Functional Analysis

Lyudmila Turowska

Department of Mathematical Sciences
Chalmers University of Technology and the University of Gothenburg

These lectures concern the theory of normed spaces and bounded linear operators between normed spaces. We shall define these terms and study their properties.

You can consult the following books.

1. G. Folland: Real Analysis. Modern Techniques and their Applications, John Wiley & Sons, 1999, Chapters 5-7 and parts of Chapter 4.

1 Vector Spaces

Definition 1.1. A *complex vector space* (or a vector space over the field $\mathbb{C}$) (or a complex linear space) is a set V with addition

$$\begin{aligned}&V \times V \to V,\\ &(x, y) \mapsto x + y \quad \text{the sum of } x \text{ and } y,\end{aligned}$$

and scalar multiplication

$$\begin{aligned}&\mathbb{C} \times V \to V,\\ &(\lambda, x) \mapsto \lambda x \quad \text{the product of } \lambda \text{ and } x,\end{aligned}$$

satisfying the following rules:

$$\begin{aligned}
&x + (y + z) = (x + y) + z \quad \forall x, y, z \in V,\\
&x + y = y + x \quad \forall x, y \in V,\\
&\exists \underline{0} \in V : \underline{0} + x = x \quad \forall x \in V,\\
&\forall x \in V\ \exists -x \in V : x + (-x) = \underline{0},\\
&\lambda(x + y) = \lambda x + \lambda y \quad \forall x, y \in V,\ \forall \lambda \in \mathbb{C},\\
&(\lambda + \mu)x = \lambda x + \mu x \quad \forall x \in V,\ \forall \lambda, \mu \in \mathbb{C},\\
&\lambda(\mu x) = (\lambda\mu)x \quad \forall x \in V,\ \forall \lambda, \mu \in \mathbb{C},\\
&1x = x \quad \forall x \in V.
\end{aligned}$$

It is easy to see that the usual arithmetic holds in V.

Remark 1.2. Similarly, we can define a *real vector space* by replacing $\mathbb{C}$ with $\mathbb{R}$ in the definition.

Examples 1.3. 1. $\mathbb{C}^n = \{(x_1, \ldots, x_n) : x_i \in \mathbb{C}\ \forall i\}$, *the complex vector space of n-tuples of complex numbers* with coordinatewise addition

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n), \quad (x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbb{C}^n$$

and coordinatewise scalar multiplication

$$\lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n), \quad \lambda \in \mathbb{C},\ (x_1, \ldots, x_n) \in \mathbb{C}^n.$$

2. An infinite-dimensional complex vector space:

$$\mathbb{C}^{\mathbb{N}} = \{(x_1, \ldots, x_n, \ldots) : x_i \in \mathbb{C}\ \forall i\}$$

is *the complex vector space of sequences of complex numbers* with coordinatewise addition and scalar multiplication:

$$x + y = (x_1 + y_1, \ldots, x_n + y_n, \ldots),$$
$$\lambda x = (\lambda x_1, \ldots, \lambda x_n, \ldots),$$

$x = (x_1, \ldots, x_n, \ldots), y = (y_1, \ldots, y_n, \ldots) \in \mathbb{C}^{\mathbb{N}},\ \lambda \in \mathbb{C}$.
3. *The complex vector space of all continuous complex-valued functions on* $[0, 1]$

$$C[0, 1] = \{f : [0, 1] \to \mathbb{C} : f \text{ is continuous on } [0, 1]\}$$

with *pointwise addition*

$$(f + g)(t) = f(t) + g(t), \quad f, g \in C[0, 1],\ t \in [0, 1],$$

and scalar multiplication

$$(\lambda f)(t) = \lambda f(t), \quad f \in C[0, 1],\ \lambda \in \mathbb{C}, t \in [0, 1].$$

These operations are well-defined due to the following fact (from the basic analysis course): If f, g are continuous functions on $[0, 1]$ and $\lambda \in \mathbb{C}$ then $(f + g)$ and λf are continuous on $[0, 1]$.
4. $M_n(\mathbb{C}) = \{(a_{ij})_{i,j=1}^n : a_{ij} \in \mathbb{C}, i, j = 1, \ldots, n\}$, *the complex vector space of* $n \times n$ *matrices* with complex entries with addition

$$(a_{ij})_{i,j=1}^n + (b_{ij})_{i,j=1}^n = (a_{ij} + b_{ij})_{i,j=1}^n,$$

and scalar multiplication

$$\lambda(a_{ij})_{i,j=1}^n = (\lambda a_{ij})_{i,j=1}^n.$$

(So we have addition and scalar multiplication by *entries*.)

1.1 Vector Subspaces

Definition 1.4. A non-empty subset of W of a vector space V is called a *vector subspace* (or a *linear subspace*) if

1. for every $x, y \in W,\ x + y \in W$;

2. for every $x \in W$ and every $\lambda \in \mathbb{C}$ (or $\mathbb{R}$), $\lambda x \in W$.

Remark 1.5. A *vector subspace* W is itself a *vector space* with respect to the addition and scalar multiplication defined on V.

