of T it follows that

$$\frac{\delta}{\|x\|_X}\|Tx\|_Y \le 1,$$

hence

$$\|T\|_{X\to Y} = \sup_{\substack{x\in X\\ x\neq 0}} \frac{\|Tx\|_Y}{\|x\|_X} \le \frac{1}{\delta} < \infty.$$

So T is bounded on X.

$\boxed{\text{Second proof.}}$ Let T be continuous at 0. Assume to the contrary that T is unbounded on X. Then $\forall\, n \in \mathbb{N}$ there exists $x_n \in X$ such that

$$\|Tx_n\|_Y \ge n\|x_n\|_X.$$

Let $y_n = \frac{1}{\sqrt{n}}\frac{x_n}{\|x_n\|_X}$. Then $\|y_n\|_X = \frac{1}{\sqrt{n}} \to 0$ as $n \to \infty$. Hence, $\|Ty_n\|_Y \to 0$ as $n \to \infty$. On the other hand, by the linearity of T

$$\|Ty_n\|_Y = \Big\| T\Big(\frac{x_n}{\sqrt{n}\|x_n\|_X}\Big) \Big\|_Y = \frac{1}{\sqrt{n}}\frac{\|Tx_n\|_Y}{\|x_n\|_X} \ge \sqrt{n},$$

hence $\|Ty_n\|_Y \to \infty$ as $n \to \infty$. Contradiction. So T is bounded on X. □

2.3 Further properties of linear operators in normed spaces

Lemma 2.3.1. *Let X, Y, Z be normed spaces, $T : X \to Y$ and $S : Y \to Z$ be bounded linear operators. Then the operator $ST : X \to Z$ is also a bounded linear operator and*

$$\|ST\|_{X\to Z} \le \|T\|_{X\to Y}\|S\|_{Y\to Z}. \tag{2.10}$$

Proof. Since

$$\|Sy\|_Z \le \|S\|_{Y\to Z}\|y\|_Y \quad \forall\, y \in Y$$

we have

$$\|ST\|_{X\to Z} = \sup_{\substack{x\in X\\ x\neq 0}} \frac{\|STx\|_Z}{\|x\|_X} \le \|S\|_{Y\to Z} \sup_{\substack{x\in X\\ x\neq 0}} \frac{\|Tx\|_Y}{\|x\|_X} = \|S\|_{Y\to Z}\|T\|_{X\to Y}.$$

□

Lemma 2.3.2. *Let X, Y be linear spaces and $T : X \to Y$ be a one-to-one linear operator. Then there exists a uniquely defined inverse operator*

$$T^{-1} : T(X) \to X$$

and this operator is linear.

Proof. $\boxed{\text{Step 1.}}$ Let $y \in T(X)$. Since T is one-to-one, there exists a uniquely defined $x \in X$ such that

$$y = Tx.$$

Let

$$T^{-1}y = x.$$

Then $T^{-1} : T(X) \to X$ and

$$T(T^{-1}y) = Tx = y \quad \forall\, y \in T(X)$$
$$\Updownarrow$$
$$TT^{-1} = I \quad \text{on} \quad T(X).$$

Next, for $x \in X$ consider $y = Tx$, then $T^{-1}(y) = x$, hence

$$T^{-1}(Tx) = x \quad \forall\, x \in X$$
$$\Updownarrow$$
$$T^{-1}T = I \quad \text{on} \quad X.$$

So T^{-1} is the inverse of T.

$\boxed{\text{Step 2.}}$ Assume that $S : T(X) \to X$ is also the inverse of T. Then, $TS = I$ and $T^{-1}T = I$ on $T(X)$, hence $T(S - T^{-1}) = 0$ on $T(X)$, that is,

$$T(S - T^{-1})y = T(Sy - T^{-1}y) = 0 \quad \forall\, y \in T(X).$$

Since T is one-to-one and $T(0) = 0$, we get $Sy - T^{-1}y = 0 \iff Sy = T^{-1}y\ \forall\, y \in T(X)$. So $S = T^{-1}$.

$\boxed{\text{Step 3.}}$ First, note that $T(X)$ is a linear space. Indeed, if $y_1, y_2 \in T(X)$, then $y_1 = Tx_1,\ y_2 = Tx_2$ for some $x_1, x_2 \in X$ and $y_1+y_2 = Tx_1+Tx_2 = T(x_1+x_2)$, hence $y_1+y_2 \in T(X)$. Similarly, if $y \in T(X)$ and α is a scalar, then $\alpha y \in T(X)$.

Let us prove that T^{-1} is linear. Let $y_1 + y_2 \in T(X)$ and $x_1 = T^{-1}y_1,\ x_2 = T^{-1}y_2,\ z = T^{-1}(y_1 + y_2)$ $(x_1, x_2, z \in X)$, then

$$Tz = y_1 + y_2 = Tx_1 + Tx_2 = T(x_1 + x_2).$$

Since T is one-to-one, $z = x_1 + x_2$, that is,

$$T^{-1}(y_1 + y_2) = T^{-1}y_1 + T^{-1}y_2 \quad \forall\, y_1, y_2 \in T(X).$$

Similarly, if $y \in T(X)$ and α is a scalar, then

$$T^{-1}(\alpha y) = \alpha T^{-1}y.$$

$\square$

Exercise. Prove this equality.

Exercise. Prove that if $T : X \to Y$ is linear and $Tx = 0$ only if $x = 0$, then T is one-to-one.

Theorem 2.3.1. *Let X, Y be linear spaces and $T : X \to Y$ be a linear operator. Then the inverse operator $T^{-1} : T(X) \to X$ exists and is bounded on $T(X)$ if and only if there exists $m > 0$ such that*

$$\|Tx\|_Y \geq m\|x\|_X \quad \forall\, x \in X. \tag{2.11}$$

Proof. $\boxed{\text{Step 1.}}$ Let (2.11) be satisfied. If $Tx = 0$, then by (2.11) $x = 0$. Hence, T is one-to-one and by Lemma 2.3.2 the inverse operator $T^{-1} : T(X) \to X$ exists, is uniquely defined and is linear.

Let $y \in T(X)$. Take as in (2.11) $x = T^{-1}y$. Then

$$\|y\|_Y \geq m\|T^{-1}y\|_X$$

and

$$\|T^{-1}\|_{T(X)\to X} = \sup_{\substack{y\in T(X)\\ y\neq 0}} \frac{\|T^{-1}y\|_X}{\|y\|_Y} \leq \frac{1}{m} < \infty.$$

So T^{-1} is bounded.

$\boxed{\text{Step 2.}}$ Let the inverse operator $T^{-1} : T(X) \to X$ exist and be bounded. Then

$$\|x\|_X = \|T^{-1}Tx\|_X \leq \|T^{-1}\|_{T(X)\to X}\|Tx\|_Y.$$

Hence inequality (2.11) is satisfied with $m = (\|T^{-1}\|_{T(X)\to X})^{-1}$. (Note that $T^{-1} \neq 0$, otherwise $T^{-1}T = 0$, hence $\|T^{-1}\|_{T(X)\to X} > 0$.) $\square$

Lemma 2.3.3. *Let X, Y be linear spaces and $T : X \to Y$ be a one-to-one bounded linear operator. Then the inverse operator $T^{-1} : T(X) \to X$ exists, is uniquely defined and*

$$\|T^{-1}\|_{T(X)\to X} \geq \frac{1}{\|T\|_{X\to Y}}. \tag{2.12}$$

Proof. By Lemma 2.3.2, $T^{-1} : T(X) \to X$ exists and is uniquely defined and is linear. If it is unbounded, then inequality (2.12) is trivial. If it is bounded, then by Lemma 2.3.1,

$$\begin{aligned} 1 = \|I\|_{X\to X} &= \|T^{-1}T\|_{X\to X} \leq \|T\|_{X\to T(X)}\|T^{-1}\|_{T(X)\to X} \\ &= \|T^{-1}\|_{T(X)\to X}\|T\|_{X\to Y}, \end{aligned}$$

which implies inequality (2.12). $\square$

2.4 Equivalent and non-equivalent norms

Definition 2.4.1. *Let a linear space X be equipped with two norms $\|\cdot\|_1$ and $\|\cdot\|_2$. They are said to be equivalent if there exist $c_1, c_2 > 0$ such that*

$$c_1\|x\|_1 \leq \|x\|_2 \leq c_2\|x\|_1 \quad \forall\, x \in X. \tag{2.13}$$

Remark. The equality $\lim_{x\to\infty} x_k = x$ holds with respect to the norm $\|\cdot\|_1$ if and only if it holds with respect to the norm $\|\cdot\|_2$ equivalent to $\|\cdot\|_1$.

A sequence $\{x_n\}_{n\in\mathbb{N}}$ is a Cauchy sequence with respect to the norm $\|\cdot\|_1$ if and only if it is a Cauchy sequence with respect to the norm $\|\cdot\|_2$ equivalent to $\|\cdot\|_1$.

Exercise. Prove these statements.

Lemma 2.4.1. *Let a linear space X be equipped with three norms $\|\cdot\|_1$, $\|\cdot\|_2$ and $\|\cdot\|_3$. If the norm $\|\cdot\|_1$ is equivalent to the norm $\|\cdot\|_2$ and the norm $\|\cdot\|_2$ is equivalent to the norm $\|\cdot\|_3$, then the norm $\|\cdot\|_1$ is equivalent to the norm $\|\cdot\|_3$.*

Proof. If for some $c_1, c_2 > 0$ inequality (2.13) is satisfied and for some $c_3, c_4 > 0$ the inequality

$$c_3\|x\|_2 \le \|x\|_3 \le c_4\|x\|_2 \quad \forall\, x \in X$$

is satisfied, then

$$c_1\|x\|_1 \le \|x\|_2 \le \frac{1}{c_3}\|x\|_3 \le \frac{c_4}{c_3}\|x\|_2 \le \frac{c_2c_4}{c_3}\|x\|_1 \quad \forall\, x \in X$$

and

$$c_1c_3\|x\|_1 \le \|x\|_3 \le c_2c_4\|x\|_1 \quad \forall\, x \in X,$$

which means that the norms $\|\cdot\|_1$ and $\|\cdot\|_3$ are equivalent. □

Theorem 2.4.1. *In a finite-dimensional linear space all norms are equivalent.*

Proof. Assume that X is a real linear space of dimension n. (The case of complex linear spaces are similar.) Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be norms on X.

Let $e_1, \dots, e_n$ be a basis in X, then $\forall\, x \in X$

$$x = \sum_{k=1}^{n} \alpha_k e_k,$$

where $\alpha_k \in \mathbb{R},\ k = 1, \dots, n$, are uniquely defined.

Let us introduce one more norm on X, namely,

$$\|x\|_0 = \Big(\sum_{k=1}^{n} |\alpha_k|^2\Big)^{\frac{1}{2}} \equiv |\alpha|.$$

By Lemma 2.4.1, it suffices to prove that the norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent to $\|\cdot\|_0$. Let us prove that the norms $\|\cdot\|_1$ and $\|\cdot\|_0$ are equivalent. (The proof of the equivalence of $\|\cdot\|_2$ and $\|\cdot\|_0$ is similar.)

Consider the function $f : \mathbb{R}^n \to \mathbb{R}$ defined by

$$f(\alpha) \equiv f(\alpha_1, \dots, \alpha_n) = \Big\|\sum_{k=1}^{n} \alpha_k e_k\Big\|_1 \quad \forall\, \alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n.$$

By the reverse and direct triangle inequalities

$$\begin{aligned}
|f(\alpha) - f(\beta)| &= |f(\alpha_1, \dots, \alpha_n) - f(\beta_1, \dots, \beta_n)| \\
= \Big|\Big\|\sum_{k=1}^{n} \alpha_k e_k\Big\|_1 - \Big\|\sum_{k=1}^{n} \beta_k e_k\Big\|_1\Big| &\le \Big\|\sum_{k=1}^{n} \alpha_k e_k - \sum_{k=1}^{n} \beta_k e_k\Big\|_1 \\
= \Big\|\sum_{k=1}^{n} (\alpha_k - \beta_k) e_k\Big\|_1 &\le \sum_{k=1}^{n} |\alpha_k - \beta_k|\|e_k\|_1 \quad \forall\, \alpha, \beta \in \mathbb{R}^n.
\end{aligned}$$

Hence, f is continuous on $\mathbb{R}^n$, because this inequality implies that $\forall\,\beta \in \mathbb{R}^n$

$$\lim_{\alpha\to\beta} |f(\alpha) - f(\beta)| = 0.$$

Let S^{n-1} be the unit sphere in $\mathbb{R}^n$. Note that $f(\alpha) > 0$ for any $\alpha \in S^{n-1}$, because $f(\alpha) \geq 0$ for any $\alpha \in \mathbb{R}^n$ and $f(\alpha) = 0$ only for $\alpha = 0$. Indeed,

$$f(\alpha) = 0 \iff \Big\| \sum_{k=1}^{n} \alpha_k e_k \Big\|_1 = 0 \iff \sum_{k=1}^{n} \alpha_k e_k = 0 \iff \alpha_1 = \ldots = \alpha_n = 0,$$

because the elements $e_1, \ldots, e_n$ are linearly independent.

Since S^{n-1} is a compact set in $\mathbb{R}^n$ and f is continuous on S^{n-1} there exist $\beta, \gamma \in S^{n-1}$ such that

$$\begin{aligned} 0 < m = f(\beta_1, \ldots, \beta_n) = \min_{\alpha\in S^{n-1}} f(\alpha_1, \ldots, \alpha_n) \leq \max_{\alpha\in S^{n-1}} f(\alpha_1, \ldots, \alpha_n) \\ = f(\gamma_1, \ldots, \gamma_n) = M < \infty. \end{aligned}$$

Note that $\forall\, x \in X,\ x \neq 0$

$$\frac{\|x\|_1}{\|x\|_0} = \frac{\Big\| \sum_{k=1}^{n} \alpha_k e_k \Big\|_1}{|\alpha|} = \Big\| \sum_{k=1}^{n} \frac{\alpha_k}{|\alpha|} e_k \Big\|_1 = f\Big(\frac{\alpha_1}{|\alpha|}, \cdots, \frac{\alpha_n}{|\alpha|}\Big)$$

and $\big(\frac{\alpha_1}{|\alpha|}, \cdots, \frac{\alpha_n}{|\alpha|}\big) \in S^{n-1}$, because

$$\sum_{k=1}^{n} \Big(\frac{\alpha_k}{|\alpha|}\Big)^2 = \frac{1}{|\alpha|^2} \sum_{k=1}^{n} \alpha_k^2 = 1.$$

Therefore, by the above inequality

$$\begin{gathered} m \leq \frac{\|x\|_1}{\|x\|_0} \leq M \quad \forall\, x \in X,\ x \neq 0 \\ \Updownarrow \\ m\|x\|_0 \leq \|x\|_1 \leq M\|x\|_0 \quad \forall\, x \in X \end{gathered}$$

since for $x = 0$ this inequality is trivial.

So the norms $\|\cdot\|_1$ and $\|\cdot\|_0$ are equivalent, which completes the proof. □

Next, we shall give several examples of norms which are not equivalent on infinite-dimensional linear spaces.

Consider the space of sequences ℓ_p where $0 < p \leq \infty$. Recall that for $1 \leq p \leq \infty$ it is a normed space and for $0 < p < 1$ it is a quasi-normed space.

Lemma 2.4.2 (Jensen's inequality). *Let $0 < p \leq q \leq \infty$. Then*

$$\ell_p \subset \ell_q$$

and

$$\|a\|_{\ell_q} \leq \|a\|_{\ell_p} \quad \forall\, a \in \ell_p. \tag{2.14}$$

Proof. First, let $0 < p < \infty$ and $\|a\|_{\ell_p} = 1$. Then

$$|a_k| \le \Big(\sum_{m=1}^{\infty} |a_m|^p\Big)^{\frac{1}{p}} = \|a\|_{\ell_p} = 1 \quad \forall\, k \in \mathbb{N}.$$

Hence, for $0 < p < q < \infty$

$$|a_k|^{q-p} \le 1 \implies |a_k|^q \le |a_k|^p \quad \forall\, k \in \mathbb{N}.$$

Therefore,

$$\sum_{k=1}^{\infty} |a_k|^q \le \sum_{k=1}^{\infty} |a_k|^p = 1$$

and

$$\|a\|_{\ell_q} \le 1 = \|a\|_{\ell_p}.$$

If $\|a\|_{\ell_p} \ne 1$ and $\|a\|_{\ell_q} \ne 0$, then consider the sequence

$$b_k = \frac{a_k}{\|a\|_{\ell_p}} \quad \forall\, k \in \mathbb{N}.$$

Note that $\|b\|_{\ell_p} = 1$, hence, as already proven

$$\|b\|_{\ell_q} \le 1 \iff \frac{\|a\|_{\ell_q}}{\|a\|_{\ell_p}} \le 1 \iff \|a\|_{\ell_q} \le \|a\|_{\ell_p}.$$

□

Example. Let

$$X = \{a = \{a_k\}_{k\in\mathbb{N}} : \exists\, k_0 \in \mathbb{N} : a_k = 0\ \forall\, k > k_0\},$$

$1 \le p, q \le \infty,\ p \ne q$. The norms $\|a\|_{\ell_p}$ and $\|a\|_{\ell_q}$ are not equivalent on X.
Assume, without loss of generality, that $p < q$. Then, for any $c > 0$ the inequality

$$c\|a\|_{\ell_p} \le \|a\|_{\ell_q} \quad \forall\, a \in X \tag{2.15}$$

does not hold.

For any $m \in \mathbb{N}$ consider the sequence

$$a_k^{(m)} = \begin{cases} k^{-\frac{1}{p}} & k = 1, \dots, m \\ 0 & k = m+1, \dots \end{cases},$$

then,

$$\|a^{(m)}\|_{\ell_q} \le \Big(\sum_{k=1}^{\infty} k^{-\frac{q}{p}}\Big)^{\frac{1}{q}} = A < \infty$$

and

$$\|a^{(m)}\|_{\ell_p} \le \Big(\sum_{k=1}^{m} \frac{1}{k}\Big)^{\frac{1}{p}}.$$

If (2.15) holds, then

$$c\Big(\sum_{k=1}^{m} \frac{1}{k}\Big)^{\frac{1}{p}} \le A,$$

which is impossible since

$$\lim_{m\to\infty} \sum_{k=1}^{m} \frac{1}{k} = \sum_{k=1}^{\infty} \frac{1}{k} = \infty.$$

Example. Let $X = L_\infty(0,1),\ 1 \le p, q \le \infty,\ p \ne q$. Then the norms $\|f\|_{L_p(0,1)}$ and $\|f\|_{L_q(0,1)}$ are not equivalent on X. Indeed, assume, without loss of generality, that $p < q$. By Hölder's inequality with the parameter $\frac{q}{p} > 1$ we get

$$\|f\|_{L_p(0,1)} = \Big(\int_0^1 |f(x)|^p \cdot 1\,dx\Big)^{\frac{1}{p}} \le \Big(\|f\|_{L_{\frac{q}{p}}(0,1)}\|1\|_{L_{(\frac{q}{p})'}(0,1)}\Big)^{\frac{1}{p}}$$

$$= \bigg(\Big(\int_0^1 (|f(x)|^p)^{\frac{q}{p}}\,dx\Big)^{\frac{p}{q}}\Big(\int_0^1 1\,dx\Big)^{1-\frac{p}{q}}\bigg)^{\frac{1}{p}} = \Big(\int_0^1 |f(x)|^q\,dx\Big)^{\frac{1}{q}} = \|f\|_{L_q(0,1)}.$$

Let us assume that for some $c_1 > 0$

$$c_1\|f\|_{L_p(0,1)} \le \|f\|_{L_q(0,1)} \quad \forall\, f \in L_\infty(0,1). \tag{2.16}$$

Consider for any $k \in \mathbb{N}$ the functions

$$f_k(x) = \begin{cases} x^{-\frac{1}{q}} & \frac{1}{k} \le x \le 1 \\ 0 & 0 < x < \frac{1}{k}. \end{cases}$$

Then,

$$\|f_k\|_{L_p(0,1)} \le \Big(\int_0^1 x^{-\frac{p}{q}}\,dx\Big)^{\frac{1}{p}} = A < \infty$$

since $\frac{p}{q} < 1$. Moreover,

$$\|f_k\|_{L_q(0,1)} = \Big(\int_{\frac{1}{k}}^1 \frac{dx}{x}\Big)^{\frac{1}{q}} = \Big(\ln 1 - \ln\frac{1}{k}\Big)^{\frac{1}{q}} = (\ln k)^{\frac{1}{q}}.$$

So $\forall\, k \in \mathbb{N}$ by (2.16),

$$c_1(\ln k)^{\frac{1}{q}} \le A,$$

which is impossible. Hence, the norms $\|f\|_{L_p(0,1)}$ and $\|f\|_{L_q(0,1)}$ are not equivalent on $L_\infty(0,1)$.

Example. Consider the space $X = C_0^\infty$ of all infinitely continuously differentiable functions on $\mathbb{R}^n$ with compact support, that is, $f \in C_0^\infty \iff f \in C^\infty(\mathbb{R})$ and there exists $a > 0$ (dependent on f) such that $f(x) = 0$ if $|x| \ge a$. Then for any $1 \le p, q \le \infty,\ p \ne q$ the norms $\|f\|_{L_p(\mathbb{R}^n)}$ and $\|f\|_{L_q(\mathbb{R}^n)}$ are not equivalent.

Indeed, assume that for some $c_1, c_2 > 0\ \forall\, f \in C_0^\infty(\mathbb{R})$

$$c_1\|f\|_{L_q(\mathbb{R}^n)} \le \|f\|_{L_p(\mathbb{R}^n)} \le c_2\|f\|_{L_q(\mathbb{R}^n)}.$$

Let $g \in C_0^\infty(\mathbb{R}^n),\ g \not\equiv 0$, and $g_k(x) = g(kx),\ k \in \mathbb{N},\ x \in \mathbb{R}^n$. Then $g_k \in C_0^\infty(\mathbb{R}^n)$, hence

$$c_1\|g(kx)\|_{L_q(\mathbb{R}^n)} \le \|g(kx)\|_{L_p(\mathbb{R}^n)} \le c_2\|g(kx)\|_{L_q(\mathbb{R}^n)}. \tag{2.17}$$

Note that

$$\|g(kx)\|_{L_p(\mathbb{R}^n)} = \Big(\int_{\mathbb{R}^n} |g(kx)|^p\,dx\Big)^{\frac{1}{p}} = \Big(x = \frac{y}{k}\Big)$$

$$= k^{-\frac{n}{p}}\Big(\int_{\mathbb{R}^n} |g(y)|^p\,dy\Big)^{\frac{1}{p}} = k^{-\frac{n}{p}}\|g\|_{L_p(\mathbb{R}^n)}.$$

Similarly,

$$\|g(kx)\|_{L_q(\mathbb{R}^n)} = k^{-\frac{n}{p}}\|g\|_{L_q(\mathbb{R}^n)}.$$

Hence, (2.16) implies that $\forall\, k \in \mathbb{N}$

$$c_1 k^{-\frac{n}{p}} \|g\|_{L_q(\mathbb{R}^n)} \le k^{-\frac{n}{p}} \|g\|_{L_p(\mathbb{R}^n)} \le c_2 k^{-\frac{n}{p}} \|g\|_{L_q(\mathbb{R}^n)}$$
$$\Updownarrow$$
$$c_1 \frac{\|g\|_{L_q(\mathbb{R}^n)}}{\|g\|_{L_p(\mathbb{R}^n)}} \le k^{n\left(\frac{1}{q}-\frac{1}{p}\right)} \le c_2 \frac{\|g\|_{L_q(\mathbb{R}^n)}}{\|g\|_{L_p(\mathbb{R}^n)}}.$$

If $q < p$, then the right-hand-side inequality is impossible, and if $q > p$, then the left-hand-side inequality is impossible.

2.5 Existence of unbounded linear functionals

Let X be a normed space. Recall that an operator $f : X \to \mathbb{R}$ or $f : X \to \mathbb{C}$ is called a functional. Consequently, the norm of a functional has the form

$$\|f\| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{|f(x)|}{\|x\|_X}.$$

Lemma 2.5.1. *Let X be a finite-dimensional normed space. Then any linear functional on X is bounded.*

Proof. Let $\dim X = n$ and $e_1, \dots, e_n$ be a basis of X. Then for any $x \in X$ there exist uniquely defined real (complex) numbers, if X is a linear space with multiplication by real (complex) numbers, such that

$$x = \sum_{k=1}^{n} \alpha_k e_k.$$

Define

$$\|x\|_0 = \sum_{k=1}^{n} |\alpha_k|.$$

Clearly, it defines a norm on X. Moreover, $\|x\|_0$ is equivalent to the norm $\|\cdot\|_X$, hence there exist $c_1, c_2 > 0$ such that

$$c_1 \|x\|_0 \le \|x\|_X \le c_2 \|x\|_0 \quad \forall\, x \in X.$$

Then,

$$\|f\| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{|f(x)|}{\|x\|_X} \le \frac{1}{c_1} \sup_{\substack{x \in X \\ x \neq 0}} \frac{|f(x)|}{\|x\|_0},$$

where

$$\begin{aligned} |f(x)| &= \left| f\Big(\sum_{k=1}^{n} \alpha_k e_k \Big) \right| = \Big| \sum_{k=1}^{n} \alpha_k f(e_k) \Big| \le \sum_{k=1}^{n} |\alpha_k| |f(e_k)| \\ &\le \Big(\sum_{k=1}^{n} |\alpha_k| \Big) \max_{k=1,\dots,n} |f(e_k)| = \|x\|_0 \max_{k=1,\dots,n} |f(e_k)|. \end{aligned}$$

So

$$\|f\| \le \frac{1}{c_1} \max_{k=1,\dots,n} |f(e_k)| < \infty.$$

□

Remark. A similar statement holds for any linear operator $T : X \to Y$, where X and Y are normed spaces and $\dim X < \infty$.

Exercise. Prove this.

The goal of this section is to prove that for infinite-dimensional normed spaces X, unbounded linear functionals exist.

Definition 2.5.1. *Let X be a linear space with multiplication by elements of a field F, in particular, real or complex numbers.*
A subset $B \subset X$ is called a Hamel basis for X if every $x \in X$, $x \neq \theta$ (θ is the zero element in X) can be uniquely represented as a finite linear combination of elements in B with non-zero coefficients.

So for every $x \in X$, $x \neq \theta$, there exists $n = n(x)$, $e_1 = e_1(x), \dots, e_n = e_n(x) \in B$ and scalars $\alpha_1 = \alpha_1(x), \dots, \alpha_n = \alpha_n(x) \in F$ such that

$$x = \sum_{k=1}^{n} \alpha_k e_k.$$

Moreover, $\alpha_1, \dots, \alpha_n$ are defined uniquely.

Lemma 2.5.2. *Let $B \subset X$ be a Hamel basis for a linear space X. Then, the elements of B are linearly independent.*

Proof. Let $n \in \mathbb{N}$, $e_1, \dots, e_n \in B$ and $\alpha_1, \dots, \alpha_n \in F$ be such that

$$\sum_{k=1}^{n} \alpha_k e_k = \theta.$$

Let $\alpha \in F$ be such that $\alpha \neq 0$, $\alpha \neq \alpha_1, \dots, \alpha \neq \alpha_n$. Consider the element

$$x = \sum_{k=1}^{n} \alpha e_k \in X.$$

Then, we also have that

$$x = \sum_{k=1}^{n} (\alpha + \alpha_k) e_k.$$

By the uniqueness of the representation for x

$$a = a + \alpha_k, \ k = 1, \dots, n \implies \alpha_1 = \ldots = \alpha_n = 0.$$

□

Theorem 2.5.1. *Each linear space has a Hamel basis.*

Without proof. The proof is based on the axiom of choice.

Theorem 2.5.2. *Let X be an infinite-dimensional normed space. Then there exists a linear functional $f : X \to \mathbb{R}$ which is not bounded ($\iff$ not continuous).*

Proof. Let B be a Hamel basis for X. Since X is infinite-dimensional, B is also infinite-dimensional, hence

$$B \supset \{e_k\}_{k=1}^{\infty}.$$

Let us define a linear functional f on X in the following way. Let

$$f(e_k) = k\|e_k\|_X, \quad k \in \mathbb{N},$$
$$f(e) = 1 \quad \forall\, e \in B \setminus \{e_k\}_{k=1}^{\infty}$$

and for any other x, f is defined by linearity: if $x = \sum\limits_{k=1}^{n} \alpha_k \psi_k$, where $\psi_k \in B$ and $\alpha_k \in \mathbb{R}$ (or $\mathbb{C}$), $k = 1, \ldots, n$, then

$$f(x) = \sum_{k=1}^{n} \alpha_k f(\psi_k).$$

Let $y \in X$, then $y = \sum\limits_{i=1}^{m} \beta_i \chi_i$, where $\chi_i \in B$, $\beta_i \in \mathbb{R}$ (or $\mathbb{C}$), $i = 1, \ldots, m$, and

$$f(y) = \sum_{i=1}^{m} \beta_i f(\chi_i).$$

Assume that some of $\psi_1, \ldots, \psi_n$ coincide with some of $\chi_1, \ldots, \chi_m$, say $\psi_1 = \chi_1, \ldots, \psi_l = \chi_l$, $l \le \min\{n, m\}$, and all other $\psi_{l+1}, \ldots, \psi_n$ do not coincide with any of $\chi_{l+1}, \ldots, \chi_m$. Then

$$\begin{aligned} x + y &= \sum_{k=1}^{l} \alpha_k \psi_k + \sum_{k=l+1}^{n} \alpha_k \psi_k + \sum_{i=1}^{l} \beta_i \chi_i + \sum_{i=l+1}^{m} \beta_i \chi_i \\ &= \sum_{k=1}^{l} (\alpha_k + \beta_k) \psi_k + \sum_{k=l+1}^{n} \alpha_k \psi_k + \sum_{i=l+1}^{m} \beta_i \chi_i \end{aligned}$$

and, by the definition of f,

$$\begin{aligned} f(x+y) &= \sum_{k=1}^{l} (\alpha_k + \beta_k) f(\psi_k) + \sum_{k=l+1}^{n} \alpha_k f(\psi_k) + \sum_{i=l+1}^{m} \beta_i f(\chi_i) \\ &= \sum_{k=1}^{l} \alpha_k f(\psi_k) + \sum_{k=1}^{l} \beta_k f(\psi_k) + \sum_{k=l+1}^{n} \alpha_k f(\psi_k) + \sum_{i=l+1}^{m} \beta_i f(\chi_i) \\ &= \sum_{k=1}^{n} \alpha_k f(\psi_k) + \sum_{i=1}^{m} \beta_i f(\chi_i) = f(x) + f(y). \end{aligned}$$

Next, let $\gamma \in \mathbb{R}$ (or $\mathbb{C}$), then

$$\gamma x = \sum_{k=1}^{n} \gamma \alpha_k \psi_k$$

and, by the definition of f,

$$f(\gamma x) = \sum_{k=1}^{n} \gamma \alpha_k f(\psi_k) = \gamma \sum_{k=1}^{n} \alpha_k f(\psi_k) = \gamma f(x).$$

So, f is a linear functional on X.
Next,

$$\|f\| = \sup_{\substack{x \in X \\ x \neq \theta}} \frac{|f(x)|}{\|x\|_X} \ge \sup_{k \in \mathbb{N}} \frac{|f(e_k)|}{\|e_k\|_X} = \sup_{k \in \mathbb{N}} k = \infty.$$

So, f is unbounded. $\square$

Next, we consider functions $f : \mathbb{R} \to \mathbb{R}$ satisfying the additivity equation

$$f(x+y) = f(x) + f(y) \quad \forall\, x, y \in \mathbb{R}. \tag{2.18}$$

Exercise 1. Prove that any twice differentiable function f, satisfying (2.18) has the form: for some $k \in \mathbb{R}$

$$f(x) = kx \quad \forall\, x \in \mathbb{R}. \tag{2.19}$$

Exercise 2. Prove that any differentiable function f, satisfying (2.18) has the form (2.19).

Exercise 3. Prove that any continuous function f, satisfying (2.18) has the form (2.19).

Step 1. Prove that $\forall\, n \in \mathbb{N}$

$$f(nx) = nf(x) \quad \forall\, x \in \mathbb{R}. \tag{2.20}$$

Step 2. Prove that $\forall\, r \in \mathbb{Q},\ r > 0$

$$f(rx) = rf(x) \quad \forall\, x \in \mathbb{R}. \tag{2.21}$$

Step 3. Prove that $\forall\, \alpha \in \mathbb{R},\ \alpha > 0$

$$f(\alpha x) = \alpha f(x). \tag{2.22}$$

Step 4. Prove that $f(0) = 0$ and $\forall\, \alpha \in \mathbb{R}$

$$f(\alpha x) = \alpha f(x). \tag{2.23}$$

Step 5. Prove formula (2.19).