Question 1.6. Let $\overline{\mathbb{D}}$ be the closed unit disc, that is

$$\overline{\mathbb{D}} = \{z \in \mathbb{C} : |z| \leq 1\}.$$

It is obvious that $\overline{\mathbb{D}}$ is a subset of $\mathbb{C}$. Is $\overline{\mathbb{D}}$ a vector subspace of $\mathbb{C}$?

No. Conditions (1) and (2) of the definition of a vector subspace are not satisfied:

1. For $z_1 = z_2 = 1$, we have $z_1 + z_2 = 2$, so $|z_1 + z_2| = 2$. Thus $z_1 + z_2 \notin \overline{\mathbb{D}}$.
2. For $z = i$, we have $|\lambda z| = |\lambda|$ for every $\lambda \in \mathbb{C}$. Thus $\lambda z \notin \overline{\mathbb{D}}$ when λ is such that $|\lambda| > 1$.

Example 1.7. Show that $W = \{f \in C[0,1] : \int_0^1 f(x)\,dx = 0\}$ is a linear subspace of $C[0,1]$.

Proof. 1) It is obvious that W is a subset of $C[0,1]$, i.e., $W \subset C[0,1]$.

2) W is a *non-empty* subset. For example, $f_0(t) = 2t - 1$ belongs to W, since $\int_0^1 f_0(t)\,dt = \int_0^1 2t - 1\,dt = (2\frac{t^2}{2} - t)|_0^1 = 0$.

3) For every $f, g \in W$, we have

$$\begin{aligned}\int_0^1 (f+g)(t)\,dt &= \int_0^1 f(t) + g(t)\,dt \\ &= \int_0^1 f(t)\,dt + \int_0^1 g(t)\,dt \quad \text{(by the linearity of integrals)} \\ &= 0 + 0 = 0.\end{aligned}$$

Thus $f + g \in W$.

4) For every $f \in W$ and $\lambda \in \mathbb{C}$, we have

$$\begin{aligned}\int_0^1 (\lambda f)(t)\,dt &= \int_0^1 \lambda f(t)\,dt \\ &= \lambda \int_0^1 f(t)\,dt \quad \text{(by the linearity of integrals)} \\ &= 0.\end{aligned}$$

Thus $\lambda f \in W$. Hence W is a vector subspace of $C[0,1]$. □

Example 1.8. Let ℓ^2 be the set of all complex sequences $x = (x_n)_{n=1}^{\infty}$ which are square summable, i.e. satisfy

$$\sum_{n=1}^{\infty} |x_n|^2 < \infty;$$

we write

$$\ell^2 = \{(x_1, \dots, x_n, \dots) : x_i \in \mathbb{C}\ \forall i \text{ and } \sum_{n=1}^{\infty} |x_n|^2 < \infty\}.$$

Show that ℓ^2 is a vector subspace of $\mathbb{C}^{\mathbb{N}}$. Recall that $\mathbb{C}^{\mathbb{N}} = \{(x_1, \dots, x_n, \dots) : x_i \in \mathbb{C}\ \forall i\}$ with $x + y = (x_1 + y_1, \dots, x_n + y_n, \dots)$ and $\lambda x = (\lambda x_1, \dots, \lambda x_n, \dots)$.

Proof. 1). It is obvious that ℓ^2 is a subset of $\mathbb{C}^{\mathbb{N}}$.

2). The sequence $x = (1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, \dots) \in \ell^2$ since the series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges. Thus ℓ^2 is non-empty.

3). We have to prove that, for any $x = (x_n)_{n=1}^{\infty}, y = (y_n)_{n=1}^{\infty} \in \ell^2$, $x + y = (x_1 + y_1, \dots, x_n + y_n, \dots)$ belongs to ℓ^2. Thus we have to show that the series $\sum_{n=1}^{\infty} |x_n + y_n|^2$ converges. We will prove this in Theorem 1.11.

4). For any $x = (x_n)_{n=1}^{\infty} \in \ell^2$ and $\lambda \in \mathbb{C}$ we have $\lambda x = (\lambda x_1, \dots, \lambda x_n, \dots)$. Note that

$$\begin{aligned}\sum_{n=1}^{\infty} |\lambda x_n|^2 &= \sum_{n=1}^{\infty} |\lambda|^2 |x_n|^2 \\ &= |\lambda|^2 \sum_{n=1}^{\infty} |x_n|^2 < \infty,\end{aligned}$$

so $\lambda x \in \ell^2$.

Hence ℓ^2 is a vector subspace of $\mathbb{C}^{\mathbb{N}}$, and ℓ^2 is a complex vector space. □

The Cauchy-Schwarz Inequality 1.9.

$$\sum_{k=1}^{n} a_k b_k \le (\sum_{k=1}^{n} a_k^2)^{\frac{1}{2}} (\sum_{k=1}^{n} b_k^2)^{\frac{1}{2}}$$

for all $a_k, b_k \in \mathbb{R}, a_k, b_k \ge 0, k = 1, 2, \dots, n, n \in \mathbb{N}$.

It follows from

$$\left(\sum_{k=1}^{n} a_k b_k\right)^2 = \sum_{k=1}^{n} a_k^2 \sum_{k=1}^{n} b_k^2 - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i b_j - b_i a_j)^2.$$

Lemma 1.10. *For any $x, y \in \ell^2$, the series $\sum_{n=1}^{\infty} x_n \overline{y_n}$ converges absolutely and*

$$\sum_{n=1}^{\infty} |x_n \overline{y_n}| \leq \left\{\sum_{n=1}^{\infty} |x_n|^2\right\}^{\frac{1}{2}} \left\{\sum_{n=1}^{\infty} |y_n|^2\right\}^{\frac{1}{2}}.$$

Proof. Using the Cauchy-Schwarz inequality, for every $k \in \mathbb{N}$, we have

$$\begin{aligned}
\sum_{n=1}^{k} |x_n \overline{y_n}| &= \sum_{n=1}^{k} |x_n||\overline{y_n}| \\
&\leq \left\{\sum_{n=1}^{k} |x_n|^2\right\}^{\frac{1}{2}} \left\{\sum_{n=1}^{k} |y_n|^2\right\}^{\frac{1}{2}} \quad \text{by the Cauchy-Schwarz inequality} \\
&\leq \left\{\sum_{n=1}^{\infty} |x_n|^2\right\}^{\frac{1}{2}} \left\{\sum_{n=1}^{\infty} |y_n|^2\right\}^{\frac{1}{2}}.
\end{aligned}$$

The latter expression is a finite number independent of k, and so the series

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

converges and

$$\sum_{n=1}^{\infty} |x_n \overline{y_n}| \leq \left\{\sum_{n=1}^{\infty} |x_n|^2\right\}^{\frac{1}{2}} \left\{\sum_{n=1}^{\infty} |y_n|^2\right\}^{\frac{1}{2}}.$$

□

Theorem 1.11. *For any $x, y \in \ell^2$,*

$$\sum_{n=1}^{\infty} |x_n + y_n|^2 \leq \left(\left\{\sum_{n=1}^{\infty} |x_n|^2\right\}^{\frac{1}{2}} + \left\{\sum_{n=1}^{\infty} |y_n|^2\right\}^{\frac{1}{2}}\right)^2$$

and $x + y \in \ell^2$.

Proof. For every $n \in \mathbb{N}$,

$$\begin{aligned} |x_n + y_n|^2 &= (x_n + y_n)(\overline{x_n} + \overline{y_n}) \\ &= x_n\overline{x_n} + y_n\overline{x_n} + x_n\overline{y_n} + y_n\overline{y_n}) \\ &= |x_n|^2 + x_n\overline{y_n} + \overline{x_n\overline{y_n}} + |y_n|^2 \\ &\leq |x_n|^2 + 2|x_n\overline{y_n}| + |y_n|^2. \end{aligned}$$

Hence, for every $k \in \mathbb{N}$,

$$\begin{aligned} \sum_{n=1}^{k} |x_n + y_n|^2 &\leq \sum_{n=1}^{k}(|x_n|^2 + 2|x_n\overline{y_n}| + |y_n|^2) \\ &= \sum_{n=1}^{k} |x_n|^2 + 2\sum_{n=1}^{k} |x_n\overline{y_n}| + \sum_{n=1}^{k} |y_n|^2 \\ &\leq \sum_{n=1}^{\infty} |x_n|^2 + 2\sum_{n=1}^{\infty} |x_n\overline{y_n}| + \sum_{n=1}^{\infty} |y_n|^2 \\ &\leq \sum_{n=1}^{\infty} |x_n|^2 + 2\left(\sum_{n=1}^{\infty} |x_n|^2\right)^{\frac{1}{2}} \left(\sum_{n=1}^{\infty} |y_n|^2\right)^{\frac{1}{2}} + \sum_{n=1}^{\infty} |y_n|^2 \\ &= \left(\left\{\sum_{n=1}^{\infty} |x_n|^2\right\}^{\frac{1}{2}} + \left\{\sum_{n=1}^{\infty} |y_n|^2\right\}^{\frac{1}{2}}\right)^2. \end{aligned}$$

The latter expression is a finite number independent of k, so the series

$$\sum_{n=1}^{\infty} |x_n + y_n|^2$$

converges and

$$\sum_{n=1}^{\infty} |x_n + y_n|^2 \leq \left(\left\{\sum_{n=1}^{\infty} |x_n|^2\right\}^{\frac{1}{2}} + \left\{\sum_{n=1}^{\infty} |y_n|^2\right\}^{\frac{1}{2}}\right)^2.$$

The theorem is proved. $\square$

2 Normed Spaces

Suppose we have a vector space V, the generalisation of $\mathbb{R}^n$, and we want to know about size of vectors $x \in V$, and about the distance between two vectors $x, y \in V$. Suppose we would also like to estimate the distance from a vector $x \in V$ to a vector subspace $W \subset V$. To deal with this kind of problem we define a *norm* on V.

Definition 2.1. A mapping $\|\cdot\| : V \to \mathbb{R} : x \mapsto \|x\|$, where V is a complex (real) vector space, is called a *norm* if it satisfies

N1 $\|x\| > 0$ for all $x \in V \setminus \{\underline{0}\}$;

N2 $\|\lambda x\| = |\lambda|\|x\|$ for all scalars $\lambda \in \mathbb{C}$ ($\lambda \in \mathbb{R}$) and $x \in V$;

N3 $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in V$ (The Triangle Inequality).