Exercise 4. Prove, by using the proof of Theorem 2.5.2, that there exists a discontinuous function $f : \mathbb{R} \to \mathbb{R}$ satisfying (2.18).
Hint: Consider the linear space $\mathbb{R}$ with multiplication by rational numbers (in Definition 2.5.1, $X = \mathbb{R}$ and $F = \mathbb{Q}$).

Remark. One can also prove that any Lebesgue measurable function f, satisfying (2.18) has the form (2.19).

3. Inner product spaces

3.1 Definition and basic properties of inner product spaces

The notion of an inner product is a generalization of the notion of a scalar product (or dot product) of vectors in geometry, which can be applied to various, not necessarily geometric objects, say to sequences and functions.

Definition 3.1.1. *An inner product space is a pair $(X, (\cdot,\cdot))$, where X is a linear space and $(\cdot,\cdot)$ is an inner product on X, that is, a function on $X \times X$, whose values are real numbers in the case of a real linear space X, and complex numbers in the case of a complex linear space X, briefly*

$$(\cdot,\cdot) : X \times X \to \mathbb{R} \text{ (or } \mathbb{C}),$$

such that for all $x, y, x_1, x_2 \in X$ and $\alpha_1, \alpha_2 \in \mathbb{R}$ (or $\mathbb{C}$)

1) $(x, x) \geq 0$ *and* $(x, x) = 0$ *if and only if* $x = 0$,

2) $(x, y) = \overline{(y, x)}$ *(in a real linear space X, $(x, y) = (y, x)$)*,

3) $(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1(x_1, y) + \alpha_2(x_2, y)$.

Note that for a complex linear space X, for any $x, y_1, y_2 \in X$ and $\alpha_1, \alpha_2 \in \mathbb{C}$

$$\begin{aligned} (x, \alpha_1 y_1 + \alpha_2 y_2) &= \overline{(\alpha_1 y_1 + \alpha_2 y_2, x)} = \overline{\alpha_1(y_1, x) + \alpha_2(y_2, x)} \\ &= \overline{\alpha_1}\,\overline{(y_1, x)} + \overline{\alpha_2}\,\overline{(y_2, x)} = \overline{\alpha_1}(x, y_1) + \overline{\alpha_2}(x, y_2) \end{aligned}$$

and for a real space X for any $x, y_1, y_2 \in X$ and $\alpha_1, \alpha_2 \in \mathbb{R}$

$$(x, \alpha_1 y_1 + \alpha_2 y_2) = \alpha_1(x, y_1) + \alpha_2(x, y_2).$$

For this reason an inner product is said to be of sesquilinear form (linear in the first variable and “half”-linear in the second variable).

Instead of $(X, (\cdot,\cdot))$ we shall often write just X if there is no danger of confusion.

If instead of 1) we have only

1′) $(x, x) \geq 0$ and $(0, 0) = 0$,

then the space $(X, (\cdot,\cdot))$ is called a semi-inner product space.

Example 3.1. Let $X = \mathbb{R}^3$ and for any vectors $x, y \in \mathbb{R}^3$

$$(x, y) = x \cdot y \equiv |x||y| \cos\alpha, \tag{3.1}$$

where α is the angle between x and y. This is the standard inner product in $\mathbb{R}^3$.

Example 3.2. Let $n \in \mathbb{N},\ X = \mathbb{R}^n$ and for any $x = (x_1, \dots, x_n),\ y = (y_1, \dots, y_n) \in \mathbb{R}^n$

$$(x, y) = \sum_{k=1}^{n} x_k y_k. \tag{3.2}$$

Example 3.3. Let $n \in \mathbb{N},\ X = \mathbb{C}^n$ and for any $x = (x_1, \dots, x_n),\ y = (y_1, \dots, y_n) \in \mathbb{C}^n$

$$(x, y) = \sum_{k=1}^{n} x_k \overline{y_k}. \tag{3.3}$$

Example 3.4. Let $X = \ell_2$ and for any $x = (x_1, \dots, x_n, \dots),\ y = (y_1, \dots, y_n, \dots) \in \ell_2$

$$(x, y) = \sum_{k=1}^{\infty} x_k \overline{y_k}. \tag{3.4}$$

(Note that by the Cauchy-Bunyakowski inequality for all $x, y \in \ell_2$,
$|(x, y)| = \Big(\sum\limits_{k=1}^{\infty} |x_k|^2\Big)^{\frac{1}{2}} \Big(\sum\limits_{k=1}^{\infty} |y_k|^2\Big)^{\frac{1}{2}} < \infty$.)
It can easily be verified that the expressions defined by formulas (3.2), (3.3), (3.4) are inner products.

Example 3.5. Let $\Omega \subset \mathbb{R}^n$ be a Lebesgue measurable set, $X = L_2(\Omega)$ and for any $f, g \in L_2(\Omega)$

$$(f, g) = \int_{\Omega} f(x)\overline{g(x)}\, dx. \tag{3.5}$$

(This integral exists and is finite, because by the Cauchy-Bunyakowski inequality
$\int\limits_{\Omega} |f(x)\overline{g(x)}|\, dx \le \Big(\int\limits_{\Omega} |f(x)|^2\, dx\Big)^{\frac{1}{2}} \Big(\int\limits_{\Omega} |g(x)|^2\, dx\Big)^{\frac{1}{2}} < \infty$.)

The expression defined by formula (3.5) is a semi-inner product, because $(f, f) = \int\limits_{\Omega} |f(x)|^2\, dx = 0$ holds if and only if f is equivalent to 0 on Ω (not f equal to 0 on Ω). Following the tradition it is usually said that it is an inner product, because the distinction between f and 0 is negligible.

Lemma 3.1.1. *Let $(X, (\cdot, \cdot))$ be an inner product space and $\|x\| = \sqrt{(x, x)}$ for any $x \in X$. Then for any $x, y \in X$*

$$\|x + y\|^2 = \|x\|^2 + 2\,\mathrm{Re}(x, y) + \|y\|^2. \tag{3.6}$$

Proof. Indeed,

$$\begin{aligned} \|x + y\|^2 &= (x + y, x + y) = (x, x + y) + (y, x + y) = (x, x) + (x, y) + (y, x) + (y, y) \\ &= (x, x) + (x, y) + \overline{(x, y)} + (y, y) = \|x\|^2 + 2\,\mathrm{Re}(x, y) + \|y\|^2. \end{aligned}$$

□

Definition 3.1.2. *Let $(X, (\cdot, \cdot))$ be an inner product space. Then the elements $x, y \in X$ are said to be orthogonal, briefly $x \perp y$, if $(x, y) = 0$.*

Lemma 3.1.2 (Pythagorean theorem). *Let $(X, (\cdot, \cdot))$ be an inner product space and $\|x\| = \sqrt{(x, x)}$ for any $x \in X$. If $x, y \in X$ and $x \perp y$, then*

$$\|x + y\|^2 = \|x\|^2 + \|y\|^2. \tag{3.7}$$

Proof. Immediately follows by (3.6). □

Remark. In fact, equality (3.7) is equivalent to the equality $\operatorname{Re}(x,y)=0$.

Lemma 3.1.3. *Let $(X,(\cdot,\cdot))$ be an inner product space and $\|x\|=\sqrt{(x,x)}$ for any $x\in X$. Then, for any $x,y\in X$ with $y\neq 0$*

$$\left\|x-\frac{(x,y)}{\|y\|^2}y\right\|^2=\|x\|^2-\frac{|(x,y)|^2}{\|y\|^2}. \tag{3.8}$$

Proof. By definition, the projection $\operatorname{Pr}_y x$ of a vector x onto a vector y is the vector $\lambda y,\ \lambda\in\mathbb{R}$ such that $x-\lambda y\perp y$, which means that

$$(x-\lambda y,y)=(x,y)-\lambda(y,y)=0 \iff \lambda=\frac{(x,y)}{\|y\|^2}$$

and

$$\operatorname{Pr}_y x=\frac{(x,y)}{\|y\|^2}y. \tag{3.9}$$

Since $x=x-\operatorname{Pr}_y x+\operatorname{Pr}_y x$ and $x-\operatorname{Pr}_y x\perp\operatorname{Pr}_y x$, by the Pythagorean theorem

$$\|x\|^2=\|x-\operatorname{Pr}_y x\|^2+\|\operatorname{Pr}_y x\|^2$$

or

$$\|x\|^2=\left\|x-\frac{(x,y)}{\|y\|^2}y\right\|^2+\frac{|(x,y)|^2}{\|y\|^2}. \tag{3.10}$$

which is equivalent to (3.8).

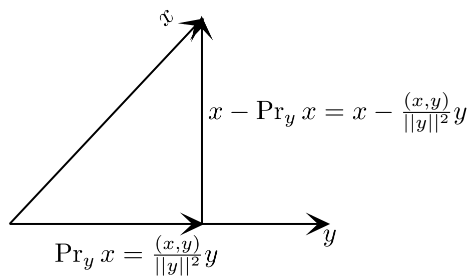


Figure 3.1 Geometric illustration of equality (3.10).

□

Lemma 3.1.4 (Cauchy-Schwarz-Bunyakowski inequality). *Let $(X,(\cdot,\cdot))$ be an inner product space and $\|x\|=\sqrt{(x,x)}$. Then $\forall\,x,y\in X$*

$$|(x,y)|\le\|x\|\cdot\|y\|. \tag{3.11}$$

Proof. If $y=0$, the inequality is obvious. If $y\neq 0$, then by Lemma 3.1.4

$$\|x\|^2-\frac{|(x,y)|^2}{\|y\|^2}\ge 0 \iff |(x,y)|^2\le\|x\|^2\|y\|^2 \iff (3.10).$$

□

Lemma 3.1.5. *Any inner product space $(X, (\cdot,\cdot))$ is a normed space $(X, \|\cdot\|)$ with the norm $\|\cdot\|$ defined for any $x \in X$ by the following equality:*

$$\|x\| = \sqrt{(x, x)}. \tag{3.12}$$

Proof. It suffices to verify the validity of the three properties of a norm.

1) $\|x\| \geq 0,\ \|x\| = 0 \iff (x, x) = 0 \iff x = 0,$

2) $\|\alpha x\| = \sqrt{(\alpha x, \alpha x)} = \sqrt{\alpha \cdot \overline{\alpha}(x, x)} = |\alpha|\|x\|,$

3) By (3.11),

$$\begin{aligned}\|x+y\|^2 &= \|x\|^2 + 2\,\mathrm{Re}(x, y) + \|y\|^2 \\ &\leq \|x\|^2 + 2|(x, y)| + \|y\|^2 \\ &\leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2 \\ &= (\|x\| + \|y\|)^2,\end{aligned}$$

hence,

$$\|x + y\| \leq \|x\| + \|y\|.$$

□

Lemma 3.1.6 (Continuity of the inner product). *For any inner product space $(X, (\cdot,\cdot))$, the inner product $(\cdot,\cdot) : X \times X \to \mathbb{C}$ is a continuous function on $X \times X$.*

Proof. By (3.11), for any $x_k, y_k \in X,\ k \in \mathbb{N},\ x, y \in X$

$$\begin{aligned}|(x_k, y_k) - (x, y)| &= |(x_k - x, y_k) + (x, y_k - y)| \leq |(x_k - x, y_k)| + |(x, y_k - y)| \\ &\leq \|x_k - x\| \cdot \|y_k\| + \|x\| \cdot \|y_k - y\|.\end{aligned} \tag{3.13}$$

If $x_k \to x$ and $y_k \to y$ in X, then the sequence $\{\|y_k\|\}$ is bounded, hence by the above inequality

$$\lim_{k\to\infty} |(x_k, y_k) - (x, y)| = 0 \iff \lim_{k\to\infty} (x_k, y_k) = (x, y).$$

□

Lemma 3.1.7 (Parallelogram identity). *Let $(X, (\cdot,\cdot))$ be an inner product space. Then for any $x, y \in X$*

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2). \tag{3.14}$$

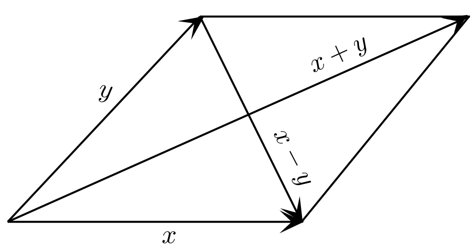


Figure 3.2 Parallelogram identity

Proof. By Lemma 3.1.1

$$\begin{aligned}\|x+y\|^2+\|x-y\|^2 &= \|x\|^2+2\operatorname{Re}(x,y)+\|y\|^2+\|x\|^2+2\operatorname{Re}(x,-y)+\|-y\|^2\\ &= 2\|x\|^2+2\|y\|^2.\end{aligned}$$

$\square$

Lemma 3.1.8. *Let $(X,(\cdot,\cdot))$ be an inner product space. Then for any $x,y\in X$*

$$\operatorname{Re}(x,y) = \frac{1}{4}(\|x+y\|^2-\|x-y\|^2), \tag{3.15}$$

$$\operatorname{Im}(x,y) = \frac{1}{4}(\|x+iy\|^2-\|x-iy\|^2). \tag{3.16}$$

Proof. By Lemma 3.1.1

$$\begin{aligned}\|x+y\|^2-\|x-y\|^2 &= \|x\|^2+2\operatorname{Re}(x,y)+\|y\|^2-(\|x\|^2+2\operatorname{Re}(x,-y)+\|-y\|^2)\\ &= 4\operatorname{Re}(x,y),\end{aligned}$$

which implies (3.15), and by (3.15) with y replaced by iy

$$\operatorname{Im}(x,y)=-\operatorname{Re}i(x,y)=\operatorname{Re}(x,iy)=\frac{1}{4}(\|x+iy\|^2-\|x-iy\|^2).$$

$\square$

Theorem 3.1.1 (Jordan-Neumann theorem, "inner productivity" of a normed space). *Let $(X,\|\cdot\|)$ be a normed space. In order that it be an inner product space $(X,(\cdot,\cdot))$ with the inner product $(\cdot,\cdot)$ satisfying for any $x,y\in X$ the equality (3.12) it is necessary and sufficient that for any $x,y\in X$ the parallelogram identity (3.14) is satisfied.*

Proof. Necessity. Follows from Lemma 3.1.7.

Sufficiency.

1. First assume that X is a real linear space. Taking into account formula (3.15) of Lemma 3.1.8, let us define the inner product for any $x,y\in X$ by the following equality

$$(x,y)=\frac{1}{4}(\|x+y\|^2-\|x-y\|^2). \tag{3.17}$$

It is required to prove that it is indeed an inner product, that is, it satisfies the three properties of an inner product.

1) By (3.17) $(x,x)=\|x\|^2\ge 0,\ (x,x)=0 \iff \|x\|=0 \iff x=0.$
2) By (3.17)

$$(y,x)=\frac{1}{4}(\|y+x\|^2-\|y-x\|^2)=\frac{1}{4}(\|y+x\|^2-\|x-y\|^2)=(x,y).$$

3) Let us prove the linearity with respect to to first variable.

$\boxed{\text{Step 1.}}$ By (3.17)

$$\begin{aligned}(2x,y) &= \frac{1}{4}(\|2x+y\|^2-\|2x-y\|^2)\\ &= \frac{1}{4}(\|x+x+y\|^2-\|x+x-y\|^2)\\ &= \frac{1}{4}(2\|x\|^2+2\|x+y\|^2-\|y\|^2\\ &\qquad -2\|x\|^2-2\|x-y\|^2+\|y\|^2)\\ &= 2\cdot\frac{1}{4}(\|x+y\|^2-\|x-y\|^2)=2(x,y).\end{aligned}$$

Replace x by $\frac{x}{2}$, then

$$(x,y) = 2\Big(\frac{x}{2}, y\Big).$$

So

$$(2x,y) = 2(x,y), \quad \Big(\frac{x}{2}, y\Big) = \frac{1}{2}(x,y). \tag{3.18}$$

$\boxed{\text{Step 2.}}$ By (3.17) and (3.18)

$$\begin{aligned}
(x_1 + x_2, y) &= \frac{1}{4}(\|x_1 + x_2 + y\|^2 - \|x_1 + x_2 - y\|^2)\\
&= \frac{1}{4}\Big(\Big\|x_1 + \frac{y}{2} + x_2 + \frac{y}{2}\Big\|^2 - \Big\|x_1 - \frac{y}{2} + x_2 - \frac{y}{2}\Big\|^2\Big)\\
&= \frac{1}{4}\Big(2\Big\|x_1 + \frac{y}{2}\Big\|^2 + 2\Big\|x_2 + \frac{y}{2}\Big\|^2 - \|x_1 - x_2\|^2\\
&\qquad - 2\Big\|x_1 - \frac{y}{2}\Big\|^2 - 2\Big\|x_2 - \frac{y}{2}\Big\|^2 + \|x_1 - x_2\|^2\Big)\\
&= 2\Big[\frac{1}{4}\Big(\Big\|x_1 + \frac{y}{2}\Big\|^2 - \Big\|x_1 - \frac{y}{2}\Big\|^2\Big)\\
&\qquad + \frac{1}{4}\Big(\Big\|x_2 + \frac{y}{2}\Big\|^2 - \Big\|x_2 - \frac{y}{2}\Big\|^2\Big)\Big]\\
&= 2\Big[\Big(x_1, \frac{y}{2}\Big) + \Big(x_2, \frac{y}{2}\Big)\Big] = (x_1, y) + (x_2, y). \qquad (3.19)
\end{aligned}$$

$\boxed{\text{Step 3.}}$ By (3.19) $\forall\, n \in \mathbb{N}$

$$(nx, y) = (\underbrace{x + \cdots + x}_{n}, y) = n(x,y).$$

Also,

$$(-x, y) = \frac{1}{4}(\| - x + y\|^2 - \| - x - y\|^2) = -(x,y).$$

So $\forall\, n \in \mathbb{Z}$

$$(nx, y) = n(x,y). \tag{3.20}$$

Replacing x by $\frac{x}{n}$ we get

$$(x,y) = n\Big(\frac{x}{n}, y\Big) \implies \Big(\frac{x}{n}, y\Big) = \frac{1}{n}(x,y).$$

Replacing in (3.20) x by $\frac{x}{m}$, $m \in \mathbb{N}$, we get

$$\Big(\frac{n}{m}x, y\Big) = n\Big(\frac{x}{m}, y\Big) = \frac{n}{m}(x,y).$$

So $\forall\, r \in \mathbb{Q}$

$$(rx, y) = r(x,y).$$

$\boxed{\text{Step 4.}}$ Let $\alpha \in \mathbb{R}$ and $r_k \in \mathbb{Q}$ be such that $\lim\limits_{k\to\infty} r_k = \alpha$, then

$$(\alpha x, y) = \Big(\lim_{k\to\infty}(r_k x), y\Big) = \lim_{k\to\infty}(r_k x, y) = \lim_{k\to\infty} r_k(x,y) = \alpha(x,y). \tag{3.21}$$

Equalities (3.20) and (3.21) imply the third property of an inner product.

2. Assume that X is a complex linear space. Taking into account formulas (3.15) and (3.16) of Lemma 3.1.8, let us define the inner product for any $x, y \in X$ by the following formula

$$(x, y) = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2) + \frac{i}{4}(\|x+iy\|^2 - \|x-iy\|^2).$$

Exercise. Complete the proof for a complex linear space X. When proving the linearity with respect to to first variable, apply already proved formulas (3.20) and (3.21).

□

3.2 Orthogonal projection

A complete inner product space is called a *Hilbert space.* A subset M of a linear space is said to be *convex* if for any $x, y \in M$ we have that for any $0 < \alpha < 1$, $\quad \alpha x + (1-\alpha)y \in M$.

Theorem 3.2.1 (Minimizing element). *Let X be a Hilbert space and M be a non-empty closed convex subset of X. Then for any $x \in X$ there exists a unique element $y \in M$ such that*

$$\operatorname{dist}(x, M) = \inf_{z \in M} \|x - z\| = \|x - y\|. \tag{3.22}$$

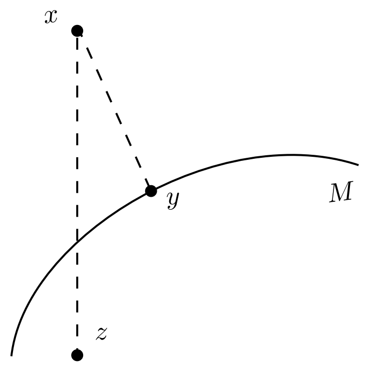


Figure 3.3

Proof. Existence. If $x \in M$, then $\operatorname{dist}(x, M) = 0$ and (3.22) holds for $y = x$. Let $x \in X \setminus M$ and $\delta = \operatorname{dist}(x, M)$. By the definition of an infimum, $\|x - z\| \le \delta$ for any $z \in M$ and for any $k \in \mathbb{N}$ there exists $y_k \in M$ such that

$$\delta - \frac{1}{k} < \|x - y_k\| \le \delta.$$

By passing to the limit as $k \to \infty$ it follows that

$$\lim_{k\to\infty} \|x - y_k\| = \delta.$$

The sequence $\{y_k\}_{k\in\mathbb{N}}$ is called a minimizing sequence.

Let us prove that any minimizing sequence is convergent to an element belonging to M. Let

$$v_k = x - y_k, \qquad \delta_k = \|v_k\| = \|x - y_k\|,$$

then $\lim_{k\to\infty} \delta_k = \delta$. As M is convex, $\frac{y_k+y_m}{2} \in M$ for any $k, m \in \mathbb{N}$. Therefore,

$$\|v_k + v_m\| = \|2x - (y_k + y_m)\| = 2\left\|x - \frac{y_k + y_m}{2}\right\| \geq 2\delta.$$

By the parallelogram identity,

$$\begin{aligned}\|y_k - y_m\|^2 &= \|x - v_k - (x - v_m)\|^2 = \|v_k - v_m\|^2 \\ &= 2\|v_k\|^2 + 2\|v_m\|^2 - \|v_k + v_m\|^2 \\ &\leq 2\delta_k^2 + 2\delta_m^2 - 4\delta,\end{aligned}$$

hence

$$\lim_{k,m\to\infty} \|y_k - y_m\| = 0.$$

By the completeness of X there exists $y \in X$ such that

$$\lim_{k\to\infty} y_k = y.$$

By the closedness of M, $y \in M$, hence

$$\|x - y\| \geq \delta.$$

Also

$$\|x - y\| \leq \|x - y_k\| + \|y_k - y\| = \delta_k + \|y_k - y\|,$$

hence

$$\delta \leq \|x - y\| \leq \lim_{k\to\infty} (\delta_k + \|y_k - y\|) = \delta.$$

So

$$\|x - y\| = \delta.$$

Uniqueness. Assume that there exists another $z \in M$ such that

$$\|x - z\| = \delta.$$

Since, by the convexity of M, $\frac{y_k+y_m}{2} \in M$, by the parallelogram identity

$$\begin{aligned}\|y - z\|^2 &= \|(y - x) - (z - x)\|^2 = 2\|y - x\|^2 + 2\|z - x\|^2 - \|y + z - 2x\|^2 \\ &= 2\delta_k^2 + 2\delta_k^2 - 4\left\|x - \frac{y + z}{2}\right\|^2 \leq 2\delta_k^2 + 2\delta^2 - 4\delta^2 = 0,\end{aligned}$$

hence $\|y - z\| = 0$ and $z = y$. □

Example 3.6. Let $X = \mathbb{R}^2$ be a Hilbert space with the standard inner product, $M = {}^{c}(B(0, 1))$, $x = 0$, then $\operatorname{dist}(0, M) = 1$ and $\|0 - y\| = 1$ for any $y \in S(0, 1)$. So the assumption of the convexity of M in Theorem 3.2.1 is essential.

Example 3.7. Let $X = \mathbb{R}^2$ be a Banach space equipped with the norm $\|(x_1, x_2)\|_\infty = \max\{|x_1|, |x_2|\}$ for any $x = (x_1, x_2) \in \mathbb{R}^2$, which is not an inner product space,

$$M = \bar{B}_\infty(0, 1) = \{(y_1, y_2) \in \mathbb{R}^2 : \max\{|y_1|, |y_2|\} \leq 1 \quad \text{and} \quad x = (2, 0).$$

Then $\operatorname{dist}(x, M) = 1$ and $\|x - y\|_\infty = \|(1, -y_2)\|_\infty = \max\{1, |y_2|\} = 1$. So, in general, in Theorem 3.2.1 Hilbert spaces cannot be replaced by Banach spaces.

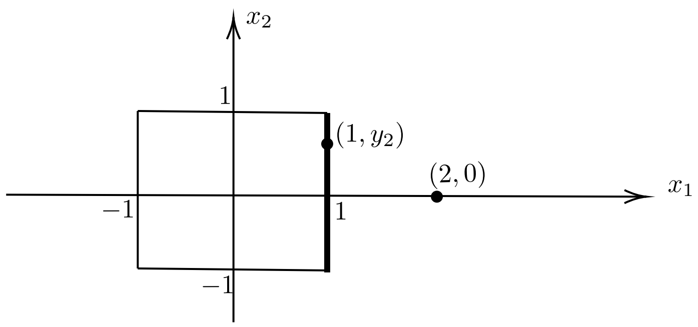


Figure 3.4

Exercise. Prove that the Banach space $\mathbb{R}^2$ equipped with the norm $\|\cdot\|_\infty$ is not a Hilbert space, that is, it is not possible to define an inner product $(\cdot,\cdot)$ on $\mathbb{R}^2$ such that $\|x\|_\infty = \sqrt{(x,x)}$ for any $x \in \mathbb{R}^2$.

Let M be a non-empty subset of an inner product space X, $x \in X$, $y \in M$. If $x - y \perp M$, then y is called an orthogonal projection of x onto M and is denoted by $\operatorname{Pr}_M x$.

Theorem 3.2.2 (Orthogonal projection). *Let X be a Hilbert space, $M \neq \oslash$ be a closed linear subspace of X, $x \in X$, $y \in M$. Then*

$$\operatorname{dist}(x, M) = \|x - y\| \iff x - y \perp M \iff y = \operatorname{Pr}_M x.$$

Proof. $\boxed{\text{Step 1.}}$ Assume that $\operatorname{dist}(x, M) = \|x - y\|$, but $x - y$ is not orthogonal to M. Then there exists $z \in M$ such that

$$(x - y, z) = \beta \neq 0.$$

Clearly, $z \neq 0$. Furthermore, for any $\alpha \in \mathbb{R}$ (or $\mathbb{C}$) by Lemma 3.1.1

$$\|x - y - \alpha z\|^2 = \|x - y\|^2 - 2\operatorname{Re}(x - y, \alpha z) + \|\alpha z\|^2.$$

Since $(x - y, \alpha z) = \overline{\alpha}(x - y, z) = \overline{\alpha}\beta$, we have

$$\|x - y - \alpha z\|^2 = \|x - y\|^2 - 2\operatorname{Re}\overline{\alpha}\beta + |\alpha|^2\|z\|^2.$$

Choose here $\alpha = \frac{\beta}{\|z\|^2}$, then for any $w = y + \alpha z \in M$

$$\|x - w\|^2 = \|x - y - \alpha z\|^2 = \|x - y\|^2 - \frac{|\beta|^2}{\|z\|^2} < \|x - y\|^2.$$

So $\|x - w\| < \|x - y\| = \inf\limits_{u \in M} \|x - u\|$. Contradiction.

$\boxed{\text{Step 2.}}$ Let $x - y \perp M$. Then for any $z \in M$ by the Pythagorean theorem

$$\|x - z\| = \|x - y + y - z\| = \sqrt{\|x - y\|^2 + \|y - z\|^2} \geq \|x - y\|.$$

Hence,

$$\|x - y\| \leq \inf_{z \in M} \|x - z\| = \operatorname{dist}(x, M) \leq \|x - y\|.$$

So $\operatorname{dist}(x, M) = \|x - y\|$. $\square$

3.3 Orthonormal sets

Definition 3.3.1. *A set M in an inner product space $(X, (\cdot,\cdot))$ is said to be orthogonal if M contains more than one element and its elements are pairwise orthogonal, that is, if $x, y \in M$, $x \neq y$, then $(x, y) = 0$. M is called an orthonormal set if it is orthogonal and for all $x \in M$, $\quad \|x\| = 1$.*

Lemma 3.3.1. *An orthonormal set in an inner product space is linearly independent.*

Proof. First, let M be a finite orthogonal set in an inner product space $(X, (\cdot,\cdot))$. Then, for some $n \in \mathbb{N}$, $\quad M = \{e_k\}_{k=1}^n$. Assume that for some scalars $\alpha_1, \ldots, \alpha_n$

$$\sum_{k=1}^{n} \alpha_k e_k = \theta,$$

where θ is the null element of X. Then for any $m \in \{1, \ldots, n\}$

$$0 = (\theta, e_m) = \Big(\sum_{k=1}^{n} \alpha_k e_k, e_m\Big) = \sum_{k=1}^{n} \alpha_k (e_k, e_m) = \alpha_m,$$

hence M is linearly independent.

If M is an infinite set, then, by definition, it is linearly independent if any of its finite subsets is linearly independent. So by the above M is linearly independent. $\square$

Example 3.8. In the space ℓ_2 the system $\{e_k\}_{k=1}^{\infty}$, where

$$e_k = \{0, \ldots, 0, \underbrace{1}_{\substack{k^{\text{th}} \\ \text{position}}}, 0, \ldots\}, \quad k \in \mathbb{N},$$

is an orthonormal set.

Example 3.9. In the space $L_2(-\pi, \pi)$ of real-valued functions the systems

$$A = \left\{\frac{\sin nx}{\sqrt{\pi}}\right\}_{n=1}^{\infty}, \quad B = \left\{\frac{1}{\sqrt{2\pi}}\right\} \cup \left\{\frac{\cos nx}{\sqrt{\pi}}\right\}_{n=1}^{\infty} \quad \text{and} \quad A \cup B$$

are orthonormal sets (or orthonormal systems).

Exercise. Prove this.

Example 3.10. In the space $L_2(-\pi, \pi)$ of all complex-valued functions the system

$$\left\{\frac{e^{inx}}{\sqrt{2\pi}}\right\}_{n \in \mathbb{Z}}$$

is an orthonormal set.

Lemma 3.3.2. *Let X be a finite-dimensional inner product space,* $\dim X = n \in \mathbb{N}$ *and* $\{e_k\}_{k=1}^n$ *be an orthonormal system in X. Then for any $x \in X$*

$$x = \sum_{k=1}^{n} (x, e_k) e_k \tag{3.23}$$

and

$$\|x\|^2 = \sum_{k=1}^{n} |(x, e_k)|^2. \tag{3.24}$$

Proof. By Lemma 3.3.1 the system $\{e_k\}_{k=1}^n$ is linearly independent, hence a basis for X. Therefore for any $x \in X$ there exist scalars $\alpha_1, \dots, \alpha_n$ such that

$$x = \sum_{m=1}^{n} \alpha_m e_m.$$

Equality (3.23) follows because for any $k \in \{1, \dots, n\}$

$$(x, e_k) = \Big(\sum_{m=1}^{n} \alpha_m e_m, e_k \Big) = \sum_{m=1}^{n} \alpha_m (e_m, e_k) = \alpha_k.$$

Also, by (3.23)

$$\begin{aligned} \|x\|^2 = (x,x) &= \Big(\sum_{k=1}^{n} (x, e_k) e_k, \sum_{m=1}^{n} (x, e_m) e_m \Big) = \sum_{k=1}^{n} (x, e_k) \Big(e_k, \sum_{m=1}^{n} (x, e_m) e_m \Big) \\ &= \sum_{k=1}^{n} (x, e_k) \Big(\sum_{m=1}^{n} \overline{(x, e_m)} (e_k, e_m) \Big) = \sum_{k=1}^{n} |(x, e_k)|^2. \end{aligned}$$

□

Lemma 3.3.3. *Let $\{e_k\}_{k=1}^n$, $n \in \mathbb{N}$ be an orthonormal system in an inner product space X and $M = \operatorname{span}\{e_1, \dots, e_n\}$. Then for any $x \in X$*

$$\operatorname{Pr}_M x = \sum_{k=1}^{n} (x, e_k) e_k. \tag{3.25}$$

Proof. The statement follows by the definition of a projection, because $y = \sum\limits_{k=1}^{n} (x, e_k) e_k \in M$ and $x - y \perp M$ since for any $m \in \{1, \dots, n\}$

$$(x - y, e_m) = (x, e_m) - (y, e_m) = (x, e_m) - \sum_{k=1}^{n} (x, e_k)(e_k, e_m) = (x, e_m) - (x, e_m) = 0.$$

□

Let $\{e_k\}_{k=1}^\infty$ be an orthonormal system in an infinite-dimensional inner product space X. Given $x \in X$, the inner products (x, e_k), $k \in \mathbb{N}$ are called the Fourier coefficients of x.