Remark 2.2. Note that N2 implies that $\|\underline{0}\| = 0$. Thus we have $\|x\| = 0$ if and only if $x = \underline{0}$.

Example 2.3. $\mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$,

$$\|\cdot\|_2 : \mathbb{R}^2 \to \mathbb{R} : (x, y) \mapsto \|(x, y)\|_2 = \sqrt{x^2 + y^2}.$$

This is called the Euclidean norm of the vector. The Euclidean norm is just one way to measure length. Now we shall check that $\|(x, y)\|_2 = \sqrt{x^2 + y^2}$ really defines a norm on $\mathbb{R}^2$, that is, we shall check that conditions N1, N2, N3 are satisfied.

N1. For any $v = (x, y) \neq \underline{0}$,

$$\|v\|_2 = \sqrt{x^2 + y^2} > 0.$$

N2. For any $v \in \mathbb{R}^2$ and $\lambda \in \mathbb{R}$,

$$\begin{aligned}
\|\lambda v\|_2 &= \|(\lambda x, \lambda y)\|_2 \\
&= \sqrt{(\lambda x)^2 + (\lambda y)^2} \\
&= \sqrt{\lambda^2 (x^2 + y^2)} \\
&= |\lambda|\sqrt{x^2 + y^2} \\
&= |\lambda|\|v\|_2.
\end{aligned}$$

N3. For any $v_1 = (x_1, y_1)$ and $v_2 = (x_2, y_2) \in \mathbb{R}^2$,

$$\begin{aligned}
\|v_1 + v_2\|_2^2 &= \|(x_1 + x_2, y_1 + y_2)\|_2^2 \\
&= (x_1 + x_2)^2 + (y_1 + y_2)^2 \\
&= x_1^2 + 2x_1x_2 + x_2^2 + y_1^2 + 2y_1y_2 + y_2^2 \\
&= (x_1^2 + y_1^2) + (x_2^2 + y_2^2) + 2(x_1x_2 + y_1y_2) \\
&= \|v_1\|_2^2 + \|v_2\|_2^2 + 2(x_1x_2 + y_1y_2).
\end{aligned}$$

By the Cauchy-Schwarz inequality,

$$\begin{aligned}|x_1x_2 + y_1y_2| &\le (x_1^2 + y_1^2)^{\frac{1}{2}}(x_2^2 + y_2^2)^{\frac{1}{2}} \\ &= \|v_1\|_2\|v_2\|_2.\end{aligned}$$

Thus

$$\begin{aligned}\|v_1 + v_2\|_2^2 &\le \|v_1\|_2^2 + 2\|v_1\|_2\|v_2\|_2 + \|v_2\|_2^2 \\ &= (\|v_1\|_2 + \|v_2\|_2)^2.\end{aligned}$$

Therefore, $\|v_1 + v_2\|_2 \le \|v_1\|_2 + \|v_2\|_2$. It is called the triangle inequality.

Definition 2.4. A pair $(V, \|\cdot\|)$, where V is a real or complex vector space and $\|\cdot\|$ is a norm on V is called a *normed space.*

Example 2.5. $(\mathbb{R}^2, \|\cdot\|_2)$ is a normed space, where $\mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$ and $\|(x, y)\|_2 = \sqrt{x^2 + y^2}$.

Example 2.6. $(\ell^2, \|\cdot\|_2)$ is a normed space, where

$$\ell^2 = \{(x_1, \ldots, x_n, \ldots) : x_i \in \mathbb{C}\ \forall i \text{ and } \textstyle\sum_{n=1}^{\infty}|x_n|^2 < \infty\}$$

and

$$\|x\|_2 = \left\{\sum_{n=1}^{\infty}|x_n|^2\right\}^{\frac{1}{2}}.$$

Proof. We proved before that ℓ^2 is a complex vector space. Thus we have to verify that $\|\cdot\|_2$ is a norm on ℓ^2.

The mapping $\|\cdot\|_2 : \ell^2 \to \mathbb{R} : x \mapsto \|x\|_2 = (\sum_{n=1}^{\infty}|x_n|^2)^{\frac{1}{2}}$ is well-defined, since the series $\sum_{n=1}^{\infty}|x_n|^2$ converges. Now we shall check that conditions N1, N2, N3 are satisfied.

N1. It is obvious that, for $x \in \ell^2 \setminus \{\underline{0}\}$, i.e., $x \neq \underline{0}$, there is an integer $i_0 \in \mathbb{N}$ such that $x_{i_0} \neq 0$. Thus

$$\|x\|_2 = \left\{\sum_{n=1}^{\infty}|x_n|^2\right\}^{\frac{1}{2}} \ge |x_{i_0}| > 0.$$

N2. We need to show that $\|\lambda x\|_2 = |\lambda|\|x\|_2$ for all $\lambda \in \mathbb{C}$ and all $x \in \ell^2$. We have

$$
\begin{aligned}
\|\lambda x\|_2 &= \left\{\sum_{n=1}^{\infty} |\lambda x_n|^2\right\}^{\frac{1}{2}} \\
&= \left\{\sum_{n=1}^{\infty} |\lambda|^2 |x_n|^2\right\}^{\frac{1}{2}} \\
&= |\lambda| \left\{\sum_{n=1}^{\infty} |x_n|^2\right\}^{\frac{1}{2}} \\
&= |\lambda| \|x\|_2
\end{aligned}
$$

for all $\lambda \in \mathbb{C}$ and all $x \in \ell^2$, as required.

N3. We need to show that $\|x+y\|_2 \leq \|x\|_2 + \|y\|_2$ for all $x, y \in \ell^2$. By Theorem 1.11, for any $x, y \in \ell^2$,

$$
\left(\sum_{n=1}^{\infty} |x_n + y_n|^2\right)^{\frac{1}{2}} \leq \left(\sum_{n=1}^{\infty} |x_n|^2\right)^{\frac{1}{2}} + \left(\sum_{n=1}^{\infty} |y_n|^2\right)^{\frac{1}{2}}.
$$

Thus we have, for all $x, y \in \ell^2$,

$$
\|x+y\|_2 \leq \|x\|_2 + \|y\|_2.
$$

Therefore, $(\ell^2, \|\cdot\|_2)$ is a normed space. □

Example 2.7. $(C[0,1], \|\cdot\|_\infty)$ is a normed space, where

$$
C[0,1] = \{f : [0,1] \to \mathbb{C} : f \text{ is continuous on } [0,1]\}
$$

and

$$
\|f\|_\infty = \sup_{0 \leq t \leq 1} |f(t)|.
$$

Proof. We showed before that $C[0,1]$ is a complex vector space. Thus we have to verify that $\|\cdot\|_\infty$ is a norm on $C[0,1]$, that is, we have to show that conditions N1, N2, N3 are satisfied.

N1. If $f \neq 0$, there is $t_0 \in [0,1]$ such that $f(t_0) \neq 0$. Thus

$$
\|f\|_\infty = \sup_{0 \leq t \leq 1} |f(t)| \geq |f(t_0)| > 0.
$$

Therefore $\|f\|_\infty > 0$ for all $f \in C[0,1] \setminus \{0\}$.

N2. Let $\lambda \in \mathbb{C}$ and $f \in C[0,1]$, then

$$\begin{aligned}
\|\lambda f\|_\infty &= \sup_{0\le t\le 1} |\lambda f(t)| \\
&= \sup_{0\le t\le 1} |\lambda||f(t)| \\
&= |\lambda| \sup_{0\le t\le 1} |f(t)| \\
&= |\lambda|\|f\|_\infty.
\end{aligned}$$

N3. Let $f, g \in C[0,1]$. It is easy to see that for every $t \in [0,1]$,

$$\begin{aligned}
|f(t)+g(t)| &\le |f(t)| + |g(t)| \\
&\le \sup_{0\le t\le 1} |f(t)| + \sup_{0\le t\le 1} |g(t)| \\
&= \|f\|_\infty + \|g\|_\infty.
\end{aligned}$$

Thus

$$\sup_{0\le t\le 1} |f(t)+g(t)| \le \|f\|_\infty + \|g\|_\infty,$$

hence

$$\|f+g\|_\infty \le \|f\|_\infty + \|g\|_\infty$$

for all $f, g \in C[0,1]$. $\square$

3 Inner Product Spaces

The main examples to keep in mind are $\mathbb{C}^n$ and ℓ^2.

Example 3.1. $\mathbb{C}^n = \{(x_1, \ldots, x_n) : x_i \in \mathbb{C} \; \forall i = 1, \ldots, n\}$. It has an inner product

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i \overline{y_i}$$

for $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$. Thus the mapping

$$\begin{aligned} \langle \cdot, \cdot \rangle &: \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C} \\ (x, y) &\mapsto \sum_{i=1}^{n} x_i \overline{y_i} \end{aligned}$$

is defined. This mapping has the following properties:

(i) For all $x, y \in \mathbb{C}^n$

$$\begin{aligned} \langle x, y \rangle &= \sum_{i=1}^{n} x_i \overline{y_i} \\ &= \sum_{i=1}^{n} \overline{y_i \overline{x_i}} \\ &= \left(\sum_{i=1}^{n} y_i \overline{x_i} \right)^{-} \\ &= \overline{\langle y, x \rangle}. \end{aligned}$$

(ii) For all $\lambda \in \mathbb{C}$ and $x, y \in \mathbb{C}^n$,

$$\begin{aligned} \langle \lambda x, y \rangle &= \sum_{i=1}^{n} \lambda x_i \overline{y_i} \\ &= \lambda \sum_{i=1}^{n} x_i \overline{y_i} \\ &= \lambda \langle x, y \rangle. \end{aligned}$$

(iii) For all $x, y, z \in \mathbb{C}^n$,

$$\begin{aligned} \langle x + y, z \rangle &= \sum_{i=1}^{n} (x_i + y_i) \overline{z_i} \\ &= \sum_{i=1}^{n} x_i \overline{z_i} + \sum_{i=1}^{n} y_i \overline{z_i} \\ &= \langle x, z \rangle + \langle y, z \rangle. \end{aligned}$$

(iv) For all $x \in \mathbb{C}^n \setminus \{0\}$,

$$\begin{aligned}\langle x, x\rangle &= \sum_{i=1}^{n} x_i \overline{x_i} \\ &= \sum_{i=1}^{n} |x_i|^2 > 0.\end{aligned}$$

Thus $\mathbb{C}^n$ with an inner product $\langle x, y\rangle = \sum_{i=1}^{n} x_i \overline{y_i}$ is an *inner product space.* Let us say precisely what this means.