Lemma 3.3.4 (The minimality of the Fourier coefficients). *Let $\{e_k\}_{k=1}^\infty$ be an orthonormal system in an infinite-dimensional inner product space X. Then for any $n \in \mathbb{N}$*

$$\Big\| x - \sum_{k=1}^{n} (x, e_k) e_k \Big\| = \min_{\alpha_1, \dots, \alpha_n \in \mathbb{R} \text{ (or } \mathbb{C})} \Big\| x - \sum_{k=1}^{n} \alpha_k e_k \Big\|. \tag{3.26}$$

Proof. Note that

$$\min_{\alpha_1, \dots, \alpha_n \in \mathbb{R} \text{ (or } \mathbb{C})} \Big\| x - \sum_{k=1}^{n} \alpha_k e_k \Big\| = \operatorname{dist}(x, M),$$

where $M = \operatorname{span}\{e_1, \dots, e_n\}$, and by (3.25)

$$\sum_{k=1}^{n} (x, e_k) e_k = \operatorname{Pr}_M x,$$

hence (3.26) is a direct corollary of Lemma 3.3.3. □

Theorem 3.3.1 (Bessel's equality). *Let $\{e_k\}_{k=1}^{\infty}$ be an orthonormal system in an infinite-dimensional inner product space X. Then for any $x \in X$ and any $n \in \mathbb{N}$*

$$\Big\| x - \sum_{k=1}^{n} (x, e_k) e_k \Big\|^2 = \|x\|^2 - \sum_{k=1}^{n} |(x, e_k)|^2. \tag{3.27}$$

Remark. If $n = 1$ this equality is a particular case of the equality

$$\Big\| x - \frac{(x, y)}{\|y\|^2} y \Big\|^2 = \|x\|^2 - \frac{|(x, y)|^2}{\|y\|^2},$$

where $x, y \in X,\ y \neq 0$, proved in Lemma 3.1.3.

Proof. By Lemma 3.3.3, $\sum\limits_{k=1}^{n} (x, e_k) e_k$ is the projection $\operatorname{Pr}_M x$ of x onto $M = \operatorname{span}\{e_1, \ldots, e_n\}$. Since $x - \operatorname{Pr}_M x \perp \operatorname{Pr}_M x$, by the Pythagorean Theorem

$$\|x\|^2 = \|x - \operatorname{Pr}_M x + \operatorname{Pr}_M x\|^2 = \|x - \operatorname{Pr}_M x\|^2 + \|\operatorname{Pr}_M x\|^2,$$

which by (3.25) and (3.24) implies (3.27). □

Corollary (Bessel's inequality). *Let $\{e_k\}_{k=1}^{\infty}$ be an orthonormal system in an inner product space X. Then for any $x \in X$ the series $\sum\limits_{k=1}^{\infty} |(x, e_k)|^2$ converges and*

$$\sum_{k=1}^{\infty} |(x, e_k)|^2 \leq \|x\|^2. \tag{3.28}$$

Proof. By Bessel's equality (3.27) for any $n \in \mathbb{N}$

$$\sum_{k=1}^{n} |(x, e_k)|^2 \leq \|x\|^2$$

which by passing to the limit as $n \to \infty$ implies inequality (3.28). □

Remark. Clearly inequality (3.28) implies that

$$\lim_{k \to \infty} (x, e_k) = 0. \tag{3.29}$$

Theorem 3.3.2. *Let $\{\varphi_k\}_{k=1}^{\infty}$ be a linearly independent system in an infinite-dimensional inner product space X. Then there exists an orthonormal system $\{e_k\}_{k=1}^{\infty}$ in X such that for any $n \in \mathbb{N}$*

$$E_n = \operatorname{span}\{e_1, \ldots, e_n\} = \operatorname{span}\{\varphi_1, \ldots, \varphi_n\}. \tag{3.30}$$

Proof. $\boxed{\text{Step 1.}}$ Let for $n = 1$

$$e_1 = \frac{\varphi_1}{\|\varphi_1\|}.$$

Clearly

$$E_1 = \operatorname{span}\{e_1\} = \operatorname{span}\{\varphi_1\}.$$

$\boxed{\text{Step 2.}}$ Let for $n = 2$

$$e_1 = \frac{\varphi_1}{\|\varphi_1\|}, \quad e_2 = \frac{\varphi_2 - \operatorname{Pr}_{E_1} \varphi_2}{\|\varphi_2 - \operatorname{Pr}_{E_1} \varphi_2\|} = \frac{\varphi_2 - (\varphi_2, e_1) e_1}{\|\varphi_2 - (\varphi_2, e_1) e_1\|}. \tag{3.31}$$

Note that $\|\varphi_2 - (\varphi_2, e_1)e_1\| > 0$. Indeed, if $\|\varphi_2 - (\varphi_2, e_1)e_1\| = 0 \iff \varphi_2 - (\varphi_2, e_1)e_1 = 0 \iff \varphi_2 - \frac{(\varphi_2,\varphi_1)}{\|\varphi_1\|^2}\varphi_1 = 0$, then φ_1 and φ_2 are linearly independent which contradicts the assumption of Lemma 3.3.1.

Formulas (3.31) imply that for some scalars $\alpha_1, \alpha_2, \beta_1, \beta_2$

$$\varphi_2 = \alpha_1 e_1 + \alpha_2 e_2, \quad e_2 = \beta_1\varphi_1 + \beta_2\varphi_2, \tag{3.32}$$

namely, for

$$\alpha_1 = (\varphi_2, e_1), \quad \alpha_2 = \|\varphi_2 - (\varphi_2, e_1)e_1\|,$$

and

$$\beta_1 = -\frac{(\varphi_2, e_1)}{\|\varphi_2 - (\varphi_2, e_1)e_1\| \cdot \|\varphi_1\|}, \quad \beta_2 = \frac{1}{\|\varphi_2 - (\varphi_2, e_1)e_1\|}.$$

By (3.32), if $x \in \operatorname{span}\{\varphi_1, \varphi_2\}$, then for some scalars γ_1, γ_2

$$\begin{aligned} x = \gamma_1\varphi_1 + \gamma_2\varphi_2 &= \gamma_1\|\varphi_1\|e_1 + \gamma_2(\alpha_1 e_1 + \alpha_2 e_2) \\ &= (\gamma_1\|\varphi_1\| + \gamma_2\alpha_1)e_1 + \gamma_2\alpha_2 e_2, \end{aligned}$$

hence, $x \in \operatorname{span}\{e_1, e_2\}$. Similarly, if $x \in \operatorname{span}\{e_1, e_2\}$, then $x \in \operatorname{span}\{\varphi_1, \varphi_2\}$. So,

$$E_2 = \operatorname{span}\{\varphi_1, \varphi_2\} = \operatorname{span}\{e_1, e_2\}.$$

$\boxed{\text{Step 3.}}$ For any arbitrary $n \in \mathbb{N},\ n \geq 3$, we set

$$e_1 = \frac{\varphi_1}{\|\varphi_1\|}, \quad e_2 = \frac{\varphi_2 - \operatorname{Pr}_{E_1}\varphi_2}{\|\varphi_2 - \operatorname{Pr}_{E_1}\varphi_2\|} = \frac{\varphi_2 - (\varphi_2, e_1)e_1}{\|\varphi_2 - (\varphi_2, e_1)e_1\|}, \dots,$$
$$e_n = \frac{\varphi_n - \operatorname{Pr}_{E_{n-1}}\varphi_n}{\|\varphi_n - \operatorname{Pr}_{E_{n-1}}\varphi_n\|} = \frac{\varphi_n - \sum\limits_{k=1}^{n-1}(\varphi_n, e_k)e_k}{\left\|\varphi_n - \sum\limits_{k=1}^{n-1}(\varphi_n, e_k)e_k\right\|}, \dots.$$

The desired equality (3.30) follows by the argument described in Step 2. $\square$

Remark. The procedure used in this proof is called the *Gram-Schmidt orthogonalization.*

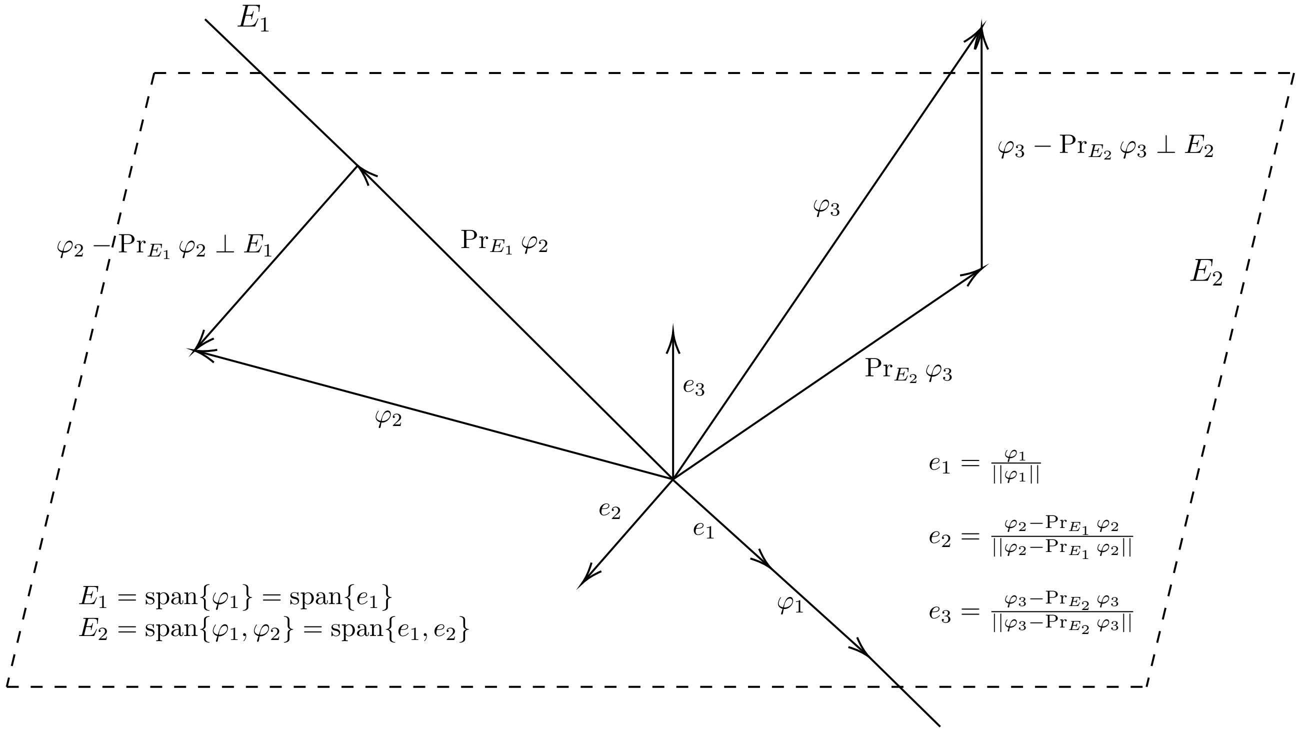


Figure 3.5 Gram-Schmidt orthogonalization
Vectors e_1, e_2, e_3 should have the same length! ($\|e_1\| = \|e_2\| = \|e_3\| = 1$)

Exercise 5. Prove that $\{x^k\}_{k=0}^{\infty}$ is a linearly independent system in $L_2(-1, 1)$.

Exercise 6. By using the Gram-Schmidt orthogonalization construct the first 3 elements of an orthonormal system in $L_2(-1, 1)$ satisfying equality (3.30) considering

1) $\varphi_1(x) = 1,\ \varphi_2(x) = x,\ \varphi_3(x) = x^2$, 2) $\varphi_1(x) = x,\ \varphi_2(x) = x,\ \varphi_3(x) = 1$ and

3) $\varphi_1(x) = x,\ \varphi_2(x) = 1,\ \varphi_3(x) = x^2$.

Theorem 3.3.3. *Let H be an infinite-dimensional Hilbert space and $\{e^k\}_{k=1}^{\infty} \subset H$ be an orthonormal system. Then the following statements are equivalent:*

1) *the set of all finite linear combinations of $e_1, \ldots, e_k, \ldots$ is dense in H, briefly,*

$$\overline{\mathrm{span}\{e^k\}_{k=1}^{\infty}} = H; \tag{3.33}$$

2) *if $x \in H$ and $(x, e_k) = 0$ for all $k \in \mathbb{N}$, then $x = 0$;*

3) *for any $x \in H$*

$$x = \sum_{k=1}^{\infty} (x, e_k) e_k; \tag{3.34}$$

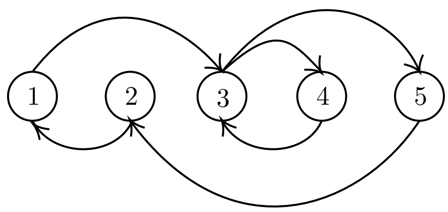
4) *$\{e^k\}_{k=1}^{\infty}$ is a basis for H;*

5) *for any $x \in H$*

$$\|x\|^2 = \sum_{k=1}^{\infty} |(x, e_k)|^2. \tag{3.35}$$

Remark. Property 1) is called the *totality* of $\{e^k\}_{k=1}^{\infty}$. Property 3) is the Fourier expansion of any $x \in H$ along $\{e^k\}_{k=1}^{\infty}$. Equality (3.35) is called the *Parseval equality* (infinite-dimensional Pythagorean theorem). Property 5) is called the closedness of $\{e^k\}_{k=1}^{\infty}$.

Proof. We shall use the following scheme of proof:



$\boxed{1) \implies 3).}$ Let $x \in H$. By 1) for any $\varepsilon > 0$ there exists $n \in \mathbb{N}$ and scalars $\alpha_1, \dots, \alpha_n$ such that

$$\Big\| x - \sum_{k=1}^{n} \alpha_k e_k \Big\| < \varepsilon.$$

Let $m \in \mathbb{N},\ m > n$ and let $\alpha_{n+1} = \ldots = \alpha_m = 0$, then also

$$\Big\| x - \sum_{k=1}^{m} \alpha_k e_k \Big\| < \varepsilon.$$

By the property of minimality of the Fourier coefficients (Lemma 3.3.4),

$$\Big\| x - \sum_{k=1}^{m} (x, e_k) e_k \Big\| \le \Big\| x - \sum_{k=1}^{m} \alpha_k e_k \Big\| < \varepsilon.$$

So $\forall\, \varepsilon > 0\ \exists\, n \in \mathbb{N} : \forall\, m \in \mathbb{N},\ m > n$

$$\Big\| x - \sum_{k=1}^{m} (x, e_k) e_k \Big\| < \varepsilon,$$

which means that $x = \lim\limits_{m\to\infty} \sum\limits_{k=1}^{m} (x, e_k) e_k$ in H or, in other words, $x = \sum\limits_{k=1}^{\infty} (x, e_k) e_k$.

$\boxed{3) \implies 5).}$ Let $x \in H$. By 3), the continuity of a norm and formula (3.24)

$$\begin{aligned} \|x\|^2 &= \Big\| \sum_{k=1}^{\infty} (x, e_k) e_k \Big\|^2 = \lim_{m\to\infty} \Big\| \sum_{k=1}^{m} (x, e_k) e_k \Big\|^2 \\ &= \lim_{m\to\infty} \sum_{k=1}^{m} |(x, e_k)|^2 = \sum_{k=1}^{\infty} |(x, e_k)|^2. \end{aligned}$$

$\boxed{5) \implies 2).}$ Let $x \in H$ be such that $(x, e_k) = 0$ for all $k \in \mathbb{N}$, then by 5)

$$\|x\|^2 = \sum_{k=1}^{\infty} |(x, e_k)|^2 = 0 \implies x = 0.$$

$\boxed{2) \implies 1).}$ Let $M = \overline{\operatorname{span}\{e_k\}_{k=1}^{\infty}}$. Assume to the contrary that $M \neq H$. Then there exists $x \in H$ such that $x \notin M$. Since M is a closed linear subspace of H by Theorem 3.3.2 there exists $y \in M$ such that $\operatorname{dist}(x, M) = \|x - y\| > 0$ and $z = x - y \perp M$. Then $z \neq 0$ and $(z, e_k) = 0$ for all $k \in \mathbb{N}$, which contradicts 2).

$\boxed{3) \implies 4).}$ Let $x \in H$. By 3)

$$x = \sum_{k=1}^{\infty} (x, e_k) e_k.$$

Moreover, this expansion of x along $\{e_k\}_{k=1}^{\infty}$ is unique. Indeed, assume that for some scalars $\alpha_k,\ k \in \mathbb{N}$

$$x = \sum_{k=1}^{\infty} \alpha_k e_k, \tag{3.36}$$

then for any $m \in \mathbb{N}$ by the continuity of an inner product

$$(x, e_m) = \Big(\sum_{k=1}^{\infty} \alpha_k e_k, e_m \Big) = \sum_{k=1}^{\infty} \alpha_k (e_k, e_m) = \alpha_m.$$

$\boxed{4) \implies 3).}$ Assume that $\{e_k\}_{k=1}^{\infty}$ is a basis for H and $x \in H$. Then there exist uniquely defined scalars $\alpha_k,\ k \in \mathbb{N}$ such that equality (3.36) holds. By the above argument $\alpha_k = (x, e_k)$ for all $k \in \mathbb{N}$, which implies equality (3.34). $\square$

Theorem 3.3.4. *Any separable Hilbert space has a basis.*

Proof. Let H be an infinite-dimensional Hilbert space and $M = \{\varphi_k\}_{k=1}^{\infty}$ be a countable dense subset of H. By applying the Gram-Schmidt orthogonalization we get an orthonormal sequence $\{e_k\}_{k=1}^{\infty} \subset H$ such that for any $n \in \mathbb{N}$

$$\operatorname{span}\{e_1, \ldots, e_n\} = \operatorname{span}\{\varphi_1, \ldots, \varphi_n\}.$$

In particular each φ_k is a finite linear combination of $e_1, \ldots, e_m, \ldots$: there exist $n_k \in \mathbb{N}$ and scalars $\alpha_{k1}, \ldots, \alpha_{kn_k}$ such that

$$\varphi_k = \sum_{m=1}^{n_k} \alpha_{km} e_m.$$

Since M is dense in H for any $x \in H$ and for any $\varepsilon > 0$, there exists φ_k such that

$$\|x - \varphi_k\| < \varepsilon \implies \Big\| x - \sum_{m=1}^{n_k} \alpha_{km} e_m \Big\| < \varepsilon,$$

which means that the system $\{e_k\}_{k=1}^{\infty}$ is *total.* Therefore, by Theorem 3.3.3 the system $\{e_k\}_{k=1}^{\infty}$ is a basis for H. $\square$

Remark. By a basis (or a Schauder basis, or a countable basis) in an infinite-dimensional Banach space B (in a Hilbert space, in particular) over a field F ($\mathbb{R}$ or $\mathbb{C}$, in particular), we mean a countable system $\{e_k\}_{k=1}^{\infty}$ such that for any $x \in B$ there exist uniquely defined scalars $\alpha_k \in F,\ k \in \mathbb{N}$, for which

$$x = \sum_{k=1}^{\infty} \alpha_k e_k \quad \text{in} \quad B \iff \lim_{m\to\infty} \Big\| x - \sum_{k=1}^{m} \alpha_k e_k \Big\| = 0.$$

This notion essentially differs from the notion of a Hamel basis, considered in Section 2.5, which is not necessarily countable and where sums are only finite.

Remark. The basis problem is the question asked by Banach, whether *every* separable Banach space has a basis. This was an open problem for many years. In 1973 this question was negatively answered by Danish mathematician Per Enflo, who constructed a separable Banach space without a basis.[2]

3.4 Representation of linear functionals

Lemma 3.4.1. *Let H be a Hilbert space, $z \in H$ and the functional $f : H \to \mathbb{C}$ be defined by*

$$f(x) = (x, z)_H \quad \forall\, x \in H.$$

Then f is a continuous linear functional (= bounded linear functional) and

$$\|f\|_{H\to\mathbb{C}} = \|z\|_H. \tag{3.37}$$

Proof. By the Cauchy-Bunyakowski inequality

$$\|f\|_{H\to\mathbb{C}} = \sup_{\substack{x\in H\\ x\neq 0}} \frac{|f(x)|}{\|x\|_H} = \sup_{\substack{x\in H\\ x\neq 0}} \frac{|(x,z)_H|}{\|x\|_H} \le \sup_{\substack{x\in H\\ x\neq 0}} \frac{\|x\|_H \|z\|_H}{\|x\|_H} = \|z\|_H.$$

On the other hand,

$$\|f\|_{H\to\mathbb{C}} \ge \frac{|f(x)|}{\|x\|_H}\bigg|_{x=z} = \frac{(z,z)_H}{\|z\|_H} = \frac{\|z\|_H^2}{\|z\|_H} = \|z\|_H.$$

Hence, equality (3.37) follows. □

Lemma 3.4.2. *Let H be a Hilbert space, $z, w \in H$ and*

$$(x, z)_H = (x, w)_H \quad \forall\, x \in H. \tag{3.38}$$

Then $z = w$.

Proof. By (3.38)

$$(x, z - w)_H = 0 \quad \forall\, x \in H.$$

Take here $x = z - w$. Then

$$(z - w, z - w)_H = 0 \implies z - w = 0 \implies z = w.$$

□

[2] A proof of this can be found here.

Lemma 3.4.3. *Let H be a Hilbert space, $M \subset H$, $M \neq H$, be a closed linear subspace. Then*

$$H = M \oplus M^{\perp},$$

that is, $\forall\, x \in H$ there exist uniquely defined $y \in M$ and $z \in M^{\perp}$ such that

$$x = y + z.$$

Proof. If $x \in M$, then $y = x,\ z = 0$. Let $x \in H \setminus M$. It was proved earlier that there exists $y \in M$ such that $\operatorname{dist}(x, M) = \|x - y\|_H$ and $z = x - y \perp M$.

So

$$x = y + x - y = y + z,$$

where $y \in M,\ z \in M^{\perp}$.

Assume also that there exist $\tilde{y} \in M,\ \tilde{z} \in M^{\perp}$ such that

$$x = \tilde{y} + \tilde{z}.$$

Then

$$\tilde{y} + \tilde{z} = y + z \iff \underbrace{\tilde{y} - y}_{\in M} = \underbrace{z - \tilde{z}}_{\in M^{\perp}},$$

hence $\tilde{y} - y \perp \tilde{y} - y \iff (\tilde{y} - y, \tilde{y} - y)_H = 0 \iff \tilde{y} = y$.

Similarly, $z - \tilde{z} \perp z - \tilde{z} \iff \tilde{z} = z$. □

Theorem 3.4.1 (Riesz representation theorem). *Let H be a Hilbert space and $f : H \to \mathbb{C}$ be a continuous linear (= bounded linear) functional. Then there exists a uniquely defined element $z \in H$ such that*

$$f(x) = (x, z)_H \quad \forall\, x \in H. \tag{3.39}$$

Proof. $\boxed{\text{Step 1.}}$ (*the case of a finite-dimensional Hilbert space.*) Let $\dim H = n$, $\{e_k\}_{k=1}^{n}$ be an orthonormal basis in H and $x = \sum\limits_{k=1}^{n} \alpha_k e_k$ for some $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$. Then

$$f(x) = f\Big(\sum_{k=1}^{n} \alpha_k e_k\Big) = \sum_{k=1}^{n} \alpha_k f(e_k)$$

and

$$\Big(x, \sum_{k=1}^{n} \overline{f(e_k)} e_k\Big) = \sum_{k=1}^{n} (x, \overline{f(e_k)} e_k) = \sum_{k=1}^{n} f(e_k)(x, e_k) = \sum_{k=1}^{n} \alpha_k f(e_k).$$

Hence,

$$f(x) = (x, z)_H, \qquad \text{where } z = \sum_{k=1}^{n} \overline{f(e_k)} e_k.$$

Next, consider the null space $\mathcal{N}$ of f, that is, the set of all elements $y = \sum\limits_{k=1}^{n} \beta_k e_k$ such that

$$f(y) = f\Big(\sum_{k=1}^{n} \beta_k e_k\Big) = \sum_{k=1}^{n} \beta_k f(e_k) = 0.$$

Note that $z \in \mathcal{N}_f^{\perp}$, because for any $y \in \mathcal{N}_f$

$$(z, y) = \Big(\sum_{k=1}^{n} \overline{f(e_k)} e_k, y\Big) = \sum_{k=1}^{n} \overline{f(e_k)}(e_k, y) = \sum_{k=1}^{n} \overline{f(e_k)}\,\overline{\beta_k} = \overline{\sum_{k=1}^{n} f(e_k) \beta_k}.$$

This gives a hint on how to prove the theorem for general Hilbert spaces.

$\boxed{\text{Step 2.}}$ (*Existence of z satisfying* (3.39).) If $f = 0$, then $z = 0$. Assume that $f \neq 0$, then $z \neq 0$ (otherwise $f(x) = (x, 0)_H = 0$ for all $x \in H$). Consider the null set $\mathcal{N}(f)$ of f, that is, the set of all $x \in H$ for which

$$f(x) = 0.$$

$\mathcal{N}(f)$ is a closed linear subspace of H. Indeed, if $x, y \in \mathcal{N}(f)$ and $\alpha, \beta \in \mathbb{C}$, then by the linearity of f

$$f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) = 0,$$

hence $\alpha x + \beta y \in \mathcal{N}(f)$.

Also, if $x_k \in \mathcal{N}(f),\ k \in \mathbb{N},\ x \in H$ and $x_k \to x$ in H, then by the continuity of f

$$f(x) = f\Big(\lim_{k\to\infty} x_k\Big) = \lim_{k\to\infty} f(x_k) = 0,$$

hence $x \in \mathcal{N}(f)$.

By Lemma 3.4.3

$$H = \mathcal{N}(f) \oplus \mathcal{N}(f)^{\perp}.$$

Since $\mathcal{N}(f) \neq H$ (otherwise $f(x) = 0$ for all $x \in H$, hence $f = 0$), it follows that $\mathcal{N}(f)^{\perp} \neq \{0\}$. Choose any $z_0 \in \mathcal{N}(f)^{\perp},\ z_0 \neq 0$.

Let x be an arbitrary element in H. Choose $\alpha \in \mathbb{C}$ so that

$$\alpha z_0 + x \in \mathcal{N}(f) \iff f(z_0 + x) = 0 \iff \alpha f(z_0) + f(x) = 0,$$

hence

$$\alpha = -\frac{f(x)}{f(z_0)}.$$

Since $z_0 \perp \alpha z_0 + x$ we get

$$\begin{aligned} -\frac{f(x)}{f(z_0)} z_0 + x \perp z_0 &\iff \Big(-\frac{f(x)}{f(z_0)} z_0 + x, z_0\Big)_H = 0 \\ &\iff -\frac{f(x)}{f(z_0)} (z_0, z_0)_H + (x, z_0)_H = 0 \\ &\iff f(x) = \frac{f(z_0)}{(z_0, z_0)_H} (x, z_0)_H \\ &\iff f(x) = \Big(x, \frac{\overline{f(z_0)}}{(z_0, z_0)_H} z_0\Big)_H. \end{aligned}$$

So for all $x \in H$

$$f(x) = (x, z)_H,$$

where

$$z = \frac{\overline{f(z_0)}}{(z_0, z_0)_H} z_0. \tag{3.40}$$

$\boxed{\text{Step 3.}}$ (*Uniqueness of z satisfying* (3.39).) Assume also for some $w \in H$ for all $x \in H$ that

$$f(x) = (x, w)_H.$$

Then

$$(x, z)_H = (x, w)_H$$

for all $x \in H$, hence by Lemma 3.4.1, $w = z$. □

Remark 3.4.1. Let $z_1 \in \mathcal{N}(f),\ z_1 \neq 0$ and $z_1 \neq z_0$. Then by formula (3.40) with z_1 replacing z_0 and by the uniqueness of z it follows that

$$\frac{\overline{f(z_1)}}{(z_1, z_1)_H} z_1 = \frac{\overline{f(z_0)}}{(z_0, z_0)_H} z_0 \iff z_1 = \alpha z_0,$$

where

$$\alpha = \frac{\overline{f(z_0)}}{\overline{f(z_1)}} \frac{(z_1, z_1)_H}{(z_0, z_0)_H}.$$

Also, for any $\alpha \in \mathbb{C},\ \alpha z_0 \in \mathcal{N}(f)^\perp$ because for any $x \in \mathcal{N}(f),\ (\alpha z_0, x) = \alpha(z_0, x) = 0$.

Therefore

$$\mathcal{N}(f)^\perp = \{\alpha z_0,\ \alpha \in \mathbb{C}\}.$$

This means that the subspace $\mathcal{N}(f)^\perp$ is one-dimensional. So for any non-zero continuous linear functional f on H the co-dimension of the null set $\mathcal{N}(f)$ (= the dimension of its orthogonal complement $\mathcal{N}(f)^\perp$) is equal to 1.

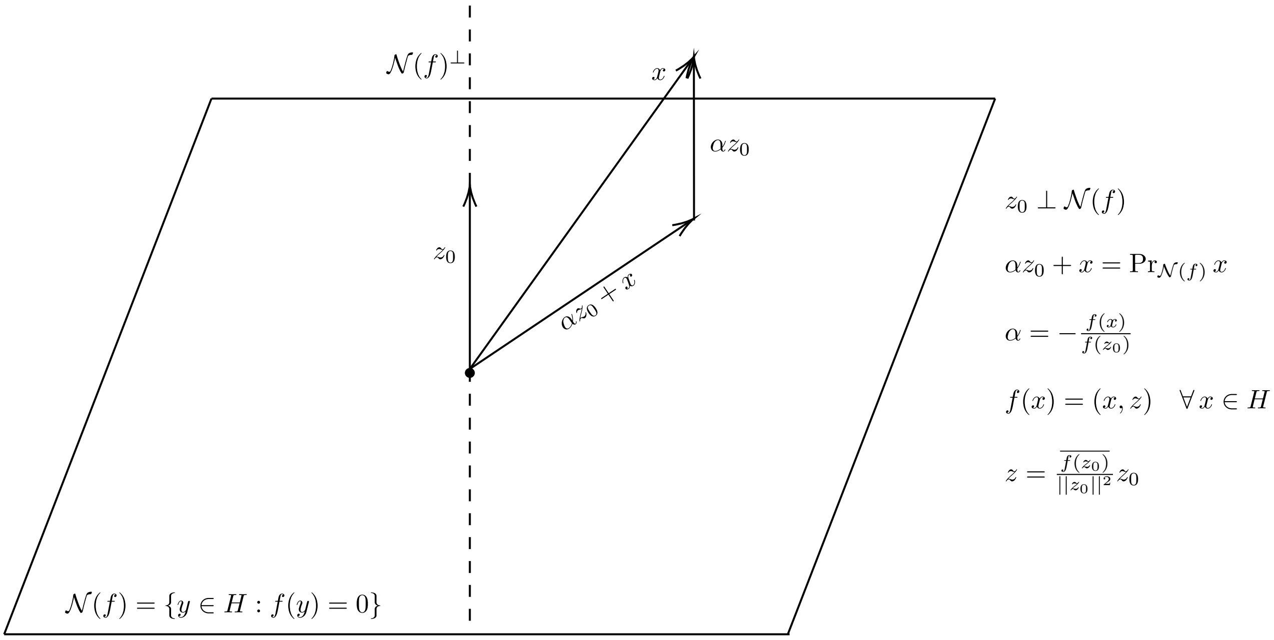


Figure 3.6 Geometric illustration to the proof of Theorem 3.4.1

Exercise. By applying Step 1 of the proof of Theorem 3.4.1, prove that for any separable infinite-dimensional Hilbert space H

$$f(x) = (x, z) \quad \forall\, x \in H,$$

where for any orthonormal basis $\{e_k\}_{k=1}^\infty$ of H

$$z = \sum_{k=1}^\infty \overline{f(e_k)} e_k.$$

3.5 Adjoint operators

Lemma 3.5.1. *Let H_1, H_2 be Hilbert spaces and $T : H_1 \to H_2$. Then*

$$\|T\|_{H_1\to H_2} = \sup_{\substack{x\in H_1,\ x\neq 0\\ y\in H_2,\ y\neq 0}} \frac{|(Tx,y)_{H_2}|}{\|x\|_{H_1}\|y\|_{H_2}}.$$

Proof. Denote the right-hand side of this equality by $\|T\|^+_{H_1\to H_2}$. Then, by the Cauchy-Bunyakowski inequality

$$\|T\|^+_{H_1\to H_2} \le \sup_{\substack{x\in H_1,\ x\neq 0\\ y\in H_2,\ y\neq 0}} \frac{\|Tx\|_{H_2}\|y\|_{H_2}}{\|x\|_{H_1}\|y\|_{H_2}} = \sup_{\substack{x\in H_1\\ x\neq 0}} \frac{\|Tx\|_{H_2}}{\|x\|_{H_1}} = \|T\|_{H_1\to H_2}.$$

On the other hand,

$$\|T\|^+_{H_1\to H_2} \ge \sup_{\substack{x\in H_1\\ x\neq 0}} \frac{|(Tx,y)_{H_2}|}{\|x\|_{H_1}\|y\|_{H_2}}\bigg|_{y=Tx} \ge \sup_{\substack{x\in H_1\\ x\neq 0}} \frac{|(Tx,Tx)_{H_2}|}{\|x\|_{H_1}\|Tx\|_{H_2}} = \sup_{\substack{x\in H_1\\ x\neq 0}} \frac{\|Tx\|_{H_2}}{\|x\|_{H_1}} = \|T\|_{H_1\to H_2}.$$

□

Definition 3.5.1. *Let H_1, H_2 be Hilbert spaces and $T : H_1 \to H_2$ be a bounded linear operator. The adjoint of T is the operator $T^* : H_2 \to H_1$ such that*

$$(Tx, y)_{H_2} = (x, T^*y)_{H_1} \quad \forall\, x \in H_1,\ y \in H_2. \tag{3.41}$$

Theorem 3.5.1. *The operator T^* in this definition exists, is uniquely defined, is a bounded linear operator and*

$$\|T^*\|_{H_2\to H_1} = \|T\|_{H_1\to H_2}. \tag{3.42}$$

Proof. $\boxed{\text{Step 1.}}$ *Existence of T^*.* For a fixed $y \in H_2$ consider the linear functional on H_1 defined by

$$f(x) = (Tx, y)_{H_2} \quad \forall\, x \in H_1.$$

Then, by Lemma 3.5.1 and the Cauchy-Bunyakowski inequality,

$$\|f\|_{H_1\to\mathbb{C}} = \sup_{\substack{x\in H_1\\ x\neq 0}} \frac{|(Tx,y)_{H_2}|}{\|x\|_{H_1}} \le \sup_{\substack{x\in H_1\\ x\neq 0}} \frac{\|Tx\|_{H_2}\|y\|_{H_2}}{\|x\|_{H_1}} = \|T\|_{H_1\to H_2}\|y\|_{H_2} < \infty,$$

so this functional is bounded.