Definition 3.2. An *inner product* (or *scalar product*) on a complex vector space V is a mapping

$$\begin{aligned}\langle \cdot, \cdot\rangle &: V \times V \to \mathbb{C} \\ (x, y) &\mapsto \langle x, y\rangle\end{aligned}$$

satisfying, for all $x, y, z \in V$ and $\lambda \in \mathbb{C}$, the following conditions:

(i) $\langle x, y\rangle = \overline{\langle y, x\rangle}$,

(ii) $\langle \lambda x, y\rangle = \lambda \langle x, y\rangle$,

(iii) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$,

(iv) $\langle x, x\rangle > 0$ whenever $x \neq 0$.

Remark 3.3. By conditions (ii) and (iii), an inner product is *linear* in the first argument.

Definition 3.4. An *inner product space* (or *pre-Hilbert space*) is a pair $(V, \langle \cdot, \cdot\rangle)$ where V is a complex vector space and $\langle \cdot, \cdot\rangle$ is an inner product on V.

We would like to find an infinite dimensional version of $\mathbb{C}^n$. One can speak of the complex vector space $\mathbb{C}^{\mathbb{N}} = \{(x_i)_{i=1}^{\infty} : x_i \in \mathbb{C} \ \forall i\}$.

Question 3.5. Does the series $\sum_{i=1}^{\infty} x_i \overline{y_i}$ converge for all x and y from $\mathbb{C}^{\mathbb{N}}$?

Unfortunately, this series does not converge for some x and y from $\mathbb{C}^{\mathbb{N}}$. Thus there is no natural way to extend the all-important notion of inner product to $\mathbb{C}^{\mathbb{N}}$.

Example 3.6. $\ell^2 = \{(x_n)_{n=1}^{\infty} : x_n \in \mathbb{C} \ \ \forall n \in \mathbb{N}$ such that $\sum_{n=1}^{\infty} |x_n|^2 < \infty\}$ with inner product

$$\langle \cdot, \cdot\rangle : \ell^2 \times \ell^2 \to \mathbb{C}$$

given by

$$\langle x, y \rangle = \sum_{n=1}^{\infty} x_n \overline{y_n}$$

is an inner product space.

Proof. In Section 1 we proved Lemma 1.10:

For any $x = (x_n)_{n=1}^{\infty}, y = (y_n)_{n=1}^{\infty} \in \ell^2$, the series $\sum_{n=1}^{\infty} x_n \overline{y_n}$ converges absolutely and

$$\sum_{n=1}^{\infty} |x_n \overline{y_n}| \leq \left\{ \sum_{n=1}^{\infty} |x_n|^2 \right\}^{\frac{1}{2}} \left\{ \sum_{n=1}^{\infty} |y_n|^2 \right\}^{\frac{1}{2}}.$$

Thus the mapping

$$\langle \cdot, \cdot \rangle : \ell^2 \times \ell^2 \to \mathbb{C} \tag{3.1}$$

$$(x, y) \mapsto \sum_{n=1}^{\infty} x_n \overline{y_n} \tag{3.2}$$

is well defined.

2) Similarly to the case of $\mathbb{C}^n$ we can check that the mapping $\langle \cdot, \cdot \rangle : \ell^2 \times \ell^2 \to \mathbb{C}$ satisfies the conditions $(i) - (iv)$ from the definition of the inner product.

[Please check them.] $\square$

Example 3.7. The complex vector space $C[0,1] = \{f : [0,1] \to \mathbb{C} : f \text{ is continuous on } [0,1]\}$ becomes an inner product space when endowed with the inner product

$$\langle \cdot, \cdot \rangle : C[0,1] \times C[0,1] \to \mathbb{C}$$

given by

$$\langle f, g \rangle = \int_0^1 f(t) \overline{g(t)}\, dt.$$

Let us check that the conditions $(i) - (iv)$ of the definition of inner product are satisfied.

(i) For all $f, g \in C[0,1]$,

$$\begin{aligned} \langle f, g \rangle &= \int_0^1 f(t) \overline{g(t)}\, dt \\ &= \int_0^1 \overline{g(t) \overline{f(t)}}\, dt \\ &= \overline{\langle g, f \rangle}. \end{aligned}$$

(ii) For all $f, g \in C[0,1]$ and $\lambda \in \mathbb{C}$,

$$\begin{aligned}\langle \lambda f, g\rangle &= \int_0^1 \lambda f(t)\overline{g(t)}\,dt \\ &= \lambda \int_0^1 f(t)\overline{g(t)}\,dt \quad \text{(by the linearity of integrals)} \\ &= \lambda\langle f, g\rangle.\end{aligned}$$

(iii) For all $f_1, f_2, g \in C[0,1]$,

$$\begin{aligned}\langle f_1 + f_2, g\rangle &= \int_0^1 (f_1(t) + f_2(t))\overline{g(t)}\,dt \\ &= \int_0^1 f_1(t)\overline{g(t)}\,dt + \int_0^1 f_2(t)\overline{g(t)}\,dt \quad \text{(by the linearity of integrals)} \\ &= \langle f_1, g\rangle + \langle f_2, g\rangle.\end{aligned}$$

(iv) When $f \in C[0,1]$ and $f \neq 0$,

$$\begin{aligned}\langle f, f\rangle &= \int_0^1 f(t)\overline{f(t)}\,dt \\ &= \int_0^1 |f(t)|^2\,dt > 0.\end{aligned}$$

Thus $(C[0,1], \langle\cdot,\cdot\rangle)$ is an inner product space.

Exercise 3.8. *Show that the formula*

$$\langle A, B\rangle = \mathrm{trace}(B^*A),$$

where B^* is the adjoint matrix of B, defines an inner product on the complex vector space $M_n(\mathbb{C})$.

Theorem 3.9. *For any x, y, z in an inner product space $(V, \langle\cdot,\cdot\rangle)$ and any $\lambda \in \mathbb{C}$,*

(i) $\langle x, \lambda y\rangle = \overline{\lambda}\langle x, y\rangle$;

(ii) $\langle x, y + z\rangle = \langle x, y\rangle + \langle x, z\rangle$;

(iii) $\langle \overline{0}, x\rangle = \langle x, \overline{0}\rangle = 0$;

(iv) *if $\langle x, z\rangle = \langle y, z\rangle$ for all $z \in V$, then $x = y$.*

Remark 3.10. By (i) and (ii), an inner product is *conjugate linear* in the *second* argument.

Proof of Theorem 3.9. (i) Let $x, y \in V$ and $\lambda \in \mathbb{C}$. Then

$$\begin{aligned}
\langle x, \lambda y\rangle &= \overline{\langle \lambda y, x\rangle} \quad \text{(by } (i) \text{ of the definition of inner product)} \\
&= \{\lambda\langle y, x\rangle\}^{-} \quad \text{(by } (ii) \text{ of the definition of inner product)} \\
&= \overline{\lambda}\ \overline{\langle y, x\rangle} \\
&= \overline{\lambda}\ \langle x, y\rangle \quad \text{(by } (i) \text{ of the definition of inner product).}
\end{aligned}$$

(ii) Let $x, y, z \in V$. Then

$$\begin{aligned}
\langle x, y+z\rangle &= \overline{\langle y+z, x\rangle} \quad \text{(by } (i) \text{ of the definition of inner product)} \\
&= \{\langle y, x\rangle + \langle z, x\rangle\}^{-} \quad \text{(by } (iii) \text{ of the definition of inner product)} \\
&= \overline{\langle y, x\rangle} + \overline{\langle z, x\rangle} \\
&= \langle x, y\rangle + \langle x, z\rangle \quad \text{(by } (i) \text{ of the definition of inner product).}
\end{aligned}$$

(iii) For all $x \in V$,

$$\begin{aligned}
\langle \underline{0}, x\rangle &= \langle 0 \cdot \underline{0}, x\rangle \\
&= 0 \cdot \langle \underline{0}, x\rangle \quad \text{by } (ii) \text{ of the definition of inner product} \\
&= 0.
\end{aligned}$$

(iv) Let $\langle x, z\rangle = \langle y, z\rangle$ for all $z \in V$. Then, by the linearity of the inner product in the first argument, for all $z \in V$,

$$0 = \langle x, z\rangle - \langle y, z\rangle = \langle x-y, z\rangle.$$

In particular, for $z = x - y$, we have $\langle x-y, x-y\rangle = 0$. Therefore, it follows from (iv) of the definition of inner product that $x - y = 0$; so $x = y$. $\square$

3.1 Inner Product Spaces as Normed Spaces

In the familiar case of $\mathbb{R}^3$, the magnitude $\|\underline{u}\|_2$ of a vector $\underline{u} \in \mathbb{R}^3$ is equal to $\langle \underline{u}, \underline{u}\rangle^{\frac{1}{2}}$:

$$\|\underline{u}\|_2 = (x^2 + y^2 + z^2)^{\frac{1}{2}} = \langle \underline{u}, \underline{u}\rangle^{\frac{1}{2}},$$

where $\underline{u} = (x, y, z)$. The Euclidean distance between points with position vectors $\underline{u}, \underline{v}$ is $\|\underline{u}-\underline{v}\|_2$. We can copy this to introduce a *natural norm* in any inner product space.

Definition 3.11. The *norm of a vector* x in an inner product space $(V, \langle\cdot,\cdot\rangle)$, written $\|x\|_2$, is defined to be $\langle x, x\rangle^{\frac{1}{2}}$.