By the Riesz representation theorem there exists a uniquely defined element $z \in H_1$ such that

$$f(x) = (Tx, y)_{H_2} = (x, z)_{H_1}.$$

Let

$$T^*y = z \quad \forall\, y \in H_2.$$

Then,

$$(Tx, y)_{H_2} = (x, T^*y)_{H_1} \quad \forall\, x \in H_1,\ y \in H_2.$$

$\boxed{\text{Step 2.}}$ *Linearity of T^*.* Given $y_1, y_2 \in H_2$ and $\alpha_1, \alpha_2 \in \mathbb{C}$, we have that for all $x \in H_1$

$$(Tx, y_1)_{H_2} = (x, T^*y_1)_{H_1}, \quad (Tx, y_2)_{H_2} = (x, T^*y_2)_{H_1}$$

and
$$(Tx, \alpha_1 y_1 + \alpha_2 y_2)_{H_2} = (x, T^*(\alpha_1 y_1 + \alpha_2 y_2))_{H_1}.$$

Then $\forall\, x \in H_1$

$$\begin{aligned}(x, T^*(\alpha_1 y_1 + \alpha_2 y_2))_{H_1} &= (Tx, \alpha_1 y_1 + \alpha_2 y_2)_{H_2} = \overline{\alpha}_1 (Tx, y_1)_{H_2} + \overline{\alpha}_2 (Tx, y_2)_{H_2} \\ &= \overline{\alpha}_1 (x, T^* y_1)_{H_1} + \overline{\alpha}_2 (x, T^* y_2)_{H_1} = (x, \alpha_1 T^* y_1 + \alpha_2 T^* y_2)_{H_1}.\end{aligned}$$

Hence, by Lemma 4.3.2
$$T^*(\alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 T^* y_1 + \alpha_2 T^* y_2.$$

$\boxed{\text{Step 3.}}$ *Uniqueness of T^*.* Indeed, assume that $\widetilde{T^*}$ is also such that

$$(Tx, y)_{H_2} = (x, \widetilde{T^*} y)_{H_1} \quad \forall\, x \in H_1,\ y \in H_2.$$

Then
$$(x, T^* y)_{H_2} = (x, \widetilde{T^*} y)_{H_1} \quad \forall\, x \in H_1,\ y \in H_2.$$

Hence, by Lemma 4.3.2
$$\widetilde{T^*} y = T^* y \quad \forall\, y \in H_2.$$

$\boxed{\text{Step 4.}}$ *Boundedness of T^* and formula* (3.42). By Lemma 3.5.1

$$\begin{aligned}\|T^*\|_{H_2 \to H_1} &= \sup_{\substack{x \in H_2,\ x \neq 0 \\ y \in H_1,\ y \neq 0}} \frac{|(T^* x, y)_{H_1}|}{\|x\|_{H_2} \|y\|_{H_1}} = \sup_{\substack{x \in H_2,\ x \neq 0 \\ y \in H_1,\ y \neq 0}} \frac{|(y, T^* x)_{H_1}|}{\|x\|_{H_2} \|y\|_{H_1}} \\ &= \sup_{\substack{x \in H_2,\ x \neq 0 \\ y \in H_1,\ y \neq 0}} \frac{|(Ty, x)_{H_2}|}{\|y\|_{H_1} \|x\|_{H_2}} = \|T\|_{H_1 \to H_2} < \infty.\end{aligned}$$

$\square$

Example. Let $X = \mathbb{C}^n$, $T : \mathbb{C}^n \to \mathbb{C}^n$ be a linear operator and $e_1, \ldots, e_n$ be a basis for $\mathbb{C}^n$. Then, for
$$x = \sum_{k=1}^n x_k e_k, \quad y = \sum_{k=1}^n y_k e_k \in \mathbb{C}^n$$

we have
$$(x, y) = \sum_{k=1}^n x_k \overline{y_k} = [x]^t \overline{[y]},$$

where
$$[x] = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},\ [y] = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},\ [x]^t = (x_1, \ldots, x_n),\ \overline{[y]} = \begin{pmatrix} \overline{y_1} \\ \vdots \\ \overline{y_n} \end{pmatrix}.$$

Let
$$Te_m = \sum_{k=1}^n a_{km} e_k, \quad a_{km} = (Te_m)_k, \quad m = 1, \ldots, n$$

and

$$[T] = (a_{km})_{k,m=1}^n = \begin{pmatrix} (Te_1)_1 & \cdots & (Te_m)_1 & \cdots & (Te_n)_1 \\ \cdot & \cdots & \cdot & \cdots & \cdot \\ (Te_1)_k & \cdots & (Te_m)_k & \cdots & (Te_n)_k \\ \cdot & \cdots & \cdot & \cdots & \cdot \\ (Te_1)_n & \cdots & (Te_m)_n & \cdots & (Te_n)_n \end{pmatrix} = ([Te_1], \ldots, [Te_n])$$

be the matrix of the operator T in the basis $e_1, \dots, e_n$. Its columns consist of the co-ordinates of the images of the vectors of the basis.

Then

$$\begin{aligned} y = Tx &\iff \sum_{k=1}^n y_k e_k = T\Big(\sum_{m=1}^n x_m e_m\Big) = \sum_{m=1}^n x_m T e_m \\ &= \sum_{m=1}^n x_m \Big(\sum_{k=1}^n a_{km} e_k\Big) = \sum_{k=1}^n \Big(\sum_{m=1}^n a_{km} x_m\Big) e_k \\ &\iff y_k = \sum_{m=1}^n a_{km} x_m, \quad k = 1, \dots, n \\ &\iff [y] = [T][x] \iff [Tx] = [T][x]. \end{aligned}$$

Let $T^* : \mathbb{C}^n \to \mathbb{C}^n$ be the adjoint of T. Then

$$[T^*x] = [T^*][x],$$

where $[T^*]$ is the matrix of T^* in the basis $e_1, \dots, e_n$.

By definition,

$$(Tx, y) = (x, T^*y) \quad \forall\, x, y \in \mathbb{C}^n$$

or

$$\begin{aligned} [Tx]^t\overline{[y]} = [x]^t\overline{[T^*y]} &\iff ([T][x])^t\overline{[y]} = [x]^t\overline{[T^*][y]} \\ &\iff [x]^t[T]^t\overline{[y]} = [x]^t\overline{[T^*][y]} \quad \forall\, x, y \in \mathbb{C}^n. \end{aligned}$$

Hence

$$[T]^t = \overline{[T^*]} \iff [T^*] = \overline{[T^t]} = [T]^*.$$

So, the matrix of the adjoint operator T^* is the adjoint of the matrix of the operator T.

Lemma 3.5.2 (Zero operator). *Let H be a Hilbert space and $T : H \to H$ be a linear operator. Then*

1) $T = 0 \iff (Tx, y)_H = 0 \quad \forall\, x, y \in H,$

2) *for a complex Hilbert space H,* $\quad T = 0 \iff (Tx, x)_H = 0 \quad \forall\, x \in H.$

Proof. $\boxed{\text{Necessity.}}$ This is obvious for both 1) and 2).

$\boxed{\text{Sufficiency.}}$ For 1). Take $y = Tx$, then $(Tx, Tx) = 0$, hence $Tx = 0\ \forall\, x \in H$ or $T = 0$.

For 2). For any $x, y \in H$,

$$\begin{aligned} 0 = (T(x+y), x+y)_H &= (Tx + Ty, x + y)_H = (Tx, x)_H + (Ty, x)_H + (Tx, y)_H + (Ty, y)_H \\ &= (Ty, x)_H + (Tx, y)_H. \end{aligned}$$

So

$$(Ty, x)_H + (Tx, y)_H = 0 \quad \forall\, x, y \in H,$$

and hence, by replacing x with ix:

$$(Ty, ix)_H + (T(ix), y)_H = -i(Ty, x)_H + i(Tx, y) = 0,$$

or

$$-(Ty, x)_H + (Tx, y) = 0.$$

Hence $(Tx, y) = 0 \quad \forall\, x, y \in H$ and by 1) $T = 0$. □

Remark 3.5.1. For a real Hilbert space 2), in general, is not valid. For example, let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the rotation of the plane through a right angle. Then T is linear, and $Tx \perp x$, hence $(Tx, x)_{\mathbb{R}^2} = 0$ for all $x \in \mathbb{R}^2$, but $T \neq 0$.

Theorem 3.5.2. *Let H_1, H_2 be Hilbert spaces. If $S : H_1 \to H_2$, $T : H_1 \to H_2$ are bounded linear operators and $\alpha \in \mathbb{R}$ (or $\mathbb{C}$), then*

1) $(S + T)^* = S^* + T^*$,

2) $(\alpha T)^* = \overline{\alpha} T^*$,

3) $(T^*)^* = T$,

4) $\|T^*T\|_{H_1\to H_1} = \|TT^*\|_{H_2\to H_2} = \|T\|^2_{H_1\to H_2}$.
If $S : H_1 \to H_2$, $T : H_2 \to H_1$ are bounded linear operators, then

5) $(ST)^* = T^*S^*$.

Proof.

Exercise. Prove Statements 1) – 4).

Proof of Statement 5). Note that $ST : H_2 \to H_2$, hence $(ST)^* : H_2 \to H_2$. On one hand,

$$(STx, y)_{H_2} = (x, (ST)^*y)_{H_2} \quad \forall\, x, y \in H_2.$$

On the other hand,

$$(STx, y)_{H_2} = (Tx, S^*y)_{H_1} = (x, T^*S^*y)_{H_2} \quad \forall\, x, y \in H_2.$$

So

$$(x, (ST)^*y)_{H_2} = (x, T^*S^*y)_{H_2} \quad \forall\, x, y \in H_2$$

and by Lemma 4.3.2,

$$(ST)^*y = T^*S^*y \quad \forall\, y \in H_2 \iff (ST)^* = T^*S^*.$$

□

3.6 Concluding remarks

Chapters 1–3 were devoted to the study of basic properties of the main types of spaces used in functional analysis, namely topological spaces, metric spaces, normed spaces, inner product spaces, and relationships between them.

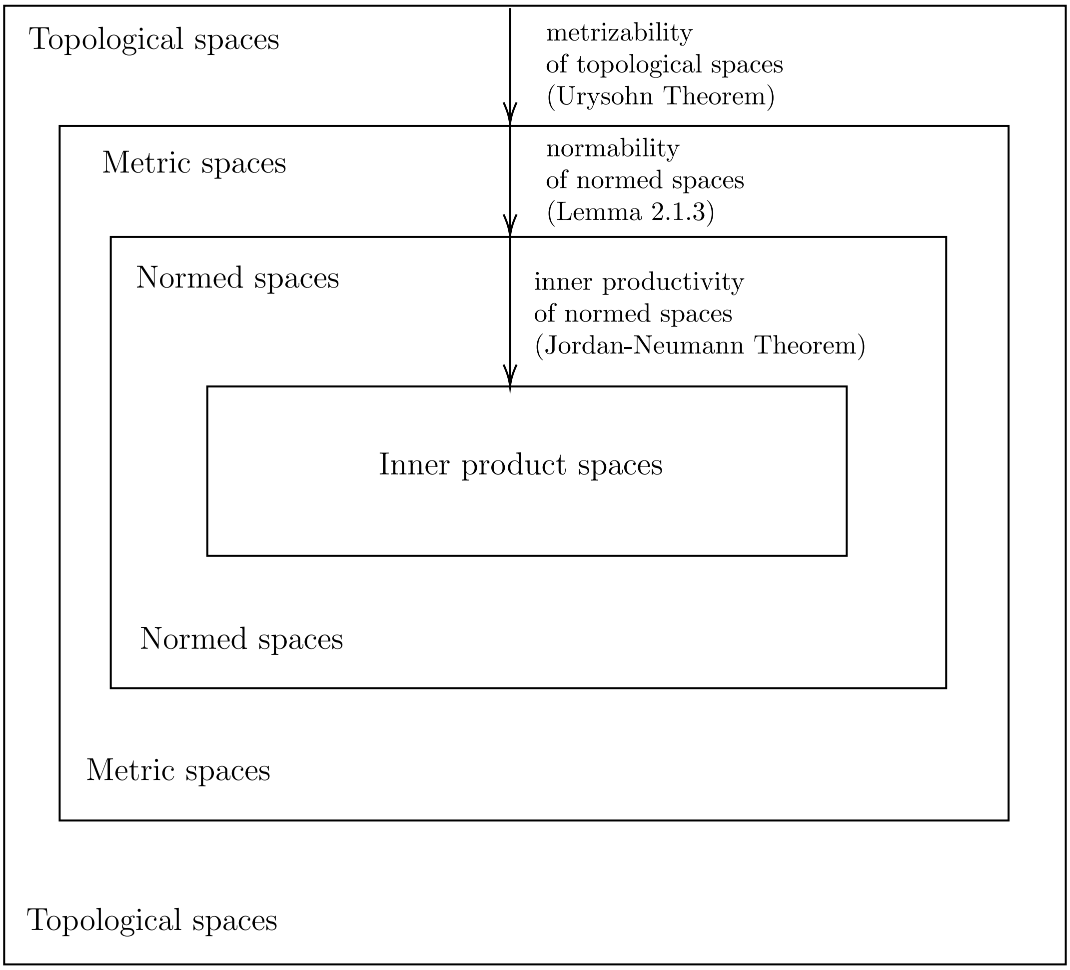


Figure 3.7 Spaces of functional analysis

4. Linear functionals in Banach spaces

4.1 Extension of linear operators by continuity

Let X and Y be normed spaces, Z_1, Z_2 be subsets of X, $Z_1 \subset Z_2$, $Z_1 \neq Z_2$ and $f_1 : Z_1 \to Y$. An operator $f_2 : Z_2 \to Y$ is said to be an extension of f_1 from Z_1 to Z_2 if

$$f_2(x) = f_1(x) \quad \forall\, x \in Z_1.$$

Theorem 4.1.1 (Extension by continuity). *Let X be a normed space and Y be Banach spaces, Z be an unclosed linear subspace of X, and $f : Z \to Y$ be a continuous linear operator. Then there exists a uniquely defined continuous linear operator $\bar{f} : \bar{Z} \to Y$ which is an extension of f from Z to $\bar{Z}$. Moreover,*

$$\|\bar{f}\|_{\bar{Z}\to Y} = \|f\|_{Z\to Y}. \tag{4.1}$$

Proof. $\boxed{\text{Step 1.}}$ (*Definition of $\bar{f}$.*) Let for $x \in \bar{Z}$

$$\bar{f}(x) = \begin{cases} f(x) & \text{if } x \in Z \\ \lim\limits_{k\to\infty} f(x_k) & \text{if } x \in \bar{Z} \setminus Z, \end{cases} \tag{4.2}$$

where $x_k \in Z$, $k \in \mathbb{N}$ and $\lim\limits_{k\to\infty} x_k = x$ in X.

$\boxed{\text{Step 2.}}$ (*Correctness of the definition of $\bar{f}$.*)

$\boxed{\text{Step 2a.}}$ (*Existence of the limit $\lim\limits_{k\to\infty} f(x_k)$ in* (4.2) *for any sequence $\{x_k\}_{k\in\mathbb{N}} \subset Z$ such that $\lim\limits_{k\to\infty} x_k = x$.*) Since the operator f is linear and bounded, $\{x_k\}_{k\in\mathbb{N}} \subset Z$, for any $k, m \in \mathbb{N}$

$$\|f(x_k) - f(x_m)\|_Y = \|f(x_k - x_m)\|_Y \leq \|f\|_{X\to Y}\|x_k - x_m\|_X.$$

Since the limit $\lim\limits_{k\to\infty} x_k$ exists, $\{x_k\}_{k\in\mathbb{N}}$ is a Cauchy sequence in X: $\lim\limits_{k,m\to\infty} \|x_k - x_m\|_X = 0$. Hence, $\lim\limits_{k\to\infty} \|f(x_k) - f(x_m)\|_Y = 0$. So, the sequence $\{f(x_k)\}_{k\in\mathbb{N}}$ is a Cauchy sequence in Y. Since Y is a complete space, the limit $\lim\limits_{k\to\infty} f(x_k)$ exists in Y.

$\boxed{\text{Step 2b.}}$ (*Coincidence of the values of the limits in* (4.2) *for different sequences.*) Let $\{x_k\}_{k\in\mathbb{N}}$, $\{y_k\}_{k\in\mathbb{N}} \subset Z$, $a, b \in Y$, $\lim\limits_{k\to\infty} x_k = \lim\limits_{k\to\infty} y_k = x$ in X and $\lim\limits_{k\to\infty} f(x_k) = a$, $\lim\limits_{k\to\infty} f(y_k) = b$ in Y. We need to prove that $a = b$. Consider the sequence $\{z_k\}_{k\in\mathbb{N}} \subset Z$ defined by

$$z_{2m-1} = x_m,\ z_{2m} = y_m, \quad m \in \mathbb{N},$$

that is, $z = \{x_1, y_1, \ldots, x_m, y_m, \ldots\}$. Then $\lim\limits_{k\to\infty} z_k = x$ in X, hence, by Step 2a, there exists $c \in Y$ such that $\lim\limits_{k\to\infty} f(z_k) = c$ in Y. Therefore,

$$c = \lim_{m\to\infty} f(z_{2m-1}) = \lim_{m\to\infty} f(x_m) = a, \quad c = \lim_{m\to\infty} f(z_{2m}) = \lim_{m\to\infty} f(y_m) = b.$$

So $a = b$.

$\boxed{\text{Step 3.}}$ (*Linearity of* $\bar f$.) Let $x, y \in \bar Z$, α, β be scalars. Then there exist $x_k, y_k \in Z$, $k \in \mathbb{N}$, such that $x_k \to x$, $y_k \to y$ in X. If $x \in Z$, then we may assume that all $x_k = x$, if $y \in Z$, then, respectively, we assume that all $y_k = y$. Since f is linear, we get by using (4.2) that

$$\bar f(\alpha x + \beta y) = \bar f\Big(\alpha \lim_{k\to\infty} x_k + \beta \lim_{k\to\infty} y_k\Big)$$
$$= \bar f\Big(\lim_{k\to\infty}(\alpha x_k + \beta y_k)\Big) = \lim_{k\to\infty} \bar f(\alpha x_k + \beta y_k) = \lim_{k\to\infty} f(\alpha x_k + \beta y_k)$$
$$= \lim_{k\to\infty}(\alpha f(x_k) + \beta f(y_k)) = \alpha \lim_{k\to\infty} f(x_k) + \beta \lim_{k\to\infty} f(y_k) = \alpha \bar f(x) + \beta \bar f(y).$$

$\boxed{\text{Step 4.}}$ (*Formula* (4.1).) Clearly, since $\bar f(x) = f(x)$ for $x \in Z$

$$\|f\|_{Z\to Y} = \sup_{\substack{x\in Z\\ x\neq 0}} \frac{\|f(x)\|_Y}{\|x\|_X} \le \sup_{\substack{x\in \bar Z\\ x\neq 0}} \frac{\|\bar f(x)\|_Y}{\|x\|_X} = \|f\|_{\bar Z\to Y}.$$

On the other hand, for any $x \in \bar Z$ there exist $x_k \in Z$ such that $\lim\limits_{k\to\infty} x_k = x$, then by (4.2) $\lim\limits_{k\to\infty} f(x_k) = \bar f(x)$ and, by the continuity of a norm, $\lim\limits_{k\to\infty} \|f(x_k)\|_Y = \|\bar f(x)\|_Y$. Hence, by passing to the limit in the inequality

$$\|f(x_k)\|_Y \le \|f\|_{Z\to Y}\|x_k\|_X, \quad \forall\, k \in \mathbb{N},$$

as $k \to \infty$, we get

$$\|\bar f(x)\|_Y \le \|f\|_{Z\to Y}\|x\|_X, \quad \forall\, x \in \bar Z$$

which implies that

$$\|\bar f\|_{\bar Z\to Y} = \sup_{\substack{x\in \bar Z\\ x\neq 0}} \frac{\|\bar f(x)\|_Y}{\|x\|_X} \le \|f\|_{Z\to Y} < \infty,$$

and equality (4.1) follows. We also proved that the operator $\bar f : \bar Z \to Y$ is bounded, and hence, due to its linearity, is continuous.

$\boxed{\text{Step 5.}}$ (*Uniqueness.*) Assume that $\bar{\bar f} : \bar Z \to Y$ is a continuous linear operator such that $\bar{\bar f}(x) = f(x)$ for any $x \in Z$. Then, by the continuity of $\bar{\bar f}$ and $\bar f$ on $\bar Z$ for any $x \in \bar Z$ and any sequence $x_k \in Z :\ \lim\limits_{k\to\infty} x_k = x$

$$\bar{\bar f}(x) = \lim_{k\to\infty} \bar{\bar f}(x_k) = \lim_{k\to\infty} f(x_k) = \lim_{k\to\infty} \bar f(x_k) = \bar f(x).$$

□

4.2 Extension of linear functionals

Let X be a normed space. The space of all bounded linear functionals $f : X \to \mathbb{C}$ (or $\mathbb{R}$) is called the *dual* of X and is denoted by X^* (sometimes by X').

In notation of Section 2.2, $X^* = \mathcal{L}(X, \mathbb{R})$ in the case of real normed spaces and $X^* = \mathcal{L}(X, \mathbb{C})$ in the case of complex normed spaces.

By Lemma 2.2.1, X^* is a normed space with the norm

$$\|f\|_{X^*} \equiv \|f\|_{X\to\mathbb{C}} = \sup_{\substack{x\in X\\ x\neq 0}} \frac{|f(x)|}{\|x\|_X}.$$

Also, by Lemma 2.2.2,

$$\|f\|_{X^*} = \sup_{\substack{x\in X\\ \|x\|_X\neq 1}} |f(x)| = \sup_{\substack{x\in X\\ \|x\|_X=1}} |f(x)|.$$

Theorem 4.2.1 (Hahn-Banach extension theorem). *Let f be a bounded linear functional on a linear subspace Z of a normed space X. Then there exists a bounded linear functional $\tilde{f}$ on X which is an extension of f and such that*

$$\|\tilde{f}\|_{X\to\mathbb{C}} = \|f\|_{Z\to\mathbb{C}}.$$

Proof. The case of a space X with an inner product.

$\boxed{\text{Step 1.}}$ Let $\bar{Z}$ be the closure of Z and $\bar{f}$ be the extension of f to $\bar{Z}$ constructed in Theorem 4.1.1. Note that $\bar{Z}$ is a Hilbert space. Hence, by the Riesz representation theorem, there exists $z \in \bar{Z}$ such that

$$\bar{f}(x) = (x, z) \quad \forall\, x \in \bar{Z}$$

and

$$\|\bar{f}\|_{\bar{Z}\to\mathbb{C}} = \|z\|_{\bar{Z}} = \|z\|_X.$$

$\boxed{\text{Step 2.}}$ Define $\tilde{f} : X \to \mathbb{C}$ by

$$\tilde{f}(x) = (x, z) \quad \forall\, x \in X.$$

Then, $\tilde{f}$ is an extension of $\bar{f}$ from $\bar{Z}$ to X, hence, an extension of f from Z to X, and by Lemma 3.4.1 and Theorem 4.1.1,

$$\|\tilde{f}\|_{X\to\mathbb{C}} = \|z\|_X = \|\bar{f}\|_{\bar{Z}\to\mathbb{C}} = \|f\|_{Z\to\mathbb{C}}.$$

The case of a normed space X (without proof).

□

Remark. The proof for normed spaces is much more complicated. It is based on the axiom of choice which implies the so-called Zorn lemma which, in turn, is directly used in the proof.

Theorem 4.2.2 (Bounded linear functionals). *Let X be a normed space and $x \in X,\ x \neq 0$. Then there exists a bounded linear functional f_x such that*

$$\|f_x\|_{X^*} = 1, \quad f_x(x) = \|x\|_X.$$

Proof. The case of an inner product space X.

Let

$$f_x(y) = \frac{(y, x)}{\|x\|_X} = (y, \frac{x}{\|x\|_X}) \quad \forall\, y \in X.$$

Then $f_x(x) = \|x\|_X$, and by the Cauchy-Bunyakowski inequality,

$$\|f_x\|_{X^*} = \sup_{\substack{y\in X\\ y\neq 0}} \frac{|(y, x)|}{\|y\|_X\|x\|_X} \leq 1$$

and

$$\|f_x\|_{X^*} \geq \frac{|(x,x)|}{\|x\|_X \|x\|_X} = 1,$$

hence $\|f_x\|_{X^*} = 1$.

The case of a normed space X.

Let

$$Z = \{\alpha x : \alpha \in \mathbb{C} \text{ (or } \mathbb{R})\}.$$

On Z we define the linear functional g_x by

$$g_x(y) = g_x(\alpha_y x) = \alpha_y \|x\|_X \quad \forall\, y \in Z,$$

where α_y is such that $y = \alpha_y x$.
Clearly, $g_x(x) = \|x\|_X,\ g_x$ is bounded and $\|g_x\|_{Z\to\mathbb{C}} = 1$, because

$$|g_x(y)| = |g_x(\alpha_y x)| = |\alpha_y|\|x\|_X = \|\alpha_y x\|_X = \|y\|_X,$$

hence

$$\|g_x\|_{Z^*} = \sup_{\substack{y\in Z\\ y\neq 0}} \frac{|g_x(y)|}{\|y\|_X} = 1.$$

By Theorem 4.1.1, there exists a continious linear extension f_x of g_x from Z to X such that

$$\|f_x\|_{X^*} = \|g_x\|_{Z^*} = 1 \quad \text{and} \quad f_x(x) = g_x(x) = \|x\|_X$$

since $x \in Z$. □

4.3 Properties of dual spaces

Theorem 4.3.1 (Duality formula). *For every x in a normed space X*

$$\|x\|_X = \sup_{\substack{f\in X^*\\ f\neq 0}} \frac{|f(x)|}{\|f\|_{X^*}}.$$

Proof. Note that $\forall\, x \in X$ and $\forall\, f \in X^*$

$$|f(x)| \leq \|f\|_{X^*}\|x\|_X,$$

hence,

$$\sup_{\substack{f\in X^*\\ f\neq 0}} \frac{|f(x)|}{\|f\|_{X^*}} \leq \|x\|_X.$$

Also

$$\sup_{\substack{f\in X^*\\ f\neq 0}} \frac{|f(x)|}{\|f\|_{X^*}} \geq \frac{|f_x(x)|}{\|f_x\|_{X^*}} = \|x\|_X,$$

and the desired statement follows. □

Theorem 4.3.2. *Let X be a normed space. Then X^* is a Banach space.*

Proof. It has been already noted that X^* is a normed space. Assume that $\{f_k\}_{k\in\mathbb{N}} \subset X^*$ is a Cauchy sequence, that is, $\forall\, \varepsilon > 0\ \exists\, n \in \mathbb{N}$ such that $\forall\, k, m \in \mathbb{N},\ k, m \geq n$

$$\|f_k - f_m\|_{X^*} < \varepsilon.$$

Since

$$|f_k(x) - f_m(x)| = |(f_k - f_m)(x)| \leq \|f_k - f_m\|_{X^*}\|x\|_X \quad \forall\, x \in X,$$

we get that for any $x \in X\ \forall\, k, m \in \mathbb{N},\ k, m \geq n$

$$|f_k(x) - f_m(x)| < \varepsilon\|x\|_X. \tag{4.3}$$

So, for any $x \in X$ the sequence $\{f_k\}_{k\in\mathbb{N}}$ is a Cauchy sequence in $\mathbb{R}$ for the case of a real normed space X or in $\mathbb{C}$ for the case of a complex normed space X. Since both $\mathbb{R}$ and $\mathbb{C}$ are complete spaces, there exists $y(x) \in \mathbb{R},\ y(x) \in \mathbb{C}$ respectively, such that $\lim\limits_{k\to\infty} f_k(x) = y(x)$.

We define the functional f on X by setting

$$f(x) = y(x) \quad \forall\, x \in X. \tag{4.4}$$

By passing to the limit in (4.3) as $m \to \infty$, we get that for any $x \in X\ \forall\, k \in \mathbb{N},\ k \geq n$

$$|f_k(x) - f(x)| \leq \varepsilon\|x\|_X,$$

hence,

$$\|f_k - f\|_{X^*} = \sup_{\substack{x\in X\\ x\neq 0}} \frac{|f_k(x) - f(x)|}{\|x\|_X} \leq \varepsilon,$$

which means that

$$\lim_{k\to\infty} \|f_k - f\|_{X^*} = 0. \tag{4.5}$$

The functional f is linear. Indeed, given $x_1, x_2 \in X$ and scalars α_1, α_2 by (4.4)

$$\lim_{k\to\infty} f_k(x_1) = y(x_1), \quad \lim_{k\to\infty} f_k(x_2) = y(x_2), \quad \lim_{k\to\infty} f_k(\alpha_1 x_1 + \alpha_2 x_2) = y(\alpha_1 x_1 + \alpha_2 x_2).$$

Since all f_k are linear,

$$\begin{aligned} \lim_{k\to\infty} f_k(\alpha_1 x_1 + \alpha_2 x_2) &= \lim_{k\to\infty} (\alpha_1 f_k(x_1) + \alpha_2 f_k(x_2)) \\ &= \alpha_1 \lim_{k\to\infty} f_k(x_1) + \alpha_2 \lim_{k\to\infty} f_k(x_2) = \alpha_1 y(x_1) + \alpha_2 y(x_2), \end{aligned}$$

hence, by the uniqueness of a limit, $y(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 y(x_1) + \alpha_2 y(x_2)$ and by (4.3) $f(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 f_k(x_1) + \alpha_2 f_k(x_2)$.

It is also bounded because for $k \geq n$

$$\|f\|_{X^*} = \|f - f_k + f_k\|_{X^*} \leq \|f - f_k\|_{X^*} + \|f_k\|_{X^*} \leq \varepsilon + \|f_k\|_{X^*} < \infty.$$

Thus, $f \in X^*$ and by (4.5), $\lim\limits_{k\to\infty} f_k = f$ in X^*. $\square$

Definition. *Two normed spaces X and Y are said to be isometrically isomorphic if there exists a one-to-one linear mapping $J : X \to Y$ onto Y ($\iff J(X) = Y$) such that*

$$\|Jx\|_Y = \|x\|_X \quad \forall\, x \in X. \tag{4.6}$$

In this case we write $X \simeq Y$.

We shall also use this notation in the case in which there exists a one-to-one conjugate linear mapping $J : X \to Y$ such that equality (4.6) holds. (Conjugate linear means that $J(\alpha_1 x_1 + \alpha_2 x_2) = \overline{\alpha_1} J(x_1) + \overline{\alpha_2} J(x_2)$ for any $x_1, x_2 \in X$ and for any $\alpha_1, \alpha_2 \in \mathbb{C}$, and we say that the spaces X and Y are isometrically conjugate isomorphic.)

4.3.1 The case of Hilbert spaces

Theorem 4.3.3. *Let H be a Hilbert space. Then H^* is also a Hilbert space and*

$$H^* \simeq H. \tag{4.7}$$

Proof. $\boxed{\text{Step 1.}}$ By the Riesz representation theorem, there exists a uniquely defined element $z \in H$ such that

$$f(x) = (x, z)_H \quad \forall\, x \in H. \tag{4.8}$$

Let the mapping $J : H^* \to H$ be defined by

$$J(f) = z \quad \forall\, f \in H^*. \tag{4.9}$$

This mapping is onto H (because for any $z \in H$ the functional f defined by (4.8) belongs to H^*), one-to-one and such that

$$\|J(f)\|_H = \|z\|_H = \|f\|_{H^*} \quad \forall\, f \in H^* \tag{4.10}$$

(by the Riesz representation theorem).

Moreover, for any $f_1, f_2 \in H^*$ and any scalars α_1, α_2

$$J(\alpha_1 f_1 + \alpha_2 f_2) = \overline{\alpha_1} J(f_1) + \overline{\alpha_2} J(f_2). \tag{4.11}$$

Indeed, let $z_1, z_2 \in H$ be such that

$$f_1(x) = (x, z_1)_H, \quad f_2(x) = (x, z_2)_H \quad \forall\, x \in H.$$

Then

$$\begin{aligned}(\alpha_1 f_1 + \alpha_2 f_2)(x) &= (\alpha_1 f_1)(x) + (\alpha_2 f_2)(x) = \alpha_1 f_1(x) + \alpha_2 f_2(x) \\ &= \alpha_1 (x, z_1)_H + \alpha_2 (x, z_2)_H = (x, \overline{\alpha_1} z_1 + \overline{\alpha_2} z_2)_H,\end{aligned}$$

hence, by (4.9),

$$J(\alpha_1 f_1 + \alpha_2 f_2) = \overline{\alpha_1} z_1 + \overline{\alpha_2} z_2 = \overline{\alpha_1} J(f_1) + \overline{\alpha_2} J(f_2).$$

Equality (4.10) means that for real Hilbert spaces, J is a linear mapping and for complex Hilbert spaces, J is a conjugate linear mapping.

$\boxed{\text{Step 2.}}$ Let the inner product on H^* be defined by

$$(f, g)_{H^*} = \overline{(Jf, Jg)_H} \quad \forall\, f, g \in H. \tag{4.12}$$

This is indeed an inner product on H^* because

1) $(f, f)_{H^*} = \overline{(Jf, Jf)_H} = 0 \iff (Jf, Jf)_H = 0 \iff Jf = 0 \iff f = 0$ since J is one-to-one and $J(0) = 0$,

2) for any $f, g \in H^*$

$$(g, f)_{H^*} = \overline{(Jg, Jf)_H} = (Jf, Jg)_H = \overline{\overline{((Jf, Jg)_H)}} = \overline{(f, g)_{H^*}},$$

3) for any $f_1, f_2, g \in H^*$ and any scalars α_1, α_2 by (4.11)

$$\begin{aligned}(\alpha_1 f_1 + \alpha_2 f_2, g)_{H^*} &= \overline{(J(\alpha_1 f_1 + \alpha_2 f_2), Jg)_H} = \overline{(\overline{\alpha_1} Jf_1 + \overline{\alpha_2} Jf_2, Jg)_H} \\ &= \overline{\overline{\alpha_1}(Jf_1, Jg)_H + \overline{\alpha_2}(Jf_2, Jg)_H} = \alpha_1 \overline{(Jf_1, Jg)_H} + \alpha_2 \overline{(Jf_2, Jg)_H} \\ &= \alpha_1 (f_1, g)_{H^*} + \alpha_2 (f_2, g)_{H^*}.\end{aligned}$$

$\boxed{\text{Step 3.}}$ The space H^* is complete. Indeed, let $\{f_k\}_{k\in\mathbb{N}} \subset H^*$ be such that

$$\lim_{k,m\to\infty}(f_k - f_m) = 0 \quad \text{in} \quad H^*$$
$$\Updownarrow$$
$$\lim_{k,m\to\infty} \|f_k - f_m\|_{H^*} = \lim_{k,m\to\infty} \|J(f_k - f_m)\|_H = \lim_{k,m\to\infty} \|Jf_k - Jf_m)\|_H = 0.$$

By the completeness of H there exists $z \in H$ such that $Jf_k \to z$ in H. Let $f = J^{-1}z$, then

$$\begin{aligned}\lim_{k\to\infty} \|f_k - f\|_{H^*} &= \lim_{k\to\infty} \|J(f_k - f)\|_H = \lim_{k\to\infty} \|Jf_k - Jf\|_H \\ &= \lim_{k\to\infty} \|Jf_k - z)\|_H = 0.\end{aligned}$$

Moreover, since
$$|f_k(x) - f(x)| \le \|f_k - f\|_{H^*}\|x\|_H \quad \forall\, x \in H,$$
it follows that
$$\lim_{k\to\infty} f_k(x) = f(x) \quad \forall\, x \in H$$
and for any $x_1, x_2 \in H$ and any scalars α_1, α_2

$$\begin{aligned}f(\alpha_1x_1 + \alpha_2x_2) &= \lim_{k\to\infty} f_k(\alpha_1x_1 + \alpha_2x_2) = \lim_{k\to\infty}(\alpha_1f_k(x_1) + \alpha_2f_k(x_2)) \\ &= \alpha_1 \lim_{k\to\infty} f_k(x_1) + \alpha_2 \lim_{k\to\infty} f_k(x_2) = \alpha_1f(x_1) + \alpha_2f(x_2),\end{aligned}$$

so f is a linear functional. Also, $\|f\|_{H^*} = \|z\|_H < \infty$. Thus, $f \in H^*$ and $\lim_{k\to\infty} f_k = f$ in H^*. $\square$

Remark 4.3.1. Property (4.11) implies that for any real Hilbert space the dual space H^* is isometrically isomorphic to H and for any complex Hilbert space the dual space H^* is isometrically conjugate isomorphic to H.

Corollary. *For any Hilbert space H*

$$H^{**} \equiv (H^*)^* \simeq H. \tag{4.13}$$

Proof. Theorem 4.3.3 implies that

$$H^{**} = (H^*)^* \simeq H^* \simeq H.$$

$\square$

Examples. 1) $(\ell_2)^* \simeq \ell_2$,

2) for any Lebesgue measurable set $\Omega \subset \mathbb{R}^n$ $\quad (L_2(\Omega))^* \simeq L_2(\Omega)$.

4.3.2 The case of spaces ℓ_p

Lemma 4.3.1. *Let $\beta = (\beta_k)_{k\in\mathbb{N}} \equiv (\beta_1, \ldots, \beta_k, \ldots) \in \ell_\infty$ and*

$$f(x) = \sum_{k=1}^{\infty} x_k\beta_k \quad \forall\, x = (x_k)_{k\in\mathbb{N}} \in \ell_1. \tag{4.14}$$

Then $f \in (\ell_1)^*$ and

$$\|f\|_{(\ell_1)^*} = \|\beta\|_{\ell_\infty}. \tag{4.15}$$

Proof. Since

$$\sum_{k=1}^{\infty}|x_k\beta_k| = \sum_{k=1}^{\infty}|x_k|\cdot|\beta_k| \le \sup_{k\in\mathbb{N}}|\beta_k|\sum_{k=1}^{\infty}|x_k| = \|\beta\|_{\ell_\infty}\|x\|_{\ell_1},$$

the series $\sum\limits_{k=1}^{\infty} x_k\beta_k$ converges and the functional f is well-defined. Moreover,

$$|f(x)| \le \|\beta\|_{\ell_\infty}\|x\|_{\ell_1},$$

hence,

$$\|f\|_{(\ell_1)^*} = \sup_{\substack{x\in\ell_1\\ x\neq 0}}\frac{|f(x)|}{\|x\|_{\ell_1}} \le \|\beta\|_{\ell_\infty}.$$

On the other hand, by choosing $x = e_k \equiv (0,\dots,0,\underbrace{1}_{\substack{k^{\text{th}}\\ \text{position}}},0,\dots)$, we get

$$\|f\|_{(\ell_1)^*} = \sup_{k\in\mathbb{N}}|f(e_k)| = \sup_{k\in\mathbb{N}}|\beta_k| = \|\beta\|_{\ell_\infty},$$

and equality (4.15) follows. □

Theorem 4.3.4. *Any functional $f \in (\ell_1)^*$ has the form* (4.15) *for some $\beta \in \ell_\infty$,*

$$(\ell_1)^* \simeq \ell_\infty \tag{4.16}$$

and $(\ell_1)^*$ is isometrically isomorphic to ℓ_∞.

Proof. $\boxed{\text{Step 1.}}$ The set $\{e_k\}_{k\in\mathbb{N}}$ is a basis for ℓ_1 and for any $x = (x_1,\dots,x_k,\dots) \in \ell_1$

$$x = \sum_{k=1}^{\infty} x_k e_k. \tag{4.17}$$

Let $f \in (\ell_1)^*$, then by the continuity and linearity of f

$$\begin{aligned} f(x) &= f\Big(\sum_{k=1}^{\infty} x_k e_k\Big) = f\Big(\lim_{m\to\infty}\sum_{k=1}^{m} x_k e_k\Big) = \lim_{m\to\infty} f\Big(\sum_{k=1}^{m} x_k e_k\Big) \\ &= \lim_{m\to\infty}\sum_{k=1}^{m} x_k f(e_k) = \sum_{k=1}^{\infty} x_k f(e_k) = \sum_{k=1}^{\infty} x_k\beta_k, \end{aligned} \tag{4.18}$$

where $\beta_k = f(e_k)$. Note that for any $k \in \mathbb{N}$

$$|\beta_k| = |f(e_k)| \le \|f\|_{(\ell_1)^*}\|e_k\|_{\ell_1} = \|f\|_{(\ell_1)^*},$$

hence,

$$\|\beta\|_{\ell_\infty} = \sup_{k\in\mathbb{N}}|\beta_k| \le \|f\|_{(\ell_1)^*} < \infty.$$

So, equality (4.14) holds and $\beta \in \ell_\infty$.

$\boxed{\text{Step 2.}}$ Let the mapping $J : (\ell_1)^* \to \ell_\infty$ be defined by the following equality

$$J(f) = (f(e_k))_{k\in\mathbb{N}} \equiv (f(e_1),\dots,f(e_k),\dots). \tag{4.19}$$

By Lemma 4.3.1, $J((\ell_1)^*) = \ell_\infty$. Moreover, J is one-to-one because if $f_1, f_2 \in (\ell_1)^*$ and $J(f_1) = J(f_2)$, then $f_1(e_k) = f_2(e_k)$ for any $k \in \mathbb{N}$, hence, for any $x \in \ell_1$ represented in form (4.17), by equality (4.18) with f replaced by f_1 and f_2

$$f_1(x) = \sum_{k=1}^{\infty} x_k f_1(e_k) = \sum_{k=1}^{\infty} x_k f_2(e_k) = f_2(x),$$

which means that $f_1 = f_2$.

Also J is a linear mapping because, given $f_1, f_2 \in (\ell_1)^*$ and $\alpha_1, \alpha_2 \in \mathbb{C}$, we have

$$\begin{aligned} J(\alpha_1 f_1 + \alpha_2 f_2) &= ((\alpha_1 f_1 + \alpha_2 f_2)(e_k))_{k\in\mathbb{N}} = (\alpha_1 f_1(e_k) + \alpha_2 f_2(e_k))_{k\in\mathbb{N}} \\ &= \alpha_1 (f_1(e_k))_{k\in\mathbb{N}} + \alpha_2 (f_2(e_k))_{k\in\mathbb{N}} = \alpha_1 J(f_1) + \alpha_2 J(f_2). \end{aligned}$$

So J is an isomorphism. Moreover, J is an isometric isomorphism, because by Lemma 4.3.1,

$$\|J(f)\|_{\ell_\infty} = \|(f(e_k))_{k\in\mathbb{N}}\|_{\ell_\infty} = \|f\|_{(\ell_1)^*}$$

for any $f \in (\ell_1)^*$. □

Lemma 4.3.2. *Let* $1 < p < \infty$, $p' = \frac{p}{p-1}$ $\left(\iff \frac{1}{p} + \frac{1}{p'} = 1\right)$, $\beta = (\beta_k)_{k\in\mathbb{N}} \in \ell_{p'}$ *and*

$$f(x) = \sum_{k=1}^{\infty} x_k \beta_k \quad \forall\, x = (x_k)_{k\in\mathbb{N}} \in \ell_p. \tag{4.20}$$

Then $f \in (\ell_1)^*$ *and*

$$\|f\|_{(\ell_1)^*} = \|\beta\|_{\ell_{p'}}. \tag{4.21}$$

Proof. By the Hölder inequality

$$\sum_{k=1}^{\infty} |x_k \beta_k| \le \|x\|_{\ell_p} \|\beta\|_{\ell_{p'}},$$

hence, the series $\sum_{k=1}^{\infty} x_k \beta_k$ converges, the functional f is well-defined and

$$|f(x)| \le \|x\|_{\ell_p} \|\beta\|_{\ell_{p'}},$$

which implies that

$$\|f\|_{(\ell_1)^*} \le \|\beta\|_{\ell_{p'}}.$$

On the other hand,

$$\|f\|_{(\ell_1)^*} \ge \frac{|f(x)|}{\|x\|_{\ell_p}},$$

where $x = (x_k)_{k\in\mathbb{N}}$,

$$x_k = \begin{cases} \dfrac{|\beta_k|^{p'}}{\beta_k} & \text{if } \beta_k \ne 0, \\ 0 & \text{if } \beta_k = 0. \end{cases} \tag{4.22}$$

Since

$$\|x\|_{\ell_p} = \Big(\sum_{\substack{k=1:\\ \beta_k\neq 0}}^{\infty} \Big(|\beta_k|^{\frac{p}{p-1}-1}\Big)^p\Big)^{\frac{1}{p}} = \Big(\sum_{k=1}^{\infty} |\beta_k|^{p'}\Big)^{\frac{1}{p}}, \tag{4.23}$$

we have

$$\|f\|_{(\ell_1)^*} \geq \frac{\Big(\sum\limits_{\substack{k=1:\\ \beta_k\neq 0}}^{\infty} \frac{|\beta_k|^{p'}}{\beta_k}\cdot\beta_k\Big)}{\Big(\sum\limits_{k=1}^{\infty} |\beta_k|^{p'}\Big)^{\frac{1}{p}}} = \frac{\|\beta\|_{\ell_{p'}}^{p'}}{\|\beta\|_{\ell_{p'}}^{\frac{p'}{p}}} = \|\beta\|_{\ell_{p'}}^{p'(1-\frac{1}{p})} = \|\beta\|_{\ell_{p'}},$$

and equality (4.21) follows. □

Theorem 4.3.5. *Let $1 < p < \infty$. Any functional $f \in (\ell_1)^*$ has form (4.21) for some $\beta \in \ell_{p'}$,*

$$(\ell_1)^* \simeq \ell_{p'}, \tag{4.24}$$

and $(\ell_1)^*$ is isometrically isomorphic to $\ell_{p'}$.

Proof. $\boxed{\text{Step 1.}}$ Let $f \in (\ell_1)^*$, $f \neq 0$ (the case $f = 0$ being trivial). As in the proof of Theorem 4.3.4, equalities (4.17) and (4.18) hold for any sequence $x = (x_k)_{k\in\mathbb{N}} \in \ell_p$. Moreover, for any $x \in \ell_p$

$$|f(x)| \leq \|f\|_{(\ell_p)^*}\|x\|_{\ell_p}. \tag{4.25}$$

Let in this inequality x be replaced by $x^{(m)} = (x_1, \ldots, x_m, 0, \ldots)$, where x_k are defined by formula (4.22) with $\beta_k = f(e_k)$ and $m \in \mathbb{N}$ is sufficiently large, so that $x^{(m)} \neq 0$. By (4.23),

$$\|x^{(m)}\|_{\ell_p} = \Big(\sum_{k=1}^{\infty} |\beta_k|^{p'}\Big)^{\frac{1}{p}} = \|\beta^{(m)}\|_{\ell_{p'}}^{\frac{p'}{p}}.$$

Furthermore,

$$|f(x^{(m)})| = \sum_{\substack{k=1:\\ \beta_k\neq 0}}^{m} \frac{|\beta_k|^{p'}}{\beta_k}\cdot\beta_k = \|\beta^{(m)}\|_{\ell_{p'}}^{p'}.$$

Hence,

$$\frac{|f(x^{(m)})|}{\|x^{(m)}\|_{\ell_p}} = \|\beta^{(m)}\|_{\ell_{p'}}^{\frac{p'}{p}-p'} = \|\beta^{(m)}\|_{\ell_{p'}} \leq \|f\|_{(\ell_p)^*}.$$

By passing here to the limit as $m \to \infty$, we get

$$\|\beta\|_{\ell_{p'}} = \lim_{m\to\infty} \|\beta^{(m)}\|_{\ell_{p'}} \leq \|f\|_{(\ell_p)^*} < \infty.$$

So, equality (4.18) holds for any $x \in \ell_p$ where $\beta = (f(e_k))_{k\in\mathbb{N}} \in \ell_{p'}$.

$\boxed{\text{Step 2.}}$ Let the mapping $J : (\ell_1)^* \to \ell_\infty$ be defined by formula (4.19) for any $f \in (\ell_1)^*$. An argument similar to the argument in Step 2 of Theorem 4.3.4 shows that J is an isometric isomorphism. □

Corollary. *For any $1 < p < \infty$*

$$(\ell_p)^{**} \simeq \ell_p. \tag{4.26}$$

Proof. Indeed,
$$(\ell_p)^{**} = ((\ell_p)^*)^* \simeq (\ell_{p'})^* \simeq \ell_{(p')'} = \ell_p.$$
□

Lemma 4.3.3. *Let $\beta = (\beta_k)_{k\in\mathbb{N}} \in \ell_1$ and*

$$f(x) = \sum_{k=1}^{\infty} x_k\beta_k \quad \forall\, x = (x_k)_{k\in\mathbb{N}} \in \ell_\infty. \tag{4.27}$$

Then $f \in (\ell_\infty)^*$ and

$$\|f\|_{(\ell_\infty)^*} = \|\beta\|_{\ell_1}. \tag{4.28}$$

Proof. Since

$$\sum_{k=1}^{\infty} |x_k\beta_k| = \sum_{k=1}^{\infty} |x_k|\cdot|\beta_k| \le \sup_{k\in\mathbb{N}} |x_k| \sum_{k=1}^{\infty} |\beta_k| = \|\beta\|_{\ell_1}\|x\|_{\ell_\infty} < \infty,$$

the series $\sum\limits_{k=1}^{\infty} x_k\beta_k$ converges, the functional f is well-defined and

$$|f(x)| \le \|\beta\|_{\ell_1}\|x\|_{\ell_\infty},$$

hence,

$$\|f\|_{(\ell_\infty)'} \le \|\beta\|_{\ell_1}.$$

On the other hand

$$\|f\|_{(\ell_\infty)'} \ge \frac{\left|\sum\limits_{k=1}^{\infty} x_k\beta_k\right|}{\|x\|_{\ell_\infty}},$$

where $x_k = \frac{|\beta_k|}{\beta_k}$ if $\beta_k \ne 0$ and $x_k = 1$ if $\beta_k = 0$ (hence $|x_k| = 1$ for all $k \in \mathbb{N}$), hence,

$$\|f\|_{(\ell_\infty)'} \ge \sum_{\substack{k=1:\\ \beta_k\ne 0}}^{\infty} \frac{|\beta_k|}{\beta_k}\cdot\beta_k = \sum_{k=1}^{\infty} |\beta_k| = \|\beta\|_{\ell_1}.$$

□

Remark 4.3.2. It can be proved that

$$(\ell_\infty)^* \not\simeq \ell_1. \tag{4.29}$$

Hence,

$$(\ell_1)^{**} = (\ell_1^*)^* \simeq (\ell_\infty)^* \not\simeq \ell_1. \tag{4.30}$$

4.3.3 The case of spaces $L_p(\Omega)$

Theorem 4.3.6 (Lebesgue). *Let $f \in L_1^{loc}(\mathbb{R}^n)$ ($f \in L_1(B(x,r))$ for all $x \in \mathbb{R}^n$ and $r > 0$), then for almost all $x \in \mathbb{R}^n$*

$$\lim_{r\to 0+} \frac{\int\limits_{B(x,r)} f(y)\,dy}{\operatorname{meas} B(x,r)} = f(x). \tag{4.31}$$

Without proof.

Corollary 4.3.1. *Let $\Omega \subset \mathbb{R}^n$ be a measurable set and* $\operatorname{meas}\Omega > 0$. *Then for almost all $x \in \Omega$*

$$\lim_{r\to 0^+} \frac{\operatorname{meas}(B(x,r)\cap\Omega)}{\operatorname{meas} B(x,r)} = 1.$$

Proof. Since $\chi_\Omega \in L_1^{loc}(\mathbb{R}^n)$ by Theorem 4.3.6 for almost all $x \in \mathbb{R}^n$

$$\lim_{r\to 0^+} \frac{\int\limits_{B(x,r)} \chi_\Omega(y)\,dy}{\operatorname{meas} B(x,r)} = \lim_{r\to 0^+} \frac{\operatorname{meas}(B(x,r)\cap\Omega)}{\operatorname{meas} B(x,r)} = \chi_\Omega(x).$$

Hence for almost all $x \in \Omega$ we get equality (4.31). □

Corollary 4.3.2. *Let $\Omega \subset \mathbb{R}^n$ be a measurable set,* $\operatorname{meas}\Omega > 0$, $f \in L_1(\Omega \cap B(x,r))$ *for all $x \in \Omega$ and $r > 0$. Then for almost all $x \in \Omega$*

$$\lim_{r\to 0^+} \frac{\int\limits_{B(x,r)\cap\Omega} f(y)\,dy}{\operatorname{meas}(B(x,r)\cap\Omega)} = f(x).$$

Proof. Let

$$\mathring{f}(x) = \begin{cases} f(x), & x \in \Omega \\ 0, & x \notin \Omega \end{cases}$$

be the extension of f by 0. Then $\mathring{f} \in L_1^{loc}(\mathbb{R}^n)$, hence, by Theorem 4.3.6 for almost all $x \in \mathbb{R}^n$

$$\lim_{r\to 0^+} \frac{\int\limits_{B(x,r)} \mathring{f}(y)\,dy}{\operatorname{meas} B(x,r)} = \mathring{f}(x).$$

Therefore, by Corollary 4.3.1 for almost all $x \in \Omega$

$$\lim_{r\to 0^+} \frac{\int\limits_{B(x,r)\cap\Omega} f(y)\,dy}{\operatorname{meas}(B(x,r)\cap\Omega)} = \lim_{r\to 0^+} \frac{\int\limits_{B(x,r)} \mathring{f}(y)\,dy}{\operatorname{meas} B(x,r)} \cdot \frac{\operatorname{meas} B(x,r)}{\operatorname{meas}(B(x,r)\cap\Omega)} = \mathring{f}(x) = f(x).$$

□

Lemma 4.3.4. *Let $\Omega \subset \mathbb{R}^n$ be a measurable set,* $\operatorname{meas}\Omega > 0$, $y \in L_\infty(\Omega)$ *and* $\forall\, x \in L_1(\Omega)$

$$f(x) = \int_\Omega x(t)y(t)\,dt.$$

Then $f \in (L_1(\Omega))^*$ and

$$\|f\|_{(L_1(\Omega))^*} = \|y\|_{L_\infty(\Omega)}.$$

Proof. Since $\forall\,\omega \subset \Omega$, with $\operatorname{meas}\omega = 0$

$$\int_\Omega |x(t)y(t)|\,dt = \int_{\Omega\setminus\omega} |x(t)y(t)|\,dt \le \sup_{t\in\Omega\setminus\omega} |y(t)| \int_\Omega |x(t)|\,dt,$$

it follows that

$$\int_\Omega |x(t)y(t)|\,dt \le \inf_{\substack{\omega\subset\Omega \\ \operatorname{meas}\omega=0}} \sup_{t\in\Omega\setminus\omega} |y(t)| \int_\Omega |x(t)|\,dt = \|x\|_{L_1(\Omega)} \|y\|_{L_\infty(\Omega)} < \infty.$$

Hence, the integral $\int\limits_\Omega x(t)y(t)\,dt$ exists and the linear functional f is well-defined.

Moreover,

$$\|f\|_{(L_1(\Omega))^*} = \sup_{\substack{x\in L_1(\Omega)\\ x\neq 0}} \frac{|f(x)|}{\|x\|_{L_1(\Omega)}} \le \|y\|_{L_\infty(\Omega)}.$$

On the other hand

$$\|f\|_{(L_1(\Omega))^*} \ge \lim_{r\to 0^+} \frac{|f(x)|}{\|x\|_{L_1(\Omega)}},$$

where

$$x(t) = \frac{\chi_{B(\tau,r)\cap\Omega}(t)}{\|\chi_{B(\tau,r)\cap\Omega}\|_{L_1(\Omega)}} = \frac{\chi_{B(\tau,r)\cap\Omega}(t)}{\operatorname{meas}(B(\tau,r)\cap\Omega)}.$$

Since $\|x\|_{L_1(\Omega)}$, by Corollary 4.3.2 for almost all $\tau\in\Omega$

$$\|f\|_{(L_1(\Omega))^*} \ge \lim_{r\to 0^+} \frac{\left|\int\limits_\Omega \chi_{B(\tau,r)\cap\Omega}(t)y(t)\,dy\right|}{\operatorname{meas}(B(\tau,r)\cap\Omega)} = \lim_{r\to 0^+} \frac{\left|\int\limits_{B(\tau,r)\cap\Omega} y(t)\,dy\right|}{\operatorname{meas}(B(\tau,r)\cap\Omega)} = |y(\tau)|.$$

Let $\Omega_0\subset\Omega$ be the set of all $\tau\in\Omega$ for which this equality holds and $\omega_0 = \Omega\setminus\Omega_0$. Then $\operatorname{meas}\omega_0 = 0$ and

$$\|f\|_{(L_1(\Omega))'} \ge \sup_{\tau\in\Omega\setminus\omega_0} |y(\tau)| \ge \inf_{\substack{\omega_0\subset\Omega\\ \operatorname{meas}\omega_0=0}} \sup_{\tau\in\Omega\setminus\omega_0} |y(\tau)| = \|y\|_{L_\infty(\Omega)}.$$

□

Theorem 4.3.7. *Let $\Omega\subset\mathbb{R}^n$ be a measurable set,* $\operatorname{meas}\Omega>0$. *Then each linear functional on $L_1(\Omega)$ has the form*

$$f(x) = \int_\Omega x(t)y(t)\,dt$$

for some $y\in L_\infty(\Omega)$,

$$(L_1(\Omega))^* \simeq L_\infty(\Omega),$$

and the dual space $(L_1(\Omega))^*$ is isometrically isomorphic to $L_\infty(\Omega)$.

Without proof.

Lemma 4.3.5. *Let $\Omega\subset\mathbb{R}^n$ be a measurable set,* $\operatorname{meas}\Omega>0$, $1<p<\infty$, *$y\in L_{p'}(\Omega)$ and $\forall\, x\in L_p(\Omega)$*

$$f(x) = \int_\Omega x(t)y(t)\,dt.$$

Then $f\in (L_p(\Omega))^*$ and

$$\|f\|_{(L_p(\Omega))^*} = \|y\|_{L_{p'}(\Omega)}.$$

Proof. By Hölder's inequality

$$|f(x)| = \int_\Omega |x(t)y(t)|\,dt \le \|y\|_{L_{p'}(\Omega)}\|x\|_{L_p(\Omega)}.$$

Hence,

$$\|f\|_{(L_p(\Omega))^*} = \sup_{\substack{x\in L_p(\Omega)\\ x\neq 0}} \frac{|f(x)|}{\|x\|_{L_p(\Omega)}} \le \|y\|_{L_{p'}(\Omega)}.$$

On the other hand,

$$\|f\|_{(L_p(\Omega))^*} \geq \lim_{k\to\infty} \frac{|f(x_k)|}{\|x_k\|_{L_p(\Omega)}} = \lim_{k\to\infty} \frac{\left|\int\limits_\Omega x_k(t)y(t)\,dt\right|}{\|x_k\|_{L_p(\Omega)}},$$

where

$$x_k(t) = x(t)\chi_k(t), \qquad x(t) = \begin{cases} \frac{|y(t)|^{p'}}{y(t)} & \text{if } y(t) \neq 0, \\ 0 & \text{if } y(t) = 0, \end{cases}$$

and $\chi_k(t) \equiv \chi_{\Omega_k}(t)$ is the characteristic function of the set

$$\Omega_k = \{x \in \Omega : |x| \leq k,\ |y(t)| \leq k\}.$$

Then

$$\|x_k\|_{L_p(\Omega)} = \Big(\int_\Omega |x_k(t)|^p\,dt\Big)^{\frac{1}{p}} = \Big(\int_\Omega |y(t)|^{(p'-1)p}\,dt\Big)^{\frac{1}{p}} = \Big(\int_\Omega |y(t)|^{p'}\,dt\Big)^{\frac{1}{p}} = \|y\|^{\frac{p'}{p}}_{L_p(\Omega)} < \infty.$$

□

Lemma 4.3.6. *Let $\Omega \subset \mathbb{R}^n$ be a measurable set,* meas $\Omega > 0$, $y \in L_1(\Omega)$ *and for each* $x \in L_\infty(\Omega)$

$$f(x) = \int_\Omega x(t)y(t)\,dt. \tag{4.32}$$

Then $f \in (L_\infty(\Omega))'$ and

$$\|f\|_{(L_\infty(\Omega))'} = \|y\|_{L_1(\Omega)}. \tag{4.33}$$

Proof. Since

$$\begin{aligned} |f(x)| &= \Big|\int_\Omega x(t)y(t)\,dt\Big| \leq \int_\Omega |x(t)|\cdot|y(t)|\,dt \\ &\leq \|x\|_{L_\infty(\Omega)} \int_\Omega |y(t)|\,dt = \|x\|_{L_\infty(\Omega)}\|y\|_{L_1(\Omega)}, \end{aligned}$$

it follows that

$$\|f\|_{(L_\infty(\Omega))'} = \sup_{\substack{x\in L_\infty(\Omega)\\ x\nsim 0}} \frac{|f(x)|}{\|x\|_{L_\infty(\Omega)}} \leq \|y\|_{L_1(\Omega)}.$$

Hence (4.33) follows with the inequality sign instead of the equality sign.

Next, let for $t \in \Omega$

$$x_0(t) = \begin{cases} \frac{|y(t)|}{y(t)} & \text{if } y(t) \neq 0, \\ 0 & \text{if } y(t) = 0. \end{cases}$$

(If y is real-valued, then $x_0(t) = \operatorname{sgn} y(t)$.) Clearly, $x_0 \in L_\infty(\Omega)$ and $\|x_0\|_{L_\infty(\Omega)} = 1$. Hence,

$$\|f\|_{(L_\infty(\Omega))'} \geq \frac{|f(x_0)|}{\|x_0\|_{L_\infty(\Omega)}} = \int_G \frac{|y(t)|}{y(t)} y(t)\,dt = \|y\|_{L_1(G)} = \|y\|_{L_1(\Omega)},$$

where $G = \{t \in \Omega : y(t) \neq 0\}$. □

Lemma 4.3.7. *Let $\Omega \subset \mathbb{R}^n$ be a compact set,* $\operatorname{meas}\Omega > 0$, $y \in L_1(\Omega)$ *and for each* $x \in C(\Omega)$ *equality* (4.32) *holds.*

Then $f \in (C(\Omega))'$ and

$$\|f\|_{(C(\Omega))'} = \|y\|_{L_1(\Omega)}. \tag{4.34}$$

Proof. The proof is similar to the proof of Lemma 4.3.6, but the problem is that, in general, $x_0 \notin C(\Omega)$. However, it can be proved that there exist $x_k \in C(\Omega)$, $k \in \mathbb{N}$, such that

$$\|x_k\|_{C(\Omega)} \le 1, \quad \lim_{k\to\infty} \|x_k\|_{C(\Omega)} = 1 \quad \text{and} \quad \lim_{k\to\infty} x_k(t) = x_0(t)$$

for almost all $t \in \Omega$.

Since $|x_k(t)y(t)| \le |y(t)| \in L_1(\Omega)\ \forall\, k \in \mathbb{N}$, by the Dominated Convergence Theorem

$$\begin{aligned} \|f\|_{(C(\Omega))'} &\ge \lim_{k\to\infty} \frac{|f(x_k)|}{\|x_k\|_{C(\Omega)}} = \lim_{k\to\infty} \Big| \int_\Omega x_k(t)y(t)\,dt \Big| \\ &= \Big| \int_\Omega x_0(t)y(t)\,dt \Big| = \|y\|_{L_1(\Omega)}, \end{aligned}$$

which completes the proof. □

Lemma 4.3.8. *Let $\Omega \subset \mathbb{R}^n$ be a compact set,* $\operatorname{meas}\Omega > 0$, $\tau \in \Omega$ *and for each* $x \in C(\Omega)$

$$\sigma_\tau(x) = x(\tau).$$

Then $\sigma_\tau \in (C(\Omega))'$ and

$$\|\sigma_\tau\|_{(C(\Omega))'} = 1. \tag{4.35}$$

This functional cannot be represented in the form

$$\sigma_\tau(x) = \int_\Omega x(t)y(\tau)\,d\tau \quad \forall\, x \in C(\Omega) \tag{4.36}$$

for any $y \in L_1(\Omega)$.

Proof. $\boxed{\text{Step 1.}}$ Clearly

$$\|\sigma_\tau\|_{(C(\Omega))'} = \sup_{\substack{x\in C(\Omega)\\ x\neq 0}} \frac{|x(\tau)|}{\|x\|_{C(\Omega)}} \le 1.$$

On the other hand,

$$\|\sigma_\tau\|_{(C(\Omega))'} \ge \frac{|x(\tau)|}{\|x\|_{C(\Omega)}}\Big|_{x(t)\equiv 1} = 1,$$

hence, (4.35) follows.

$\boxed{\text{Step 2.}}$ Assume that for some $y \in L_1(\Omega)$, (4.36) holds. Take there $x(t) = z(t)|t-\tau|$ with an arbitrary $z \in C(\Omega)$. Then $x \in C(\Omega)$ and $x(\tau) = 0$. Hence

$$\int_\Omega z(t)|t-\tau|y(t)\,dt = 0 \quad \forall\, z \in C(\Omega).$$

By the main lemma of variational calculus

$$|t-\tau|y(t) = 0$$

for almost all $t \in \Omega$, hence

$$y(t) = 0$$

for almost all $t \in \Omega$. So

$$\sigma_\tau(x) = \int_\Omega x(t)y(t)\,dt = 0 \neq x(\tau)$$

if $x(\tau) \neq 0$. $\square$

Definition. *Let $-\infty < a < b < \infty$. A real-valued function w is said to be of bounded variation on $[a,b]$ if*

$$\mathrm{Var}(w) = \sup \sum_{k=1}^{n} |w(t_k) - w(t_{k-1})| < \infty,$$

where the supremum is taken over all partitions

$$a = t_0 < t_1 < \ldots < t_n = b$$

of the interval $[a,b]$.

Example. If w is a non-decreasing function, then

$$\mathrm{Var}(w) = w(b) - w(a),$$

if w is a non-increasing function, then

$$\mathrm{Var}(w) = w(a) - w(b).$$

Example. Let $[a,b] = [0,\pi]$, and $w(t) = \sin t$. Then

$$\mathrm{Var}(w) = 2.$$

Example. If $w \in C^1[a,b]$, then

$$\mathrm{Var}(w) = \int_a^b |w'(t)|\,dt.$$

Remark. One can prove that each function w of bounded variation on $[a,b]$ can be represented in the form

$$w(t) = w_1(t) - w_2(t),$$

where $w_1(t)$ and $w_2(t)$ are non-negative non-increasing functions.

Let $-\infty < a < b < \infty$, $x \in [a,b]$ and w be of bounded variation on $[a,b]$, briefly $w \in BV[a,b]$. Let P_n be a partition of $[a,b]$ ($a = t_0 < t_1 < \ldots < t_n = b$) and

$$\eta(P_n) = \max\{t_1 - t_0, t_2 - t_1, \ldots, t_n - t_{n-1}\}.$$

If there exists a number I with the property: $\forall\, \varepsilon > 0\ \exists\, \delta > 0$ such that if

$$\eta(P_n) < \delta,$$

then

$$|I - S(P_n)| < \varepsilon,$$

where

$$S(P_n) = \sum_{k=1}^{n} x(\xi_k)(w(t_k) - w(t_{k-1})), \quad t_{k-1} \le \xi_k \le t_k$$

is the Riemann-Stieltjes sum, then I is called the Riemann-Stieltjes integral of x over $[a,b]$ with respect to w and is denoted by

$$\int_a^b x(t)\,dw(t),$$

briefly

$$\int_a^b x(t)\,dw(t) = \lim_{\eta(P_n)\to 0} \sum_{k=1}^{n} x(\xi_k)(w(t_k) - w(t_{k-1})).$$

If $w(t) = t$, then this is the standard Riemann integral.

Properties:

1) linearity in x:

$$\int_a^b (\alpha_1 x_1(t) + \alpha_2 x_2(t))\,dw(t) = \alpha_1 \int_a^b x_1(t)\,dw(t) + \alpha_2 \int_a^b x_2(t)\,dw(t),$$

2) linearity in w:

$$\int_a^b x(t)\,d(\beta_1 w_1(t) + \beta_2 w_2(t)) = \beta_1 \int_a^b x(t)\,dw_1(t) + \beta_2 \int_a^b x(t)\,dw_2(t),$$

3) If $x \in C[a,b]$ and w is of bounded variation, then the integral $\int\limits_a^b x(t)\,dw(t)$ exists and

$$\Big| \int_a^b x(t)\,dw(t) \Big| \le \|x\|_{C[a,b]} \operatorname{Var} w.$$

4) If $x \in C[a,b]$ and $w \in C^1[a,b]$, then

$$\int_a^b x(t)\,dw(t) = \int_a^b x(t)w'(t)\,dt.$$

Theorem 4.3.8 (Riesz). *Every bounded linear functional on $C[a,b]$ can be represented by a Riemann-Stieltjes integral:*

$$f(x) = \int_a^b x(t)\,dw(t) \quad \forall\, x \in C[a,b]$$

for some $w \in BV[a,b]$.

Moreover,

$$\|f\|_{(C[a,b])'} = \operatorname{Var} w.$$

Without proof.

Lemma 4.3.9. *Let $-\infty < a < b < \infty,\ \tau \in (a,b]$,*

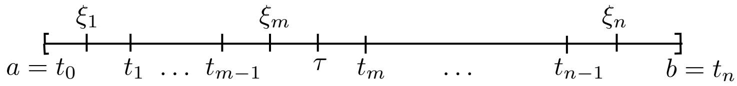
$$\sigma_\tau(x) = x(\tau) \quad \forall\, x \in C[a,b].$$

Then

$$\sigma_\tau(x) = \int_a^b x(t)\,d\chi_{[\tau,b]}(t).$$

(Recall that $\sigma_\tau(x)$ cannot be represented in the form $\sigma_\tau(x) = \int\limits_a^b x(t)y(t)\,dt \quad \forall\, x \in C[a,b]$ for any function $y \in L_1[a,b]$.)

Proof. Assume that $a < \tau < b$. For a given partition P_n we choose $m \in \mathbb{N},\ m \le n$, such that $t_{m-1} < \tau \le t_m$:



Note that

$$\begin{aligned}
S(P_n) &= \sum_{k=1}^{n} x(\xi_k)(\chi_{[\tau,b]}(t_k) - \chi_{[\tau,b]}(t_{k-1})) \\
&= \sum_{k=1}^{m-1} x(\xi_k)(\underbrace{\chi_{[\tau,b]}(t_k)}_{=0} - \underbrace{\chi_{[\tau,b]}(t_{k-1})}_{=0}) \\
&+ \ x(\xi_m)(\underbrace{\chi_{[\tau,b]}(t_m)}_{=1} - \underbrace{\chi_{[\tau,b]}(t_{m-1})}_{=0}) \\
&+ \sum_{k=m+1}^{n} x(\xi_k)(\underbrace{\chi_{[\tau,b]}(t_k)}_{=1} - \underbrace{\chi_{[\tau,b]}(t_{k-1})}_{=1}) = x(\xi_m).
\end{aligned}$$

Since $|\xi_m - \tau| \le |t_m - t_{m-1}| \le \eta(P_n) \to 0$ as $n \to \infty$, it follows that

$$\begin{aligned}
\int_a^b x(t)\, d\chi_{[\tau,b]}(t) &= \lim_{\eta(P_n)\to 0} S(P_n) = \lim_{\eta(P_n)\to 0} x(\xi_m) \\
&= \lim_{\xi_m \to \tau} x(\xi_m) = x(\tau).
\end{aligned}$$

□

Exercise. Give the proof for the case $\tau = b$.

4.4 Reflexivity

Definition. *Let X be a normed space and X' be its dual space (the space of all bounded linear functionals on X). The space*

$$X'' = (X')'$$

is called the second dual of X.

Let $x \in X$. Define the functional g_x on X' by

$$g_x(f) = f(x) \quad \forall\, f \in X'.$$

Lemma 4.4.1. *For every fixed x in a normed space X the functional g_x is a bounded linear functional on X' ($\iff g_x \in X''$) and*

$$\|g_x\|_{X''} = \|x\|_X.$$

Proof. Let $f_1, f_2 \in X'$ and $\alpha_1, \alpha_2 \in \mathbb{R}$ (or $\mathbb{C}$). Then

$$\begin{aligned} g_x(\alpha_1 f_1 + \alpha_2 f_2) = (\alpha_1 f_1 + \alpha_2 f_2)(x) &= \alpha_1 f_1(x) + \alpha_2 f_2(x) \\ &= \alpha_1 g_x(f_1) + \alpha_2 g_x(f_2). \end{aligned}$$

Also, by the duality formula,

$$\|g_x\|_{X''} = \sup_{\substack{f \in X' \\ f \neq 0}} \frac{|g_x(f)|}{\|f\|_{X'}} = \sup_{\substack{f \in X' \\ f \neq 0}} \frac{|f(x)|}{\|f\|_{X'}} = \|x\|_X.$$

□

Remark. Let the mapping $C : X \to X''$ be defined by

$$C(x) = g_x \quad \forall\, x \in X.$$

Then C is a linear one-to-one mapping from X to $C(X)$. (Indeed, if $C(x_1) = C(x_2) \iff g_{x_1} = g_{x_2} \iff f(x_1) = f(x_2)\ \forall\, f \in X'$, then $x_1 = x_2$.) By Lemma 4.4.1

$$X \simeq C(X).$$

C is called the canonical mapping from X to X''.

Definition. *A normed space X is said to be reflexive if*

$$C(X) = X''.$$

Theorem 4.4.1. *Any Hilbert space H is reflexive.*

Proof. It is required to prove that $\forall\, h \in H''$ there exists $x \in H$ such that

$$h(f) = g_x(f) \quad \forall\, f \in H'. \tag{4.37}$$

By the Riesz representation theorem $\forall\, f \in H'$ there exists $z_f \in H$ such that

$$f(x) = (x, z_f)_H \quad \forall\, x \in H. \tag{4.38}$$

It was proved that H' is a Hilbert space with the inner product defined by

$$(f, g)_{H'} = (z_g, z_f)_H \quad \forall\, f, g \in H'.$$

Since $H'' = (H')'$ and H' is a Hilbert space, then by the Riesz representation theorem $\forall\, h \in H''$ there exists $w_h \in H'$ such that

$$h(f) = (f, w_h)_{H'} = (z_{w_h}, z_f)_H \quad \forall\, f \in H'.$$

By (4.38)

$$h(f) = f(z_{w_h}) = g_{z_{w_h}}(f) \quad \forall\, f \in H',$$

hence (4.37) holds with $x = z_{w_h} \in H$. □

Remark. By Theorem 4.3.3

$$H'' = (H')' \simeq H' \simeq H.$$

This means that the space H'' is isometric to H. However, this is not a proof of reflexivity of H, because the canonical mapping C was not used.

Remark. The spaces $L_p(\Omega)$, where $\Omega \subset \mathbb{R}^n$ is a measurable set, and the spaces ℓ_p are reflexive if $1 < p < \infty$. The spaces $L_1(\Omega), L_\infty(\Omega), \ell_1, \ell_\infty$ are not reflexive. The space $C[a,b]$, $-\infty < a < b < \infty$, is not reflexive. (Without proofs.)

Theorem 4.4.2. *If a normed space X is reflexive, then it is complete (hence, a Banach space).*

Proof. Since X is reflexive, the canonical mapping $C : X \to X^{**}$ is linear, bounded, one-to-one, $C(X) = X^{**}$ and

$$\|C(x)\|_{X^{**}} = \|x\|_X \quad \forall\, x \in X,$$

hence there exists the inverse $C^{-1} : X^{**} \to X$ of C and

$$\|y\|_{X^{**}} = \|C^{-1}(y)\|_X \quad \forall\, y \in X^{**}.$$

Let $\{x_k\}_{k\in\mathbb{N}}$ be a Cauchy sequence in X:

$$\lim_{k,m\to\infty} \|x_k - x_m\|_X = 0.$$

Then

$$\begin{aligned}\lim_{k,m\to\infty} \|C(x_k) - C(x_m)\|_{X^{**}} &= \lim_{k,m\to\infty} \|C(x_k - x_m)\|_{X^{**}} \\ &= \lim_{k,m\to\infty} \|x_k - x_m\|_X = 0.\end{aligned}$$

Since the space X^{**}, being the dual of X^*, is complete, there exist $y \in X^{**}$ such that

$$\lim_{k\to\infty} \|C(x_k) - y\|_{X^{**}} = 0.$$

Hence,

$$\begin{aligned}\lim_{k\to\infty} \|x_k - C^{-1}(y)\|_X &= \lim_{k\to\infty} \|C^{-1}(C(x_k) - y)\|_X \\ &= \lim_{k\to\infty} \|C(x_k) - y\|_{X^{**}} = 0.\end{aligned}$$

So

$$\lim_{k\to\infty} x_k = C^{-1}(y) \quad \text{in} \quad X.$$

□

5. Fundamental theorems for Banach spaces

5.1 Baire's Category Theorem

Lemma 5.1.1 (Principle of Embedded Closed Balls). *Let X be a complete metric space, $\overline{B}_k \equiv \overline{B}_k(x_k, r_k) \subset X,\ k \in \mathbb{N}$, be closed balls such that*

$$\overline{B}_{k+1} \subset \overline{B}_k,\ k \in \mathbb{N}, \quad \lim_{k\to\infty} r_k = 0.$$

Then there exists a unique $x \in X$ such that $x \in \overline{B}_k$ for all $k \in \mathbb{N}$, in other words,

$$\bigcap_{k=1}^{\infty} \overline{B}_k = \{x\}.$$

Moreover,

$$\lim_{k\to\infty} x_k = x \quad \text{in} \quad X.$$

Proof. Let $k, m \in \mathbb{N}$. Then $\overline{B}(x_m, r_m) \subset \overline{B}(x_k, r_k)$ if $m > k$ and $\overline{B}(x_k, r_k) \subset \overline{B}(x_m, r_m)$ if $m < k$. Hence,

$$d(x_k, x_m) \le \max\{r_k, r_m\}. \tag{5.1}$$

Therefore,

$$\lim_{k,m\to\infty} d(x_k, x_m) = 0.$$

So $\{x_k\}_{k\in\mathbb{N}}$ is a Cauchy sequence. Since X is complete there exists $x \in X$ such that

$$\lim_{m\to\infty} d(x_m, x) = 0.$$

By passing to the limit in (5.1) as $m \to \infty$ (note that for $m > k$, (5.1) takes the form $d(x_k, x_m) \le r_k$), we get that

$$d(x_k, x) \le r_k \quad \forall\, k \in \mathbb{N}.$$

Thus $x_k \in \overline{B}_k\ \forall\, k \in \mathbb{N}$.

If there is also $y \in X$ such that $y \in \overline{B}_k\ \forall\, k \in \mathbb{N}$, then

$$d(x, y) \le d(x, x_k) + d(x_k, y) \le 2r_k.$$

By passing to the limit as $k \to \infty$, we get

$$d(x, y) = 0 \iff y = x.$$

□

Remark. In Lemma 5.1.1 one cannot replace closed balls $\overline{B}_k$ with open balls B_k as the following simple example shows. Let $X = \mathbb{R}$ and $B_k = (0, \frac{1}{k})$, $k \in \mathbb{N}$. Then

$$B_{k+1} \subset B_k, \; k \in \mathbb{N}, \quad \lim_{k\to\infty} r_k = \lim_{k\to\infty} \frac{1}{2k} = 0,$$

but

$$\bigcap_{k=1}^{\infty} B_k = \varnothing.$$

However, one can replace closed balls $\overline{B}_k$ by open balls B_k if the assumption $B_{k+1} \subset B_k$ is replaced by a stronger assumption $\overline{B}_{k+1} \subset B_k$.

Corollary. *Let X be a complete metric space, $B_k = B(x_k, r_k) \subset X$, $k \in \mathbb{N}$, be open balls such that*

$$\overline{B}_{k+1} \subset B_k, \; k \in \mathbb{N}, \quad \lim_{k\to\infty} r_k = 0. \tag{5.2}$$

Then there exists a unique point $x \in X$ such that

$$\bigcap_{k=1}^{\infty} B_k = \{x\}. \tag{5.3}$$

Proof. Condition (5.2) implies condition (5.1). Hence, by Lemma 5.1.1 there exists $x \in X$ such that

$$\bigcap_{k=1}^{\infty} \overline{B}_k = \{x\}.$$

By condition (5.2),

$$\bigcap_{k=1}^{\infty} \overline{B}_k = \bigcap_{k=2}^{\infty} \overline{B}_k = \bigcap_{k=1}^{\infty} \overline{B}_{k+1} \subset \bigcap_{k=1}^{\infty} B_k \subset \bigcap_{k=1}^{\infty} \overline{B}_k.$$

Therefore,

$$\bigcap_{k=1}^{\infty} B_k = \bigcap_{k=1}^{\infty} \overline{B}_k = \{x\}.$$

□

Definition. *A subset M of a metric space X is said to be nowhere dense in X if its closure $\overline{M}$ has no inner points of X.*

Examples. Any finite subset of $\mathbb{R}$, $\{\frac{1}{k}\}_{k\in\mathbb{N}}$ in $\mathbb{R}$, the uncountable Cantor set in $\mathbb{R}$.

Definition. *A metric space X is called a space of the first category if X is a union of countably many nowhere dense subsets of X, that is, $\exists\, M_k \subset X$, $k \in \mathbb{N}$, nowhere dense in X such that*

$$X = \bigcup_{k=1}^{\infty} M_k. \tag{5.4}$$

Otherwise, X is called a space of the second category. This means that if for some $M_k \subset X$, $k \in \mathbb{N}$, equality holds, then at least one of them is such that $\overline{M}_k$ has an inner point of X $\iff$ contains an open ball in X.

Theorem (Baire's Category Theorem). *If a non-empty metric space X is complete, then it is a space of the second category.*

Proof. Assume to the contrary that X is a space of the first category, that is, $\exists\, M_k \subset X,\ k \in \mathbb{N}$, nowhere dense in X such that (5.4) holds.

Then $\overline{M}_1$ does not contain inner points, but X does (X is open and all its points are inner points). Hence, $\overline{M}_1 \neq X$. Therefore,

$$^{\mathrm{c}}\overline{M}_1 = X \setminus \overline{M}_1 \neq \oslash \text{ and is open.}$$

Let $x_1 \in {}^{\mathrm{c}}\overline{M}_1$. Then there exists $0 < r_1 \leq \frac{1}{2}$ such that

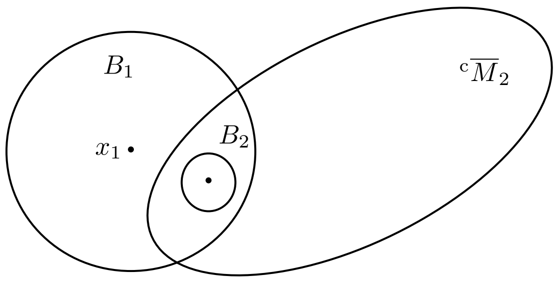
$$\overline{B}_1 \equiv \overline{B}(x_1, r_1) \subset {}^{\mathrm{c}}\overline{M}_1.$$

Next, $\overline{M}_2$ does not contain an open ball. Hence $B_1 \not\subset \overline{M}_2 \iff$ there exists $x_2 \in B_1$ such that $x_2 \notin \overline{M}_2 \iff x_2 \in {}^{\mathrm{c}}\overline{M}_2$. So

$$^{\mathrm{c}}\overline{M}_2 \cap B_1 \neq \oslash \text{ is open} \quad \text{and} \quad x_2 \in {}^{\mathrm{c}}\overline{M}_2 \cap B_1.$$

Therefore, there exists $0 < r_2 \leq \frac{1}{4}$ such that

$$\overline{B}_2 \equiv \overline{B}(x_2, r_2) \subset {}^{\mathrm{c}}\overline{M}_2 \quad \text{and} \quad \overline{B}_2 \subset \overline{B}_1.$$



By induction $\forall\, k \in \mathbb{N},\ k \geq 2$, there exist $x_k \in {}^{\mathrm{c}}\overline{M}_k \cap B_{k-1}$ and $0 < r_k \leq 2^{-k}$ such that

$$\overline{B}_k = \overline{B}(x_k, r_k) \subset {}^{\mathrm{c}}\overline{M}_k \ (\implies \overline{B}_k \cap \overline{M}_k = \oslash),\ \overline{B}_{k+1} \subset \overline{B}_k, \quad \lim_{k\to\infty} r_k = 0.$$

By Lemma 5.1.1, there exists $x \in X$ such that $\forall\, k \in \mathbb{N} \quad x_k \in B_k$ and $x_k \notin M_k$. Hence

$$x \notin \bigcup_{k=1}^{\infty} M_k = X$$

and we arrive at a contradiction. □

5.2 Uniform boundedness theorem

Theorem. *Let $\{T_n\}_{n\in\mathbb{N}}$ be a sequence of bounded linear operators $T_n : X \to Y$ acting from a Banach space X to a normed space Y such that the sequence $\{\|T_n x\|_Y\}_{n\in\mathbb{N}}$ is bounded for every $x \in X$, that is, for some $c_x > 0$*

$$\|T_n x\|_Y \leq c_x \quad \forall\, n \in \mathbb{N}. \tag{5.5}$$

Then the sequence of norms $\{\|T_n\|_{X\to Y}\}_{n\in\mathbb{N}}$ is bounded, that is, there is a $c > 0$ such that

$$\|T_n\|_{X\to Y} \leq c \quad \forall\, n \in \mathbb{N}. \tag{5.6}$$

Proof. For every $n \in \mathbb{N}$, let $M_k \subset X$ be the set of all x for which

$$\|T_n x\|_Y \le k \quad \forall\, n \in \mathbb{N}.$$

The set M_k is closed. Indeed, let $\{x_j\}_{j\in\mathbb{N}} \subset M_k$ be a sequence converging to some $x \in X$. Then for every $n \in \mathbb{N}$

$$\|T_n x_j\|_Y \le k.$$

Hence for every $n \in \mathbb{N}$

$$\|T_n x\|_Y = \left\| T_n\Big(\lim_{j\to\infty} x_j\Big)\right\|_Y = \left\|\lim_{j\to\infty} T_n x_j\right\|_Y = \lim_{j\to\infty} \|T_n x_j\|_Y \le k.$$

So $x \in M_k$, and M_k is closed.

By (5.5), each $x \in X$ belongs to some M_k. Hence

$$X = \bigcup_{k=1}^{\infty} M_k.$$

Since X is complete, Baire's Category Theorem implies that some M_k contains an open ball, say,

$$B_0 \equiv B(x_0, r) \subset M_{k_0}. \tag{5.7}$$

For an arbitrary $x \in X,\ x \ne 0$, we set

$$z = x_0 + \gamma x, \quad \text{where } \gamma = \frac{r}{2\|x\|_X}. \tag{5.8}$$

Then $\|z - x_0\|_X < r \implies x \in B_0$. By (5.7) and by the definition of M_{k_0} we have $\|T_n z\|_Y \le k_0$ for all $n \in \mathbb{N}$. Also $\|T_n x_0\|_Y \le k_0$ for all $n \in \mathbb{N}$ since $x_0 \in B_0$. From (5.8) we obtain

$$x = \frac{1}{\gamma}(z - x_0),$$

hence

$$\|T_n x\|_Y = \frac{1}{\gamma}\|T_n(z-x_0)\|_Y = \frac{1}{\gamma}\|T_n z - T_n x_0\|_Y \le \frac{1}{\gamma}(\|T_n z\|_Y + \|T_n x_0\|_Y) \le \frac{4\|x\|_X k_0}{r}.$$

Therefore

$$\|T_n\|_{X\to Y} = \sup_{\substack{x\in X\\ x\ne 0}} \frac{\|T_n x\|_Y}{\|x\|_X} \le \frac{4k_0}{r} \quad \forall\, n \in \mathbb{N},$$

which is inequality (5.6) with $c = \frac{4k_0}{r}$. □

5.3 Weak Convergence

Definition. *A sequence $\{x_n\}_{n\in\mathbb{N}}$ in a normed space X is said to be weakly convergent if there exists $x \in X$ such that for every $f \in X^*$*

$$\lim_{n\to\infty} f(x_n) = f(x).$$

Briefly,

$$x_n \xrightarrow{w} x \quad \text{as} \quad n \to \infty.$$

The element x is called the weak limit of $\{x_n\}_{n\in\mathbb{N}}$, and we say that the sequence $\{x_n\}_{n\in\mathbb{N}}$ converges weakly in X to x.

First, we prove several statements for weak convergence having the same form as for ordinary convergence, which in the sequel, will be called strong convergence in order to distinguish it from weak convergence.

Lemma. *Let $\{x_n\}_{n\in\mathbb{N}}$ be a weakly convergent sequence in a normed space X, say $x_n \xrightarrow{w} x$ as $n \to \infty$. Then*

1) *the weak limit x of $\{x_n\}_{n\in\mathbb{N}}$ is unique,*
2) *every subsequence of $\{x_n\}_{n\in\mathbb{N}}$ converges weakly to x,*
3) *the sequence $\{\|x_n\|_X\}_{n\in\mathbb{N}}$ is bounded.*

Proof. 1) Suppose also that $x_n \xrightarrow{w} y$ as $n \to \infty$ for some $y \in X$. Then for every $f \in X^*$

$$\lim_{n\to\infty} f(x_n) = f(x) \quad \text{and} \quad \lim_{n\to\infty} f(x_n) = f(y).$$

Since $\{f(x_n)\}_{n\in\mathbb{N}}$ is a numerical sequence, its limit is unique. Hence, $f(x) = f(y)$. So, for every $f \in X'$

$$f(x) - f(y) = f(x-y) = 0.$$

Recall that there exists a functional $g \in X^*$ such that $\|g\|_{X^*} = 1$ and $g(x-y) = \|x-y\|_X$. So,

$$\|x-y\|_X = g(x-y) = 0 \implies y = x.$$

2) Let $\{x_{n_k}\}_{k\in\mathbb{N}}$ be a subsequence of $\{x_n\}_{n\in\mathbb{N}}$ and $f \in X^*$. Since $\{f(x_n)\}_{n\in\mathbb{N}}$ is a numerical sequence, for any $f \in X^*$

$$\lim_{k\to\infty} f(x_{n_k}) = \lim_{n\to\infty} f(x_n) = f(x)$$

which implies that $x_{n_k} \xrightarrow{w} x$ as $k \to \infty$.

3) Let $f \in X^*$. Since $\{f(x_n)\}_{n\in\mathbb{N}}$ is a convergent numerical sequence, it is bounded, that is, for some $c_f > 0$

$$|f(x_n)| \le c_f \quad \forall\, n \in N.$$

Define $g_n : X^* \to \mathbb{C}$ (or $\mathbb{R}$) by

$$g_n(f) = f(x_n) \quad \forall\, f \in X^*,\ n \in \mathbb{N}.$$

Then $\{g_n\}_{n\in\mathbb{N}}$ is a sequence of bounded linear operators (functionals) with $\|g_n\|_{X^*\to\mathbb{C}} = \|x_n\|_X$ (duality formula) such that the sequence $\{\|g_n(f)\|_{\mathbb{C}}\}_{n\in\mathbb{N}}$ is bounded for every $f \in X^*$, because

$$\|g_n(f)\|_{\mathbb{C}} = |g_n(f)| = |f(x_n)| \le c_f \quad \forall\, n \in \mathbb{N}.$$

Since X^* is complete, the Uniform Boundedness Theorem implies that the sequence $\{\|g_n\|_{X^*\to\mathbb{C}}\}_{n\in\mathbb{N}} = \{\|x_n\|_X\}_{n\in\mathbb{N}}$ is bounded.

□

Theorem. *Let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence in a normed space X. Then*

1) *strong convergence implies weak convergence with the same limit,*
2) *the converse of 1) is not generally true, in particular, any orthonormal sequence $\{e_n\}_{n\in\mathbb{N}}$ in a Hilbert space H weakly converges in H to 0, but does not converge strongly in H,*
3) *if $\dim X < \infty$, then weak convergence in X is equivalent to strong convergence in X.*

Proof. 1) Assume that $x \in X$ and $x_n \to x$ as $n \to \infty$ in X, that is, $\lim_{n\to\infty} \|x_n - x\|_X = 0$. Since $\forall\, f \in X^*$

$$|f(x_n) - f(x)| = |f(x_n - x)| \le \|f\|_{X^*}\|x_n - x\|_X,$$

it follows that $\lim_{n\to\infty} f(x_n) = f(x)$, hence $x_n \xrightarrow{w} x$ as $n \to \infty$.

2) Let $\{e_n\}_{n\in\mathbb{N}}$ be an orthonormal sequence in a Hilbert space H and $f \in H^*$. By the Riesz representation theorem there exists $z \in H$ such that $f(x) = (x, z)_H$ $\forall\, x \in H$, in particular, $f(e_n) = (e_n, z)_H$ $\forall\, n \in \mathbb{N}$. Recall that

$$\lim_{n\to\infty} f(e_n) = \lim_{n\to\infty} (e_n, z)_H = 0.$$

(This is a corollary of Bessel's inequality.) Since $f \in H^*$ was arbitrary, we see that $e_n \xrightarrow{w} 0$. However, $\{e_n\}_{n\in\mathbb{N}}$ does not converge strongly, because for any $m, n \in \mathbb{N},\ m \neq n$

$$\|e_m - e_n\|^2 = (e_m - e_n, e_m - e_n) = (e_m, e_m) - 2\,\mathrm{Re}(e_m, e_n) + (e_n, e_n) = 2.$$

3) By 1) it suffices to prove that weak convergence implies strong convergence. Suppose that $\dim X = k,\ k \in \mathbb{N}$, and $\{x_n\}_{n\in\mathbb{N}} \subset X,\ x \in X,\ x_n \xrightarrow{w} x$ as $n \to \infty$. Let $\{e_1, \ldots, e_k\}$ be any basis for X, then for some $\alpha_j^{(n)} \in \mathbb{C},\ \alpha_j,\ j = 1, \ldots, k$

$$x_n = \alpha_1^{(n)} e_1 + \ldots + \alpha_k^{(n)} e_k$$

and

$$x = \alpha_1 e_1 + \ldots + \alpha_k e_k.$$

By assumption, $f(x_n) \to f(x)$ as $n \to \infty$ for every $f \in X^*$. We take, in particular, the linear functionals $f_j,\ j = 1, \ldots, k$, defined by

$$f_j(e_j) = 1, \quad f_j(e_m) = 0, \quad m = 1, \ldots, j-1, j+1, \ldots, k.$$

Then $f \in X^*$ and

$$f_j(x_n) = \alpha_j^{(n)}, \quad f_j(x) = \alpha_j, \quad j = 1, \ldots, k.$$

Since $f_j(x_n) \to f_j(x)$ as $n \to \infty$, we have $\alpha_j^{(n)} \to \alpha_j$ as $n \to \infty,\ j = 1, \ldots, k$. Therefore

$$\|x_n - x\|_X = \Big\| \sum_{j=1}^{k} (\alpha_j^{(n)} - \alpha_j) e_j \Big\|_X \le \sum_{j=1}^{k} |\alpha_j^{(n)} - \alpha_j| \cdot \|e_j\|_X \to 0$$

as $n \to \infty$, which shows that $\{x_n\}_{n\in\mathbb{N}}$ converges strongly to x.

□

Exercise. Prove that the functionals f_j are bounded.

Remark. It is interesting to note that there also exist infinite dimensional normed spaces such that strong and weak convergence are equivalent. An example is the space ℓ_1 (without proof).

5.4 Open mapping theorem

Bounded inverse theorem

Definition. *Let X and Y be metric spaces. Then $T : X \to Y$ is called an open mapping if for every open set in X its image is an open set in Y.*

Example. Let $T_k : \mathbb{R} \to \mathbb{R}$, where $k \in \mathbb{N}$ and $T_k(x) = x^k,\ x \in \mathbb{R}$. If k is odd, then T_k is an open set. Indeed, let $M \subset \mathbb{R}$ be an arbitrary open set, then $M = \bigcup\limits_{m=1}^{s} (a_m, b_m)$, where $s \in \mathbb{N}$ or $s = \infty$ and (a_m, b_m) are non-intersecting intervals (finite or infinite). Then

$$T_k(M) = T_k\Big(\bigcup_{m=1}^{s} (a_m, b_m)\Big) = \bigcup_{m=1}^{s} T_k((a_m, b_m)) = \bigcup_{m=1}^{s} (a_m^k, b_m^k)$$

is an open set. If k is even, then T_k is not an open mapping, because, for example, $T_k((-1, 1)) = [0, 1)$ is not an open set.

Example. Recall that a mapping $T : X \to Y$ is continuous if and only if for any open set in Y its preimage is open. Assume that T is bijective (one-to-one + onto $\iff T(X) = Y$). Then there exists the inverse $T^{-1} : Y \to X$, and the continuity of T is equivalent to the openness of T^{-1}.

Lemma. *Let X and Y be metric spaces and $T : X \to Y$. Then the following statements are equivalent:*

a) *T is open,*

b) *for any ball $B_X(x, r) \subset X \quad T(B_X(x, r))$ is open in Y,*

c) *for any ball $B_X(x, r) \subset X$ there exists $\rho > 0$ such that $T(B_X(x, r)) \supset B_Y(Tx, \rho)$.*

Proof. Clearly a) $\implies$ b). Also b) $\implies$ c), because $Tx \in T(B_X(x, r))$ and $T(B_X(x, r))$ is open, hence there exists $\rho > 0$ such that $B_Y(Tx, \rho) \subset T(B_X(x, r))$.

$\boxed{\text{c)} \implies \text{a).}}$ Let $A \subset X$ be an open set, then $\forall\, x \in A\ \exists\, r_x > 0$ such that $B_X(x, r_x) \subset A$. Hence

$$A = \bigcup_{x \in A} B_X(x, r_x).$$

Next

$$T(A) = T\Big(\bigcup_{x \in A} B_X(x, r_x)\Big) = \bigcup_{x \in A} T(B_X(x, r_x)).$$

By c) $\exists\, \rho_x : T(B_X(x, r_x)) \supset B_Y(Tx, \rho_x)$. Hence

$$\bigcup_{x \in A} B_Y(Tx, \rho_x) = T(A). \tag{5.9}$$

Indeed,

1) $B_Y(Tx, \rho_x) \subset T(B_X(x, r_x)) \subset T(A)$, so $\bigcup\limits_{x \in A} B_Y(Tx, \rho_x) \subset T(A)$,

2) if $y \in T(A)$ then $\exists\, x \in A : y = Tx$, hence $y \in B_Y(Tx, \rho_x)$ and $y \in \bigcup\limits_{z \in A} B_Y(Tz, \rho_z)$, so $T(A) \subset \bigcup\limits_{z \in A} B_Y(Tz, \rho_z)$.

Equality (5.9) implies that $T(A)$ is open. □

Example. Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function which is strictly increasing or strictly decreasing. Then f is an open mapping. This follows by the above lemma, because in the first case $\forall\, x \in \mathbb{R}$ and $\forall\, r > 0$

$$f((x-r, x+r)) = (f(x-r), f(x+r))$$

and in the second case

$$f((x-r, x+r)) = (f(x+r), f(x-r)).$$

Lemma. *Let X and Y be normed spaces and $T : X \to Y$ be a linear operator. Then the equivalent conditions* a), b), c) *of the previous lemma are also equivalent to the condition*

d) *there exists $\rho > 0$ such that $T(B_X(0,1)) \supset B_Y(0,\rho)$.*

Proof. Clearly c) $\implies$ d). Let us prove that d) $\implies$ c). Let $x \in X$ and $r > 0$, then

$$T(B_X(x,r)) = Tx + rT(B_X(0,1)). \tag{5.10}$$

Indeed,

$$\begin{aligned} T(B_X(x,r)) &= T(\{y : \|x-y\|_X < r\}) = \{Ty : \|x-y\|_X < r\} \\ &= \{Tx + T(y-x) : \|x-y\|_X < r\} = (x-y=z) = \{Tx + Tz : \|z\|_X < r\} \\ &= \Big\{Tx + rT\Big(\frac{z}{r}\Big) : \Big\|\frac{z}{r}\Big\|_X < 1\Big\} = \Big(\frac{z}{r} = w\Big) = \{Tx + rTw : \|w\|_X < 1\} \\ &= Tx + r\{Tw : \|w\|_X < 1\} = Tx + rT\{w : \|w\|_X < 1\} = Tx + rT(B_X(0,1)). \end{aligned}$$

By applying equality (5.10) and condition d) we get

$$T(B_X(x,r)) = Tx + rT(B_X(0,1)) \supset Tx + rB_Y(0,\rho) = B_Y(Tx, r\rho).$$

□

Lemma. *Let X and Y be Banach spaces and $T : X \to Y$ be a bounded linear operator. Then the equivalent conditions* a), b), c), d) *of the previous lemmas are equivalent to the condition*

e) *there exists $\rho > 0$ such that $\overline{T(B_X(0,1))} \supset B_Y(0,\rho)$.*

Proof. Clearly d) $\implies$ e). We shall prove that

$$T(B_X(0,1)) \supset B_Y\Big(0, \frac{\rho}{2}\Big). \tag{5.11}$$

To do this we first prove that the condition e) implies that $\forall\, n \in \mathbb{N}$

$$\overline{T(B_X(0,2^{-n}))} \supset B_Y(0, \rho 2^{-n}). \tag{5.12}$$

Indeed, by (5.10) with $x = 0$ and $r = 2^{-n}$

$$\begin{aligned} \overline{T(B_X(0,2^{-n}))} &= \overline{T(2^{-n}B_X(0,1))} = 2^{-n}\overline{T(B_X(0,1))} \\ &\supset 2^{-n}B_Y(0,\rho) = B_Y(0, 2^{-n}\rho). \end{aligned}$$

Let $y \in B_Y(0, \frac{\rho}{2})$. By (5.12) $y \in \overline{T(B_X(0,\frac{1}{2}))}$, hence $\exists\, v \in T(B_X(0,\frac{1}{2}))$ such that $\|y - v\| < \frac{\rho}{4}$ and $\exists\, x_1 \in B_X(0, \frac{1}{2})$ such that $v = Tx_1$. So

$$\|y - Tx_1\|_Y < \frac{\rho}{4}$$

which implies that $y - Tx_1 \in B_X(0, \frac{1}{4})$. As before we conclude that there exists $x_2 \in B_X(0, \frac{1}{4})$ such that

$$\|(y - Tx_1) - Tx_2\|_Y < \frac{\rho}{8}.$$

So $y - (Tx_1 + Tx_2) \in B_X(0, \frac{1}{8})$, and so on.

Therefore $\forall\, k \in \mathbb{N}\ \exists\, x_k \in B_X(0, 2^{-k})$ such that

$$\Big\| y - T\Big(\sum_{k=1}^{n} x_k\Big)\Big\|_Y < \rho 2^{-n-1}. \tag{5.13}$$

Let $z_n = x_1 + \ldots + x_n,\ n \in \mathbb{N}$. Since $x_k \in B_X(0, 2^{-k}),\ \|x_k\|_X < 2^{-k}$. Therefore $\forall\, m, n \in \mathbb{N},\ n > m$

$$\|z_n - z_m\|_X \le \sum_{k=m+1}^{n} \|x_k\|_X < \sum_{k=m+1}^{\infty} 2^{-k} = 2^{-m}.$$

Similarly, for $m > n$

$$\|z_m - z_n\|_X < 2^{-n}.$$

Therefore

$$\lim_{m,n\to\infty} \|z_m - z_n\|_X = 0.$$

Since the space X is complete, $\exists\, x \in X$ such that $\lim_{m\to\infty} z_m = x$ in X. Also

$$\|x\|_X = \Big\|\sum_{k=1}^{\infty} x_k\Big\|_X \le \sum_{k=1}^{\infty} \|x_k\|_X < \sum_{k=1}^{\infty} 2^{-k} = 1.$$

Hence $x \in B_X(0,1)$. By the continuity of T and (5.13)

$$\|y - Tx\|_Y = \lim_{n\to\infty} \Big\| y - T\Big(\sum_{k=1}^{m} x_k\Big)\Big\|_Y = 0.$$

So $y = Tx$ and $y \in T(B_X(0,1))$, which implies (5.11). $\square$

Theorem (Open mapping theorem). *Let X and Y be Banach spaces. A bounded linear operator $T : X \to Y$ which maps X onto Y $(T(X) = Y)$ is an open mapping.*

Proof. $\boxed{\text{Step 1.}}$ ($\overline{T(B_X(0,1))}$ *contains an open ball.*) Clearly

$$X = \bigcup_{k=1}^{\infty} B_X(0,k),$$

hence

$$Y = T(X) = T\Big(\bigcup_{k=1}^{\infty} B_X(0,k)\Big) = \bigcup_{k=1}^{\infty} T(B_X(0,k)).$$

By Baire's Category Theorem $\exists\, k_0 \in \mathbb{N}$ such that $\overline{T(B_X(0,k))}$ contains an open ball in Y, say, $B_Y(z_0, r_0)$. Then

$$\overline{T(B_X(0,1))} = \frac{1}{k_0}\overline{T(B_X(0,k_0))} \supset \frac{1}{k_0} B_Y(z_0, r_0) = B_Y\Big(\frac{z_0}{k_0}, \frac{r_0}{k_0}\Big). \tag{5.14}$$

$\boxed{\text{Step 2.}}$ ($\overline{T(B_X(0,1))}$ *contains an open ball in* Y *centered at* 0.) By (5.14) it follows that

$$\overline{T(B_X(0,1/2))} = \frac{1}{2}\overline{T(B_X(0,1))} \supset B_Y(w_0, \rho_0) = B_Y(0, \rho_0) + w_0, \tag{5.15}$$

where $w_0 = \frac{z_0}{2k_0}$, $\rho_0 = \frac{r_0}{2k_0}$. Let us prove that

$$\overline{T(B_X(0,1))} \supset \overline{T(B_X(0,1/2))} - w_0. \tag{5.16}$$

Let $y \in \overline{T(B_X(0,1/2))} - w_0$, and $w_0 \in \overline{T(B_X(0,1/2))}$. Then there exist $u_n, v_n \in B_X(0,1/2)$, $n \in \mathbb{N}$, such that $Tu_n, Tv_n \in T(B_X(0,1/2))$ and $u_n \to y + w_0$, $v_n \to w_0$ as $n \to \infty$. Then

$$\|Tu_n - Tv_n\|_Y \le \|Tu_n\|_Y + \|Tv_n\|_Y < \frac{1}{2} + \frac{1}{2} = 1,$$

hence $T(u_n - v_n) = Tu_n - Tv_n \in B_X(0,1)$ and

$$T(u_n - v_n) = Tu_n - Tv_n \to y + w_0 - w_0 = y.$$

So, $y \in \overline{T(B_X(0,1))}$ which implies (5.16).

Finally, (5.15) and (5.16) imply that

$$\overline{T(B_X(0,1))} \supset B_Y(0, \rho_0),$$

and the desired statement follows by the above lemma. □

Theorem (Bounded Inverse Theorem). *Let X and Y be Banach spaces and $T : X \to Y$ be a bijective bounded linear operator. Then T^{-1} is bounded.*

Proof. By the Open Mapping Theorem T is open. The continuity of T^{-1} is equivalent to the openness of $(T^{-1})^{-1} = T$. So T^{-1} is continuous $\iff$ bounded. □

5.5 Closed Graph Theorem

5.5.1 Closed operators

Let X and Y be normed spaces. Consider the normed space $X \times Y$, where the two algebraic operations of a vector space in $X \times Y$ are defined as usual, that is,

$$\begin{aligned} (x_1, y_1) + (x_2, y_2) &= (x_1 + x_2, y_1 + y_2) \\ \alpha(x, y) &= (\alpha x, \alpha y) \end{aligned}$$

(α is a scalar) and the norm on $X \times Y$ is defined by

$$\|(x, y)\|_{X \times Y} = \|x\|_X + \|y\|_Y.$$

Note that

$$\lim_{n\to\infty} (x_n, y_n) = (x, y) \quad \text{in} \quad X \times Y \iff \lim_{n\to\infty} x_n = x \quad \text{in} \quad X, \quad \lim_{n\to\infty} y_n = y \quad \text{in} \quad Y.$$

Lemma 5.5.1. *If X and Y are Banach spaces, then $X \times Y$ is also a Banach space.*

Proof. Let $\{z_n\}_{n\in\mathbb{N}}$ be a Cauchy sequence in $X \times Y$, where $z_n = (x_n, y_n),\ x_n \in X,\ y_n \in Y$. Then

$$\lim_{m,n\to\infty} \|z_m - z_n\|_{X\times Y} = \lim_{m,n\to\infty} (\|x_m - x_n\|_X + \|y_m - y_n\|_Y) = 0.$$

Hence

$$\lim_{m,n\to\infty} \|x_m - x_n\|_X = \lim_{m,n\to\infty} \|y_m - y_n\|_Y = 0,$$

which means that $\{x_n\}_{n\in\mathbb{N}}$ and $\{y_n\}_{n\in\mathbb{N}}$ are Cauchy sequences in $X,\ Y$ respectively. Since X and Y are complete spaces, there exist $x \in X$ and $y \in Y$ such that $\lim\limits_{n\to\infty} x_n = x$ in X and $\lim\limits_{n\to\infty} y_n = y$ in Y. This implies that

$$\lim_{n\to\infty} z_n = \lim_{n\to\infty} (x_n, y_n) = (x, y).$$

Since the Cauchy sequence $\{z_n\}_{n\in\mathbb{N}}$ was arbitrary, $X \times Y$ is complete. □

Definition 5.5.1 (Closed operator). *Let X and Y be normed spaces and $T : \mathfrak{D}(T) \to Y$ an operator with domain $\mathfrak{D}(T) \subset X$. Then T is called a closed operator if its graph*

$$\mathfrak{G}(T) = \{(x, y) | x \in \mathfrak{D}(T),\ y = Tx\} \quad \text{is closed in} \quad X \times Y. \tag{5.17}$$

By the definition of a closed set this means that if $z_n \in \mathfrak{G}(T),\ n \in \mathbb{N}$ and $z_n \to z$ as $n \to \infty$ in $X \times Y$, then $z \in \mathfrak{G}(T)$. In other words, if $x_n \in \mathfrak{D}(T)$ and

$$z_n = (x_n, Tx_n) \to z = (x, y) \text{ in } X \times Y \text{ as } n \to \infty \iff x_n \to x \text{ in } X \quad \text{and} \quad y_n \to y \text{ in } Y,$$

then $z \in \mathfrak{G}(T) \iff x \in \mathfrak{D}(T)$ and $y = Tx$.

Remark 5.5.1. Condition (5.17) in the definition of a closed operator can be replaced by

$$\begin{cases} x_n \in \mathfrak{D}(T),\ n \in \mathbb{N},\ x \in X,\ y \in Y \\ x_n \to x \text{ in } X,\ Tx_n \to y \text{ in } Y \text{ as } n \to \infty \end{cases} \implies x_n \in \mathfrak{D}(T),\ Tx = y \tag{5.18}$$

(equivalent definition in terms of sequences).

Importantly, even if $\mathfrak{D}(T)$ is a normed space, $x_n \to x$ in X, not in $\mathfrak{D}(T)$.

Remark 5.5.2 (relationship between continuity and closedness). Let $T : \mathfrak{D}(T) \to Y$ be a continuous operator.

If $\mathfrak{D}(T)$ is closed, then the operator T is closed. Indeed, let the assumptions of (5.18) be satisfied. Since $\mathfrak{D}(T)$ is closed x belonging to X, in fact, belongs to $\mathfrak{D}(T)$ and by the continuity of T

$$Tx = T\Big(\lim_{n\to\infty} x_n\Big) = \lim_{n\to\infty} Tx_n = y.$$

If $\mathfrak{D}(T)$ is not closed, then the operator T is not necessarily closed. Let, for example, $T = I$, the identity operator. Since $\mathfrak{D}(T)$ is not closed there exists $x \in X \setminus \mathfrak{D}(T)$ and $\{x_n\}_{n\in\mathbb{N}} \subset \mathfrak{D}(T)$ such that $x_n \to x$ as $n \to \infty$ in X. Then $Ix_n = x_n \to x$ as $n \to \infty$ and $Ix = x$, but $x \notin \mathfrak{D}(T)$.

Remark 5.5.3. Let $-\infty < a < b < \infty$. The natural domain $\mathfrak{D}(\frac{d}{dt})$ of the differential operator $\frac{d}{dt}$ is the set of all real-valued functions x on (a, b) for which for any $t \in (a, b)$ there exists the derivative $x'(t)$.

We shall consider the action of $\frac{d}{dt}$ from one normed space X of functions defined on (a, b) to another one Y. To do this we consider its domain $\mathfrak{D}_{X,Y}(\frac{d}{dt})$ generated by the spaces X and Y:

$$\mathfrak{D}_{X,Y}\Big(\frac{d}{dt}\Big) = \Big\{x \in \mathfrak{D}\Big(\frac{d}{dt}\Big) : x \in X,\ x' \in Y\Big\}. \tag{5.19}$$

Then $\frac{d}{dt} : \mathfrak{D}_{X,Y}(\frac{d}{dt}) \to Y$ and

$$\Big\|\frac{d}{dt}\Big\|_{\mathfrak{D}_{X,Y}(\frac{d}{dt})\to Y} = \sup_{\substack{x\in\mathfrak{D}_{X,Y}(\frac{d}{dt})\\ x\neq 0}} \frac{\|x'\|_Y}{\|x\|_X}. \tag{5.19$'$}$$

Example 5.1. Let $-\infty < a < b < \infty,\ X = C^1(a,b)$ and $Y = C(a,b)$. Recall that

$$\|x\|_{C(a,b)} = \sup_{t\in(a,b)} |x(t)|, \quad \|x\|_{C^1(a,b)} = \|x\|_{C(a,b)} + \|x'\|_{C(a,b)}. \tag{5.20}$$

Then $\mathfrak{D}_{X,Y}(\frac{d}{dt}) = C^1(a,b)$ and by (5.19$'$)

$$\Big\|\frac{d}{dt}\Big\|_{C^1(a,b)\to C(a,b)} = \sup_{\substack{x\in C^1(a,b)\\ x\neq 0}} \frac{\|x'\|_{C(a,b)}}{\|x\|_{C^1(a,b)}}. \tag{5.20$'$}$$

Then the operator

$$\frac{d}{dt} : C^1(a,b) \to C(a,b) \tag{5.21}$$

is bounded and closed. Indeed, by (5.20$'$) $\|\frac{d}{dt}\|_{C^1(a,b)\to C(a,b)} \le 1$, hence operator (5.21) is bounded. It is also closed because if a sequence $\{x_n\}_{n\in\mathbb{N}} \subset C^1(a,b)$ and $x \in C^1(a,b),\ y \in C(a,b)$ are such that

$$\lim_{n\to\infty} x_n = x \text{ in } C^1(a,b) \iff \lim_{n\to\infty} x_n = x \text{ in } C(a,b) \text{ and } \lim_{n\to\infty} x_n' = x' \text{ in } C(a,b)$$

and

$$\lim_{n\to\infty} x_n' = y \quad \text{in} \quad C(a,b),$$

then $x \in C^1(a,b)$ and, by the uniqueness of the limit, $y = x'$.

Example 5.2. Let $-\infty < a < b < \infty$ and $X = Y = C(a,b)$. Then by (5.19)

$$\mathfrak{D}_{X,Y}\Big(\frac{d}{dt}\Big) = \Big\{x \in \mathfrak{D}\Big(\frac{d}{dt}\Big) : x \in C(a,b),\ x' \in C(a,b)\Big\} = C^1(a,b) \tag{5.22}$$

and by (5.19$'$)

$$\Big\|\frac{d}{dt}\Big\|_{C^1(a,b)\to C(a,b)} = \sup_{\substack{x\in C^1(a,b)\\ x\neq 0}} \frac{\|x'\|_{C(a,b)}}{\|x\|_{C(a,b)}}. \tag{5.22$'$}$$

(Compared with (5.20$'$) we have $\|x\|_{C(a,b)}$ and not $\|x\|_{C^1(a,b)}$.)

Then operator (5.21) is unbounded and closed. Indeed, by (5.22$'$)

$$\Big\|\frac{d}{dt}\Big\|_{C^1(a,b)\to C(a,b)} \ge \lim_{n\to\infty} \frac{\|(e^{nt})'\|_{C(a,b)}}{\|e^{nt}\|_{C(a,b)}} = \lim_{n\to\infty} n = \infty,$$

hence, operator (5.21) is unbounded. It is a closed operator because if $\{x_n\}_{n\in\mathbb{N}} \subset C^1(a,b),\ x,y \in C(a,b),\ x_n \to x$ and $x_n' \to y$ as $n \to \infty$, then, by the appropriate theorem in Calculus, $x \in C^1(a,b)$ and $x' = y$.

Example 5.3. Let $-\infty < a < b < \infty$ and $X = Y = L_1(a,b)$. Then by (5.19)

$$\mathfrak{D}_1\Big(\frac{d}{dt}\Big) \equiv \mathfrak{D}_{X,Y}\Big(\frac{d}{dt}\Big) = \Big\{x \in \mathfrak{D}\Big(\frac{d}{dt}\Big) : x \in L_1(a,b),\ x' \in L_1(a,b)\Big\} \tag{5.23}$$

and by (5.19′)

$$\Big\|\frac{d}{dt}\Big\|_{\mathfrak{D}_1(\frac{d}{dt})\to L_1(a,b)} = \sup_{\substack{x\in\mathfrak{D}_1(\frac{d}{dt})\\ x\not\equiv 0}} \frac{\|x'\|_{L_1(a,b)}}{\|x\|_{L_1(a,b)}}. \tag{5.23′}$$

Then the operator

$$\frac{d}{dt} : \mathfrak{D}_1\Big(\frac{d}{dt}\Big) \to L_1(a,b) \tag{5.24}$$

is unbounded and unclosed. Indeed, by (5.23′)

$$\Big\|\frac{d}{dt}\Big\|_{\mathfrak{D}_1(\frac{d}{dt})\to L_1(a,b)} \ge \lim_{n\to\infty} \frac{\|(e^{nt})'\|_{L_1(a,b)}}{\|e^{nt}\|_{L_1(a,b)}} = \lim_{n\to\infty} n = \infty,$$

hence, operator (5.24) is unbounded. In order to prove that it is not closed we choose

$$x_n(t) = \sqrt{\Big(t - \frac{a+b}{2}\Big)^2 + \frac{1}{n}},\ n \in \mathbb{N}, \quad x(t) = \Big|t - \frac{a+b}{2}\Big|, \quad y(t) = \operatorname{sgn}\Big(t - \frac{a+b}{2}\Big).$$

Clearly $x \notin \mathfrak{D}_1(\frac{d}{dt})$. However,

$$\lim_{n\to\infty} x_n = x, \quad \lim_{n\to\infty} x_n' = y \quad \text{in} \quad L_1(a,b). \tag{5.25}$$

Indeed,

$$\|x_n - x\|_{L_1(a,b)} = \int_a^b \Big|\sqrt{\Big(t - \frac{a+b}{2}\Big)^2 + \frac{1}{n}} - \Big|t - \frac{a+b}{2}\Big|\Big|\,dt = \Big(t - \frac{a+b}{2} = \tau\Big)$$

$$= \int_{\frac{a-b}{2}}^{\frac{b-a}{2}} \Big|\sqrt{\tau^2 + \frac{1}{n}} - |\tau|\Big|\,d\tau = 2\int_0^{\frac{b-a}{2}} \Big(\sqrt{\tau^2 + \frac{1}{n}} - \tau\Big)\,d\tau = 2\int_0^{\frac{b-a}{2}} \frac{\frac{1}{n}}{\sqrt{\tau^2 + \frac{1}{n}} + \tau}\,d\tau.$$

Since $\sqrt{\tau^2 + \frac{1}{n}} + \tau \ge \frac{1}{\sqrt{n}}$, we get

$$\|x_n - x\|_{L_1(a,b)} \le \frac{b-a}{\sqrt{n}},$$

hence, follows the first equality in (5.25). Moreover,

$$\|x_n' - y\|_{L_1(a,b)} = \int_a^b \Bigg|\frac{t - \frac{a+b}{2}}{\sqrt{\Big(t - \frac{a+b}{2}\Big)^2 + \frac{1}{n}}} - \operatorname{sgn}\Big(t - \frac{a+b}{2}\Big)\Bigg|\,dt = \Big(t - \frac{a+b}{2} = \tau\Big)$$

$$= \int_{\frac{a-b}{2}}^{\frac{b-a}{2}} \Big|\frac{\tau}{\sqrt{\tau^2 + \frac{1}{n}}} - \operatorname{sgn}\tau\Big|\,d\tau = 2\int_0^{\frac{b-a}{2}} \Big(1 - \frac{\tau}{\sqrt{\tau^2 + \frac{1}{n}}}\Big)\,d\tau$$

$$= 2\int_0^{\frac{b-a}{2}} \frac{\sqrt{\tau^2 + \frac{1}{n}} - \tau}{\sqrt{\tau^2 + \frac{1}{n}}}\,d\tau = 2\int_0^{\frac{b-a}{2}} \frac{\frac{1}{n}}{\sqrt{\tau^2 + \frac{1}{n}}\big(\sqrt{\tau^2 + \frac{1}{n}} + \tau\big)}\,d\tau.$$

Since $\sqrt{\tau^2+\frac{1}{n}}(\sqrt{\tau^2+\frac{1}{n}}+\tau) \geq \frac{1}{\sqrt{n}}(\frac{1}{\sqrt{n}}+\tau) \geq \frac{1}{\sqrt{n}}(2\frac{1}{\sqrt[4]{n}}\sqrt{\tau})$, we get

$$\|x_n' - y\|_{L_1(a,b)} \leq \frac{1}{\sqrt[4]{n}} \int_0^{\frac{b-a}{2}} \frac{d\tau}{\sqrt{\tau}} = \frac{\sqrt{2(b-a)}}{\sqrt[4]{n}},$$

hence, follows the second equality in (5.25).

Since (5.25) holds, but $x \notin \mathfrak{D}_1(\frac{d}{dt})$, it follows that operator (5.24) is not closed.

5.5.2 Closable operators

If an operator $T : \mathfrak{D}(T) \to Y$ is not closed, can it be slightly altered to become a closed operator? With this aim we introduce the following definition.

Definition 5.5.2. *Let X, Y be normed spaces, $\mathfrak{D}(T) \subset X$ and $T : \mathfrak{D}(T) \to Y$. The operator T is called closable, if there exists a closed extension of T, that is, a closed operator $\bar{T} : \mathfrak{D}(\bar{T}) \to Y$ such that*

$$\mathfrak{D}(T) \subset \mathfrak{D}(\bar{T}) \quad \text{and} \quad \bar{T}x = Tx \quad \forall\, x \in \mathfrak{D}(\bar{T}). \tag{5.26}$$

Lemma 5.5.2. *Let X, Y be normed spaces, $\mathfrak{D}(T) \subset X$ and $T : \mathfrak{D}(T) \to Y$ be a linear operator. The operator T is closable if and only if the condition*

$$y \in Y,\ x_k \in \mathfrak{D}(T),\ k \in \mathbb{N}, \quad x_k \to 0 \quad \text{in} \quad X \quad \text{and} \quad Tx_k \to y \quad \text{in} \quad Y \tag{5.27}$$

implies that $y = 0$.

Proof. $\boxed{\text{Step 1.}}$ Let $\bar{T} : \mathfrak{D}(\bar{T}) \to Y$ be a closed extension of T. Assume that condition (5.27) is satisfied. Then by (5.26)

$$y \in Y,\ x_k \in \mathfrak{D}(\bar{T}),\ k \in \mathbb{N}, \quad x_k \to 0 \quad \text{in} \quad X, \quad \bar{T}x_k = Tx_k \to y \quad \text{in} \quad Y.$$

Since T is closed, $0 \in \mathfrak{D}(\bar{T})$ and $\bar{T}(0) = y$. Since $0 \in \mathfrak{D}(T)$, $T(0) = \bar{T}(0) = y$ and $y = T(0) = 0$.

$\boxed{\text{Step 2.}}$ Let condition (5.27) be satisfied. Define

$$\mathfrak{D}(\bar{T}) = \{x \in X : \text{there exist } \{x_k\}_{k\in\mathbb{N}} \subset X \text{ and } y \in Y \text{ such that } x_k \to x \text{ in } X,\ Tx_k \to y \text{ in } Y \text{ as } k \to \infty\}$$

and let

$$\bar{T}x = y \quad \forall\, x \in \mathfrak{D}(\bar{T}).$$

$\boxed{\text{Step 3.}}$ Condition (5.27) ensures that the operator $\bar{T}$ is well-defined. Indeed, let $\{\widetilde{x_k}\}_{k\in\mathbb{N}} \subset X$, $\widetilde{y} \in Y$ also be such that $\widetilde{x_k} \to x$ in X, $T\widetilde{x_k} \to y$ in Y as $k \to \infty$. Then

$$x_k - \widetilde{x_k} \to 0, \quad T(x_k - \widetilde{x_k}) = Tx_k - T\widetilde{x_k} \to y - \widetilde{y} \quad \text{as} \quad k \to \infty.$$

By (5.27) $y - \widetilde{y} = 0 \iff \widetilde{y} = y$.

$\boxed{\text{Step 4.}}$ *$\bar{T}$ is an extension of T.* Let $x \in \mathfrak{D}(T)$. Consider $x_k = x$, $k \in \mathbb{N}$. Then $x_k \to x$ in X, $Tx_k \to Tx$ in Y as $k \to \infty$. Hence, by the definition of $\bar{T}$, $x \in \mathfrak{D}(\bar{T})$ and $\bar{T}x = Tx$.

$\boxed{\text{Step 5.}}$ *$\bar{T}$ is a closed operator.* Let $x_k \in \mathfrak{D}(\bar{T})$, $k \in \mathbb{N}$, $x \in X$, $y \in Y$ and $x_k \to x$ in X, $\bar{T}x_k \to y$ in Y as $k \to \infty$. Since $x_k \in \mathfrak{D}(\bar{T})$ there exist $x_{km} \in \mathfrak{D}(T)$, $m \in \mathbb{N}$, and $y_k \in Y$

such that $x_{km} \to x_k$ in X, $Tx_{km} \to y_k$ in Y as $m \to \infty$ and $\bar{T}x_k = y_k$. Hence, for any $k \in \mathbb{N}$ there exists $m(k) \in \mathbb{N}$ such that

$$\|x_{km(k)} - x_k\|_X < \frac{1}{k}, \quad \|Tx_{km(k)} - y_k\|_Y < \frac{1}{k}.$$

Let $\widetilde{x_k} = x_{km(k)}$. Then

$$\begin{aligned}\|\widetilde{x_k} - x\|_X &= \|\widetilde{x_k} - x_k + x_k - x\|_X \\ &\leq \|\widetilde{x_k} - x_k\|_X + \|x_k - x\|_X < \frac{1}{k} + \|x_k - x\|_X.\end{aligned}$$

Hence, $\widetilde{x_k} \to x$ as $k \to \infty$. Moreover,

$$\begin{aligned}\|T\widetilde{x_k} - y\|_Y &= \|T\widetilde{x_k} - y_k + \bar{T}x_k - y\|_Y \\ &\leq \|T\widetilde{x_k} - y_k\|_Y + \|\bar{T}x_k - y\|_Y < \frac{1}{k} + \|\bar{T}x_k - y\|_Y.\end{aligned}$$

Hence, $T\widetilde{x_k} \to y$ in Y as $k \to \infty$. So,

$$\widetilde{x_k} \in \mathfrak{D}(T), \quad \widetilde{x_k} \to x \quad \text{in} \quad X, \quad T\widetilde{x_k} \to y \quad \text{in} \quad Y \text{ as } k \to \infty.$$

Therefore, by the definition of $\bar{T}$

$$x \in \mathfrak{D}(\bar{T}) \quad \text{and} \quad \bar{T}x = y.$$

□

Remark 5.5.4. Let condition (5.27) be satisfied, then the operator $\bar{T}$ constructed in Step 2 of the proof of Lemma 5.5.2 is the minimal closed extension operator of T. Indeed, let $\mathfrak{D}(T) \subset \mathfrak{D}(T_1)$, $T_1 x = Tx \;\forall\, x \in \mathfrak{D}(T)$ and T_1 be closed. Let $x \in \mathfrak{D}(\bar{T})$ and $\bar{T}x = y$, then there exist $\{x_k\}_{k\in\mathbb{N}} \subset \mathfrak{D}(T)$, $x \in X$, $y \in Y$ such that $x_k \to x$ in X, $Tx_k \to x$ in Y as $k \to \infty$. Since T_1 is closed, $x \in \mathfrak{D}(T_1)$, $T_1 x = y = \bar{T}x$.

The operator $\bar{T}$ is called the closure of the operator T.

Example. Let $-\infty < a < b < \infty$, $X = Y = L_1(a, b)$. Then the operator

$$\frac{d}{dt} : C^1[a, b] \to L_1[a, b]$$

is closable. Its closure is the so-called weak differentiation operator $\overline{\frac{d}{dt}} \equiv (\frac{d}{dt})_w$, and $(\frac{d}{dt})_w \equiv f'_w$ is the weak (or generalized) derivative of f.

We shall use the following statement.

Lemma (du Bois-Reymond). *Let $f \in L_1[a, b]$ and*

$$\int_a^b f\varphi\, dt = 0 \quad \forall\, \varphi \in C^1[a, b] : \varphi(a) = \varphi(b) = 0,$$

then

$$f \sim 0 \quad \text{on} \quad [a, b].$$

Proof of closability. Let $x_k \in C^1[a,b],\ x_k \to 0$ in $L_1[a,b],\ y \in L_1[a,b],\ x_k' \to y$ in $L_1[a,b]$. Integrating by parts we get that $\forall\, \varphi \in C^1[a,b] : \varphi(a) = \varphi(b) = 0$

$$\begin{aligned} \int_a^b x_k'(t)\varphi(t)\,dt &= x_k(t)\varphi(t)\Big|_a^b - \int_a^b x_k(t)\varphi'(t)\,dt \\ &= -\int_a^b x_k(t)\varphi'(t)\,dt. \end{aligned}$$

By passing to the limit as $k \to \infty$ we obtain that

$$\begin{aligned} \Big|\int_a^b x_k(t)\varphi'(t)\,dt\Big| &\le \int_a^b |x_k(t)| \cdot |\varphi'(t)|\,dt \\ &\le \|\varphi'\|_{C[a,b]}\|x_k - 0\|_{L_1[a,b]} \to 0 \end{aligned}$$

and

$$\Big|\int_a^b x_k'(t)\varphi(t)\,dt - \int_a^b y(t)\varphi(t)\,dt\Big| \le \|x_k' - y\|_{L_1[a,b]}\|\varphi\|_{C[a,b]} \to 0.$$

So

$$\int_a^b y(t)\varphi(t)\,dt = 0 \quad \forall\, \varphi \in C^1[a,b] : \varphi(a) = \varphi(b) = 0$$

and

$$y(t) \sim 0 \quad \text{on} \quad [a,b].$$

□

Example.

$$|t|_w' = \operatorname{sgn} t \quad \text{on} \quad [-1,1].$$

We already proved that

$$\begin{aligned} x_k(t) &= \sqrt{t^2 + \frac{1}{k}} \to |t| \quad \text{in} \quad L_1[-1,1] \\ x_k'(t) &= \frac{t}{\sqrt{t^2 + \frac{1}{k}}} \to \operatorname{sgn} t \quad \text{in} \quad L_1[-1,1]. \end{aligned}$$

Hence

$$\overline{\frac{d}{dt}}(|t|) = \operatorname{sgn} t \iff |t|_w' = \operatorname{sgn} t.$$

5.5.3 Closed graph theorem

A natural question arises: "Under what conditions is a closed linear operator bounded?" An answer is given by the following important statement.

Theorem (Closed Graph Theorem). *Let X and Y be Banach spaces and $T : X \to Y$ be a closed linear operator. Then the operator T is bounded.*

Proof. By assumption the graph $\mathfrak{G}(T)$ is closed in $X \times Y$. Hence $\mathfrak{G}(T)$ is a Banach space. We now consider the mapping

$$P : \mathfrak{G}(T) \to X$$

defined by

$$P((x, Tx)) = x \quad \forall\, x \in X.$$

Clearly, P is linear. P is bounded because

$$\|P((x,Tx))\|_X = \|x\|_X \le \|x\|_X + \|Tx\|_Y = \|(x,Tx)\|_{X\times Y}.$$

The equality $P((x_1,Tx_1)) = P((x_2,Tx_2))$ means that $x_1 = x_2$, hence, P is one-to-one. Also

$$P(\mathfrak{G}(T)) = \bigcup_{x\in X} P((x,Tx)) = \bigcup_{x\in X} x = X,$$

hence, P maps $\mathfrak{G}(T)$ onto X. So, P is bijective and the inverse mapping is

$$P^{-1} : X \to \mathfrak{G}(T),$$

where

$$P^{-1}(x) = (x,Tx) \quad \forall\, x \in X.$$

Since $\mathfrak{G}(T)$, being a closed linear subspace of the Banach space $X \times Y$, is a Banach space, and X is a Banach space by the Bounded Inverse Theorem it follows that the operator P^{-1} is bounded. So, for some $c > 0$

$$\|P^{-1}x\|_{X\times Y} = \|(x,Tx)\|_{X\times Y} \le c\|x\|_X \quad \forall\, x \in X.$$

Hence, T is bounded because

$$\|Tx\|_Y \le \|x\|_X + \|Tx\|_Y = \|(x,Tx)\|_{X\times Y} \le c\|x\|_X \quad \forall\, x \in X.$$

□

Corollary. *Let X and Y be Banach spaces and $T : \mathfrak{D}(T) \to Y$ be a closed linear operator, where $\mathfrak{D}(T)$ is a linear subspace of X. If $\mathfrak{D}(T)$ is closed in X, then the operator T is bounded.*

Proof. It suffices to note that $\mathfrak{D}(T)$, being a closed linear subspace of the Banach space X, is itself a Banach space and apply the Closed Graph Theorem. □

5.6 Compact operators

Definition 5.6.1. *Let X, Y be normed spaces and $T : X \to Y$ be a linear operator. T is called a compact operator (or completely continuous) if for any bounded sequence $\{x_k\}_{k\in\mathbb{N}} \subset X$ the sequence $\{Tx_k\}_{k\in\mathbb{N}} \subset Y$ contains a convergent subsequence.*

Lemma 5.6.1. *Let X, Y be normed spaces. Then*

1) *every compact linear operator $T : X \to Y$ is bounded,*

2) *if $\dim X = \infty$, then the identity operator $I : X \to X$ (which is bounded) is not compact.*

Proof. 1) Let $M \subset X$ be a bounded set. It is required to prove that the set $T(M) \subset Y$ is also bounded. Assume to the contrary that $T(M)$ is unbounded, that is, there exists a sequence $\{x_k\}_{k\in\mathbb{N}} \subset M$ such that the sequence $\{Tx_k\}_{k\in\mathbb{N}} \subset Y$ is unbounded. Then there exist $k_1 \in \mathbb{N}$ such that $\|Tx_{k_1}\|_Y \ge 1$, $k_2 \in \mathbb{N}$, $k_2 > k_1$ such that $\|Tx_{k_2}\|_Y \ge \|Tx_{k_1}\|_Y + 1, \ldots,$ $k_m \in \mathbb{N}$, $k_m > k_{m-1}$ such that

$$\|Tx_{k_m}\|_Y \ge \|Tx_{k_1}\|_Y + \|Tx_{k_2}\|_Y + \ldots + \|Tx_{k_{m-1}}\|_Y + 1.$$

Therefore, for any $m, s \in \mathbb{N},\ m > s$

$$\begin{aligned}\|Tx_{k_m} - Tx_{k_s}\|_Y &\ge \|Tx_{k_m}\|_Y - \|Tx_{k_s}\|_Y \\ &\ge \|Tx_{k_1}\|_Y + \|Tx_{k_2}\|_Y + \ldots + \|Tx_{k_{m-1}}\|_Y - \|Tx_{k_s}\|_Y + 1 \ge m - 1.\end{aligned}$$

Hence, any subsequence of the sequence $\{Tx_{k_m}\}$ is not a Cauchy sequence, hence, is not convergent, which contradicts the compactness of T. So, $T(M)$ is bounded.

2) (proof for the case of inner-product spaces). Let $\dim X = \infty$. Then there exists an orthonormal basis $\{x_k\}_{k\in\mathbb{N}}$ for X. Clearly, the set $\{x_k\}_{k\in\mathbb{N}}$ is bounded. However, for any $k, m \in \mathbb{N}$, $k \neq m$,

$$\begin{aligned}\|Ix_k - Ix_m\|_X &= \|x_k - x_m\|_X = \sqrt{(x_k - x_m, x_k - x_m)_X} \\ &= \sqrt{(x_k, x_k) - (x_m, x_k) - (x_k, x_m) + (x_m, x_m)} = \sqrt{2}.\end{aligned}$$

Hence, any subsequence of the sequence $\{Ix_k\}_{k\in\mathbb{N}}$ is not a Cauchy sequence, hence, is not convergent. So, the operator I is not compact.

□

Theorem 5.6.1 (Weak convergence). *Let X and Y be normed spaces and $T : X \to Y$ a compact linear operator. Suppose that a sequence $\{x_k\}_{k\in\mathbb{N}} \subset X$ weakly converges to x in X as $k \to \infty$. Then the sequence $\{Tx_k\}_{k\in\mathbb{N}}$ converges strongly to Tx in Y as $k \to \infty$.*

Proof. $\boxed{\text{Step 1.}}$ First we show that $Tx_k \xrightarrow{w} Tx$ as $k \to \infty$. Let g be any bounded linear functional on Y. We define the functional f on X by setting

$$f(z) = g(Tz), \quad z \in X. \tag{5.28}$$

Since g and T are linear, f is also linear. Since T, being compact by Lemma 5.6.1, is bounded, f is bounded, because

$$|f(z)| = |g(Tz)| \le \|g\|_{Y^*}\|Tz\|_Y \le \|g\|_{Y^*}\|T\|_{X\to Y}\|z\|_X,$$

hence

$$\|f\|_{X^*} = \sup_{\substack{z\in X\\ z\neq 0}} \frac{|f(z)|}{\|z\|_X} \le \|g\|_{Y^*}\|T\|_{X\to Y} < \infty.$$

Since $x_k \xrightarrow{w} x$ as $k \to \infty$, we have

$$\lim_{k\to\infty} g(Tx_k) = \lim_{k\to\infty} f(x_k) = f(x) = g(Tx).$$

Since $g \in Y^*$ was arbitrary, it follows that $Tx_k \xrightarrow{w} Tx$ as $k \to \infty$.

$\boxed{\text{Step 2.}}$ Now we prove that $Tx_k \to Tx$ in Y (strong convergence). Assume the contrary. Then there exists $\eta > 0$ and a subsequence $\{x_{k_m}\}_{m\in\mathbb{N}}$ such that

$$\|Tx_{k_m} - Tx\|_Y \ge \eta. \tag{5.29}$$

Since the sequence $\{x_k\}_{k\in\mathbb{N}}$ is weakly convergent, it is bounded, and so is $\{x_{k_m}\}_{m\in\mathbb{N}}$. The compactness of T implies that $\{Tx_{k_m}\}_{m\in\mathbb{N}}$ has a convergent subsequence, say, $\{Tx_{k_{m_s}}\}_{s\in\mathbb{N}}$ converges to $y \in Y$. By Step 1 $\{Tx_{k_{m_s}}\}_{s\in\mathbb{N}}$ weakly converges to Tx, hence $y = Tx$. $\lim\limits_{s\to\infty} \|Tx_{k_{m_s}} - Tx\|_Y = 0$, which contradicts (5.29). So $Tx_k \to Tx$ in Y. □

Lemma 5.6.2. *Let X be a normed space, $T : X \to X$ a compact linear and $S : X \to X$ a bounded linear operator. Then the operators TS and ST are compact linear operators.*

Proof. $\boxed{\text{Step 1.}}$ Let $\{x_k\}_{k\in\mathbb{N}} \subset X$ be a bounded sequence. Since S is bounded, the sequence $\{Sx_k\}_{k\in\mathbb{N}} \subset X$ is also bounded. Since T is compact $\{TSx_k\}_{k\in\mathbb{N}}$ contains a convergent subsequence. So TS is compact.

$\boxed{\text{Step 2.}}$ Let $\{x_k\}_{k\in\mathbb{N}} \subset X$ be a bounded sequence. Since T is compact the sequence $\{Tx_k\}_{k\in\mathbb{N}}$ contains a convergent subsequence, say, $Tx_{k_m} \to y$ in X as $m \to \infty \iff \lim\limits_{m\to\infty} \|Tx_{k_m} - y\|_X = 0$. Since S is bounded and linear

$$\|STx_{k_m} - Sy\|_X = \|S(Tx_{k_m} - y)\|_X \le \|S\|_{X\to Y}\|Tx_{k_m} - y\|_X,$$

hence $\lim\limits_{m\to\infty} \|STx_{k_m} - Sy\|_X = 0 \iff STx_{k_m} \to Sy$ in X as $m \to \infty$. So ST is compact. □

Lemma 5.6.3. *Eigenvectors $x_1, \dots, x_n$ corresponding to different eigenvalues $\lambda_1 \dots, \lambda_n$ of a linear operator T on a linear space X constitute a linearly independent set.*

Proof. We assume that the set $\{x_1, \dots, x_n\}$ is linearly dependent and derive a contradiction. Let x_m be the first of the vectors which is a linear combination of $x_1, \dots, x_{m-1}$, say,

$$x_m = \alpha_1 x_1 + \ldots + \alpha_{m-1} x_{m-1},$$

where $x_1, \dots, x_{m-1}$ are linearly independent. Then

$$0 = (T - \lambda_m I)x_m = \sum_{j=1}^{m-1} \alpha_j (T - \lambda_m I)x_j = \sum_{j=1}^{m-1} \alpha_j (\lambda_j - \lambda_m)x_j.$$

Since $x_1, \dots, x_{m-1}$ are linearly independent

$$\alpha_j(\lambda_j - \lambda_m) = 0 \implies \alpha_j = 0, \quad j = 1, \dots, m-1.$$

Then $x_m = 0$ which contradicts the fact that $x_m \neq 0$ since x_m is an eigenvalue. $\square$

Theorem 5.6.2 (Eigenvalues). *The set of eigenvalues of a compact linear operator on a normed space X is countable (perhaps finite or even empty), and the only possible point of accumulation is $\lambda = 0$.*

In the proof of this theorem we shall use the following statement.

Theorem 5.6.3 (F. Riesz's Lemma). *Let X be a normed space, $M \neq \{0\}$, X be a closed linear subspace of X. Then for any $0 < \theta < 1$ there exists $x \in X$ such that*

$$\|x\|_X = 1 \quad \text{and} \quad \|x - y\|_X \geq \theta \quad \forall\, y \in M.$$

Proof of Theorem 5.6.3. Let $v \in X \setminus M$ be such that $\delta = \operatorname{dist}(v, M) = \inf_{y \in M} \|v - y\|_X$. Such $v \in X \setminus M$ exists, because otherwise $\delta = 0$ for all $x \in X \setminus M$, hence, due to the closedness of M, $x \in M$ and $X = M$, which contradicts the assumption $M \neq X$.

By definition of infimum, for any $0 < \theta < 1$ there exists $y_\theta \in M$ such that

$$\delta \leq \|v - y_\theta\|_X \leq \frac{\delta}{\theta}. \tag{5.30}$$

Let

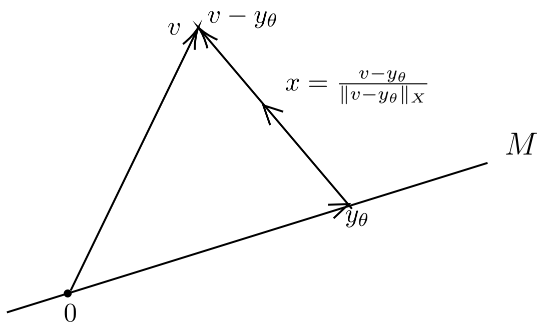
$$x = \frac{v - y_\theta}{\|v - y_\theta\|_X}. \tag{5.31}$$

Then $\|x\|_X = 1$ and for any $y \in M$

$$\|x - y\|_X = \left\| \frac{x - y_\theta}{\|x - y_\theta\|_X} - y \right\|_X = \frac{1}{\|x - y_\theta\|_X} \|x - (y_\theta + \|x - y_\theta\|_X y)\|_X \geq \frac{\delta}{\|x - y_\theta\|_X}$$

because $y_\theta + \|x - y_\theta\|_X y \in M$. Since $\|v - y_\theta\|_X \leq \frac{\delta}{\theta}$, it follows that

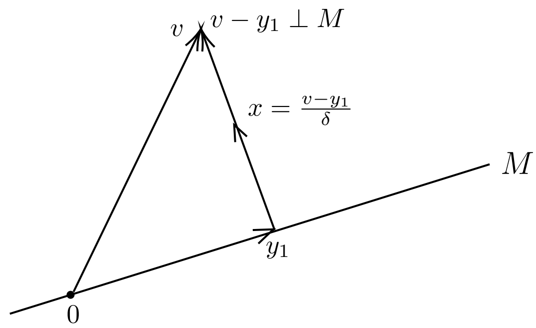
$$\|x - y\|_X \geq \frac{\delta}{(\frac{\delta}{\theta})} = \theta.$$



□

Remark 5.6.1. If there exists $y_1 \in X$ such that $\lim\limits_{\theta \to 1^-} y_\theta = y_1$ in X, then, due to the closedness of M, $y_1 \in M$ and, by passing to the limit in (5.30), it follows that $\|v - y_1\|_X = 1$. So the statement of the theorem holds also for $\theta = 1$. However, in general, the limit $\lim\limits_{\theta \to 1^-} y_\theta$ may not exist.

If X is a Hilbert space, then by Theorem 3.3.2 there exists $y_1 \in M$ such that $\|v - y_1\|_X = \delta$ (moreover, $v - y_1 \perp M$) and the statement of the theorem holds also for $\theta = 1$.



Proof of Theorem 5.6.2. It suffices to show that for any $r > 0$ the set of all eigenvalues λ such that $|\lambda| \geq r$ is finite.

Suppose to the contrary that for some $r_0 > 0$ the set of all eigenvalues λ such that $|\lambda| \geq r_0$ is infinite. Then there exists a sequence $\{\lambda_n\}_{n\in\mathbb{N}}$ of distinct eigenvalues λ_n such that $|\lambda_n| \geq r_0$. Also $Tx_n = \lambda_n x_n$ for some $x_n \in X$, $x_n \neq 0$. By Lemma 5.6.2 the set of all eigenvectors is linearly independent. Let $M_n = \mathrm{span}\{x_1, \ldots, x_n\}$. Then every $x \in M_n$ has a unique representation

$$x = \alpha_1 x_1 + \ldots + \alpha_n x_n,$$

where the α_j's are scalars. We apply $T - \lambda_n I$ and use the equalities $Tx_j = \lambda_j x_j$, $j = 1, \ldots, n$:

$$(T - \lambda_n I)x = \alpha_1(\lambda_1 - \lambda_n)x_1 + \ldots + \alpha_{n-1}(\lambda_{n-1} - \lambda_n)x_{n-1}. \tag{5.32}$$

Hence, $(T - \lambda_n I)x \in M_{n-1}$.

By Theorem 5.6.3 with $\theta = \frac{1}{2}$, taking into account that the sets M_n are closed, we get that there exists a sequence $\{y_n\}_{n\in\mathbb{N}}$ such that

$$y_n \in M_n, \quad \|y_n\|_X = 1, \quad \|y_n - x\|_X \geq \frac{1}{2} \quad \forall\, x \in M_n. \tag{5.33}$$

We shall prove that $\forall\, m, n \in \mathbb{N}$, $m < n$

$$\|Ty_n - Ty_m\|_X \geq \frac{r_0}{2} > 0, \tag{5.34}$$

so that no subsequence $\{Ty_{n_k}\}_{k\in\mathbb{N}}$ is fundamental, hence, convergent. This contradicts the compactness of T because the sequence $\{y_n\}_{n\in\mathbb{N}}$ is bounded.

We note that

$$Ty_n - Ty_m = \lambda_n y_n - \lambda_n y_n + Ty_n - Ty_m = \lambda_n y_n - \tilde{x},$$

where

$$\tilde{x} = \lambda_n y_n - Ty_n + Ty_m. \tag{5.35}$$

We show that $\tilde{x} \in M_{n-1}$. Indeed, $y_m \in M_m \subset M_{n-1}$:

$$y_m = \beta_1 x_1 + \ldots + \beta_{n-1} x_{n-1}$$

for some scalars β_j, hence

$$Ty_m = \beta_1 Tx_1 + \ldots + \beta_{n-1} Tx_{n-1} = \beta_1 \lambda_1 x_1 + \ldots + \beta_{n-1}\lambda_{n-1} x_{n-1} \in M_{n-1}.$$

By (5.32)

$$\lambda_n y_n - Ty_n = -(T - \lambda_n I) y_n \in M_{n-1},$$

hence by (5.35)

$$z = \frac{\tilde{x}}{\lambda_n} \in M_{n-1}$$

and by (5.33)

$$\|\lambda_n y_n - x\|_X = |\lambda_n| \|y_n - z\|_X \geq \frac{|\lambda_n|}{2} \geq \frac{r_0}{2},$$

we get inequality (5.34).

So, the assumption that there are infinitely many eigenvalues satisfying $|\lambda| \geq r_0$ for some $r_0 > 0$ is false and the theorem is proved. □

Remark. Theorem 5.6.2 shows that if a compact linear operator on a normed space has infinitely many eigenvalues, we can arrange these eigenvalues in a sequence with non-increasing moduli converging to 0:

$$|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_n| \geq \ldots,$$
$$\lim_{n\to\infty} \lambda_n = 0.$$

In order to prove our next theorem we shall need the following statement about compact sets.

Theorem 5.6.4. *If the closed unit ball of a normed space is compact, then this space is finite-dimensional.*

Proof. Let X be a normed space and $\bar{B} \equiv \bar{B}_X(0,1) = \{x \in X : \|x\|_X \leq 1\}$ the closed unit ball. Assume to the contrary that $\bar{B}$ is compact but $\dim X = \infty$.

We choose any $x_1 \in X$ with $\|x_1\|_X = 1$. Then x_1 generates a one-dimensional subspace X_1 of X, which is closed and is a proper subspace of X since $\dim X = \infty$. By Theorem 5.6.3 with $\theta = \frac{1}{2}$ there exists $x_2 \in X$ such that

$$\|x_2\|_X = 1, \quad \|x_2 - x_1\|_X \geq \frac{1}{2}.$$

The elements x_1, x_2 generate a two-dimensional proper closed subspace X_2 of X. By Theorem 5.6.3 there exists $x_3 \in X$ such that

$$\|x_3\|_X = 1, \quad \|x_3 - x\|_X \geq \frac{1}{2} \quad \forall\, x \in X_2.$$

In particular,

$$\|x_3\|_X = 1, \quad \|x_3 - x_1\|_X \geq \frac{1}{2}, \quad \|x_3 - x_2\|_X \geq \frac{1}{2}.$$

Proceeding by induction, we obtain a sequence $\{x_n\}_{n\in\mathbb{N}} \subset \bar{B}$ such that

$$\|x_m - x_n\|_X \geq \frac{1}{2} \quad \forall\, m, n \in \mathbb{N},\ m \neq n.$$

Hence, $\{x_n\}_{n\in\mathbb{N}}$ cannot have a convergent subsequence. This contradicts the compactness of $\bar{B}$. Hence, the assumption $\dim X = \infty$ is false, and $\dim X < \infty$. $\square$

Theorem 5.6.5. *Let $T : X \to X$ be a compact linear operator on a normed space X. Then for every $\lambda \neq 0$ the null space $\mathcal{N}(T_\lambda)$, where $T_\lambda = T - \lambda I$ is finite-dimensional.*

Proof. Consider the normed space $\mathcal{N}(T_\lambda)$ and let $\{x_n\}_{n\in\mathbb{N}} \subset \bar{B} \equiv \bar{B}_{\mathcal{N}(T_\lambda)}(0,1)$. Then $\{x_n\}_{n\in\mathbb{N}}$ is bounded, hence, the sequence $\{Tx_n\}_{n\in\mathbb{N}}$ has a convergent subsequence $\{Tx_{n_k}\}_{k\in\mathbb{N}}$. Since $Tx_{n_k} = \lambda x_{n_k}$ and $\lambda \neq 0$ the sequence $\{x_{n_k}\}_{k\in\mathbb{N}} = \{\frac{Tx_{n_k}}{\lambda}\}_{k\in\mathbb{N}}$ also converges. Its limit belongs to $\bar{B}$ since $\bar{B}$ is closed. Hence $\bar{B}$ is compact as $\{x_n\}_{n\in\mathbb{N}}$ was arbitrarily chosen in $\bar{B}$. By Theorem 5.6.4 $\dim \mathcal{N}(T_\lambda) < \infty$. $\square$

Remark. Recall that the dimension of the set of all eigenvectors of a linear operator $T : X \to X$ corresponding to an eigenvalue λ is called the multiplicity of λ. By Theorem 5.6.5 all eigenvalues $\lambda \neq 0$ of a compact linear operator T (if any) have finite multiplicities.

Example (Integral operators). Let $-\infty < a < b < \infty$, $k \in C([a,b] \times [a,b])$ and $\forall\, f \in C[a,b]$

$$(Tf)(x) = \int_a^b k(x,y) f(y)\, dy.$$

Then $T : C[a,b] \to C[a,b]$ is a bounded linear operator. Indeed, let $M \subset C[a,b]$ be a bounded set. Then $\forall\, f \in M$

$$\begin{aligned}
|(Tf)(x)| &= \Big| \int_a^b k(x,y) f(y)\, dy \Big| \\
&\leq \int_a^b |k(x,y)||f(y)|\, dy \\
&\leq \|k\|_{C([a,b]\times[a,b])} \|f\|_{C[a,b]} (b-a) \\
\sup_{f\in M} \|Tf\|_{C[a,b]} &\leq \|k\|_{C([a,b]\times[a,b])} (b-a) \sup_{f\in M} \|f\|_{C[a,b]}.
\end{aligned}$$

So the set $T(M)$ is bounded.

Since $k \in C([a,b] \times [a,b])$ it is uniformly continuous on $[a,b] \times [a,b]$. In particular, $\forall\, \varepsilon > 0\ \exists\, \delta > 0$ such that $\forall\, x, z, y \in [a,b]$ with $|x - z| < \delta \quad |k(x,y) - k(z,y)| < \varepsilon$. Therefore $\forall\, x, z \in [a,b]$ and $\forall\, f \in M$

$$\begin{aligned}
&|(Tf)(x) - (Tf)(z)| = \Big| \int_a^b (k(x,y) - k(z,y)) f(y)\, dy \Big| \\
&\leq \int_a^b |k(x,y) - k(z,y)||f(y)|\, dy < \varepsilon \int_a^b |f(y)|\, dy \\
&\leq \varepsilon \|f\|_{C[a,b]} (b-a) \leq \Big(\sup_{f\in M} \|f\|_{C[a,b]} (b-a) \Big) \varepsilon.
\end{aligned}$$

Hence, the functions $f \in T(M)$ are equicontinuous.

Hence, by the Arzelà-Ascoli theorem the set $T(M)$ is precompact, which means that for any bounded sequence $\{f_k\}_{k\in\mathbb{N}} \subset C[a,b]$ the sequence $\{Tf_k\}_{k\in\mathbb{N}} \subset C[a,b]$ contains a subsequence convergent in $C[a,b]$.

So T is a compact (= completely continuous) linear operator.

5.7 Elements of spectral theory

Theorem (Eigenvalues of an operator). *All matrices representing a given linear operator $T : X \to X$ on a finite dimensional normed space X relative to various bases for X have the same eigenvalues.*

Proof. We must see what happens in the transition from one basis for X to another. Let $e = (e_1, \dots, e_n)$ and $\tilde{e} = (\tilde{e}_1, \dots, \tilde{e}_n)$ be any bases for X, written as row vectors. By the definition of a basis, each e_j is a linear combination of the $\tilde{e}_k$'s and conversely. We can write this

$$\tilde{e} = eC \quad \text{or} \quad \tilde{e}^T = C^T e^T, \tag{5.36}$$

where C is a nonsingular n-rowed square matrix. Every $x \in X$ has a unique representation with respect to each of the two bases, say,

$$x = ex_1 = \sum \xi_j e_j = \tilde{e}x_2 = \sum \tilde{\xi}_k \tilde{e}_k,$$

where $x_1 = (\xi_j)$ and $x_2 = (\tilde{\xi}_k)$ are column vectors. From this and (5.36) we have $ex_1 = \tilde{e}x_2 = eCx_2$. Hence

$$x_1 = Cx_2. \tag{5.37}$$

Similarly, for $Tx = y = ey_1 = \tilde{e}y_2$ we have

$$y_1 = Cy_2. \tag{5.38}$$

Consequently, if T_1 and T_2 denote the matrices which represent T with respect to e and $\tilde{e}$, respectively, then

$$y_1 = T_1 x_1 \quad \text{and} \quad y_2 = T_2 x_2,$$

and from this and (5.37) and (5.38),

$$CT_2x_2 = Cy_2 = y_1 = T_1x_1 = T_1Cx_2.$$

Premultiplying by C^{-1} we obtain the transformation law

$$T_2 = C^{-1}T_1C \tag{5.39}$$

with C determined by the bases according to (5.36) (and independent of T). Using (5.39) and $\det(C^{-1})\det C = 1$, we can now show that the characteristic determinants of T_2 and T_1 are equal:

$$\begin{aligned} \det(T_2 - \lambda I) &= \det\big(C^{-1}T_1C - \lambda C^{-1}IC\big) \\ &= \det\big(C^{-1}(T_1 - \lambda I)C\big) \\ &= \det\big(C^{-1}\big)\det(T_1 - \lambda I)\det C \\ &= \det(T_1 - \lambda I). \end{aligned}$$

□

Corollary. *A linear operator on a finite-dimensional normed space has at least one eigenvalue.*

Next, we consider linear operators in infinite dimensional normed spaces.

Let $X \neq \{0\}$ be a complex normed space and $T : \mathfrak{D}(T) \to Y$ a linear operator with domain $\mathfrak{D}(T) \subset X$. With T we associate the operator

$$T_\lambda = T - \lambda I,$$

where λ is a complex number and I is the identity operator on $\mathfrak{D}(T)$. If T_λ has an inverse, we denote it by $R_\lambda(T)$, that is,

$$R_\lambda(T) = T_\lambda^{-1} = (T - \lambda I)^{-1}$$

and call it the *resolvent operator* of T or, simply, the resolvent of T. Instead of $R_\lambda(T)$ we also write simply R_λ if it is clear to what operator T we refer in a specific discussion.

The name "resolvent" is appropriate, since $R_\lambda(T)$ helps to solve the equation $T_\lambda x = y$. Thus, $x = T_\lambda^{-1} y = R_\lambda(T) y$ provided $R_\lambda(T)$ exists.

More important, the investigation of properties of R_λ will be basic for an understanding of the operator T itself. Naturally, many properties of T_λ and R_λ, depend on λ. And *spectral theory* is concerned with those properties. For instance, we shall be interested in the set of all λ in the complex plane such that R_λ exists. Boundedness of R_λ is another property that will be essential. We shall also ask for what λ's the domain of R_λ is dense in X, to name just a few aspects.

Definition 5.7.1 (Regular value, resolvent set, spectrum). *Let $X \neq \{0\}$ be a complex normed space and $T : \mathfrak{D}(T) \to X$ a linear operator with domain $\mathfrak{D}(T) \subset X$. A regular value λ of T is a complex number such that*

(R1) $R_\lambda(T)$ *exists,*

(R2) $R_\lambda(T)$ *is bounded,*

(R3) $R_\lambda(T)$ *is defined on a set which is dense in X.*

The resolvent set $\rho(T)$ of T is the set of all regular values λ of T. Its complement $\sigma(T) = \mathbb{C} - \rho(T)$ in the complex plane $\mathbb{C}$ is called the spectrum of T, and a $\lambda \in \sigma(T)$ is called a spectral value of T. Furthermore, the spectrum $\sigma(T)$ is partitioned into three disjoint sets as follows.

The **point spectrum** or *discrete spectrum* $\rho_p(T)$ is the set such that $R_\lambda(T)$ does not exist. A $\lambda \in \sigma_p(T)$ is called an eigenvalue of T.

The **continuous spectrum** $\rho_c(T)$ is the set such that $R_\lambda(T)$ exists and satisfies (R3) but not (R2), that is, $R_\lambda(T)$ is unbounded.

The **residual spectrum** $\rho_r(T)$ is the set such that $R_\lambda(T)$ exists (and may be bounded or not) but does not satisfy (R3), that is, the domain of $R_\lambda(T)$ is not dense in X.

To avoid trivial misunderstandings, let us say that some of the sets in this definition may be empty. This is an existence problem which we shall have to discuss. For instance, $\rho_c(T) = \rho_r(T) = \emptyset$ in the finite dimensional case.

The conditions stated in Definition 5.7.1 can be summarized in the following table.

Satisfied	Not satisfied	λ belongs to:
(R1), (R2), (R3)		$\rho(T)$
	(R1)	$\sigma_p(T)$
(R1) (R3)	(R2)	$\sigma_c(T)$
(R1)	(R3)	$\sigma_r(T)$

If X is infinite dimensional, then T can have spectral values which are not eigenvalues:

Example 5.4 (Operator with a spectral value which is not an eigenvalue). On the Hilbert sequence space $X = \ell^2$ we define a linear operator $T : \ell^2 \to \ell^2$ by

$$(\xi_1, \xi_2, \ldots) \mapsto (0, \xi_1, \xi_2, \ldots), \tag{5.40}$$

where $x = (\xi_j) \in \ell^2$. The operator T is called the *right-shift operator*. T is bounded (and $\|T\| = 1$) because

$$\|Tx\|^2 = \sum_{j=1}^{\infty} |\xi_j|^2 = \|x\|^2.$$

The operator $R_0(T) = T^{-1} : T(X) \to X$ exists; in fact, it is the *left-shift operator* given by

$$(\xi_1, \xi_2, \ldots) \mapsto (\xi_2, \xi_3, \ldots).$$

But $R_0(T)$ does not satisfy (R3), because (5.40) shows that $T(X)$ is not dense in X; indeed, $T(X)$ is the subspace Y consisting of all $y = (\eta_j)$ with $\eta_1 = 0$. Hence, by definition, $\lambda = 0$ is a spectral value of T. Furthermore, $\lambda = 0$ is not an eigenvalue. We can see this directly from (5.40) since $Tx = 0$ implies $x = 0$ and the zero vector is not an eigenvector.