Remark 3.12. Later we will show that the mapping

$$\begin{aligned} \|\cdot\|_2 &: V \to \mathbb{R} \\ x &\mapsto \|x\|_2 = \langle x, x\rangle^{\frac{1}{2}} \end{aligned}$$

is always a norm.

Example 3.13. $(\ell^2, \langle\cdot,\cdot\rangle)$. We have

$$\|x\|_2 = \langle x, x\rangle^{\frac{1}{2}} = \left(\sum_{i=1}^{\infty} x_i\overline{x_i}\right)^{\frac{1}{2}} = \left(\sum_{i=1}^{\infty} |x_i|^2\right)^{\frac{1}{2}}.$$

Find $\|x_0\|_2$, where $x_0 = (3i, 0, -4, 0, \ldots)$. By the above,

$$\|x_0\|_2 = (|3i|^2 + |-4|^2)^{\frac{1}{2}} = (9+16)^{\frac{1}{2}} = 5.$$

Example 3.14. $(C[0,1], \langle\cdot,\cdot\rangle)$. We have

$$\|f\|_2 = \left(\int_0^1 f(t)\overline{f(t)}\, dt\right)^{\frac{1}{2}} = \left(\int_0^1 |f(t)|^2\, dt\right)^{\frac{1}{2}}.$$

Find $\|f_0\|_2$, where $f_0(t) = 1 - it$. By the above

$$\begin{aligned} \|f_0\|_2 &= \left(\int_0^1 |1-it|^2\, dt\right)^{\frac{1}{2}} \\ &= \left(\int_0^1 (1+t^2)\, dt\right)^{\frac{1}{2}} \\ &= \left(t\Big|_0^1 + \frac{t^3}{3}\Big|_0^1\right)^{\frac{1}{2}} \\ &= \left(1 + \frac{1}{3}\right)^{\frac{1}{2}} = \sqrt{\frac{4}{3}}. \end{aligned}$$

Remark 3.15. Consider $C[0,1]$, the vector space of continuous complex valued functions on $[0,1]$. Note that the normed spaces:

1. $(C[0,1], \|\cdot\|_\infty)$ where $\|f\|_\infty = \sup_{0\le t\le 1} |f(t)|$ and

2. $(C[0,1], \|\cdot\|_2)$ where $\|f\|_2 = (\int_0^1 |f(t)|^2\, dt)^{\frac{1}{2}}$,

have the same underlying vector space, but different norms.

Theorem 3.16. (Cauchy-Schwarz Inequality) *For any x, y in an inner product space $(V, \langle\cdot,\cdot\rangle)$,*

$$|\langle x, y\rangle| \leq \|x\|_2\|y\|_2$$

with equality if and only if x and y are linearly dependent.

Proof. When $y = \lambda x$ for some $\lambda \in \mathbb{C}$, we have

$$\begin{aligned}
|\langle x, y\rangle| &= |\langle x, \lambda x\rangle| \\
&= |\overline{\lambda}\langle x, x\rangle| \\
&= |\lambda|\|x\|_2^2 \\
&= \|\lambda x\|_2\|x\|_2 \\
&= \|y\|_2\|x\|_2.
\end{aligned}$$

Suppose now that x and y are *not linearly dependent.* Then, for all $\lambda \in \mathbb{C}$, $x + \lambda y \neq 0$, so

$$\langle x + \lambda y, x + \lambda y\rangle > 0.$$

Thus, by the definition of inner product, for all $\lambda \in \mathbb{C}$, we have

$$\begin{aligned}
\langle x + \lambda y, x + \lambda y\rangle &= \langle x, x + \lambda y\rangle + \langle \lambda y, x + \lambda y\rangle \\
&= \langle x, x\rangle + \langle x, \lambda y\rangle + \langle \lambda y, x\rangle + \langle \lambda y, \lambda y\rangle \\
&= \langle x, x\rangle + \overline{\lambda}\langle x, y\rangle + \lambda\langle y, x\rangle + \lambda\overline{\lambda}\langle y, y\rangle.
\end{aligned}$$

Pick $\lambda = -\frac{\langle x,y\rangle}{\langle y,y\rangle}$. Then

$$\langle x, x\rangle - \frac{\langle x, y\rangle}{\langle y, y\rangle}.\langle y, x\rangle > 0,$$

therefore

$$\langle x, x\rangle\langle y, y\rangle > \langle x, y\rangle\langle y, x\rangle.$$

Since $\langle y, x\rangle = \overline{\langle x, y\rangle}$ and $\|x\|_2^2 = \langle x, x\rangle$, we have

$$|\langle x, y\rangle| < \|x\|_2\|y\|_2$$

for all linearly independent x and y. $\square$

Remark 3.17. Recall in $\mathbb{R}^3$

$$\langle \underline{v}, \underline{w}\rangle = \|\underline{v}\|_2\|\underline{w}\|_2 \cos\theta,$$

where θ is the angle between $\underline{v}$ and $\underline{w}$. Thus we can use $\langle\cdot,\cdot\rangle$ to test if vectors are orthogonal.

Inner products are good because they are *linear* in the *first* argument, *conjugate linear* in the *second* argument and related to *norms*, which are not linear.

Example 3.18. Prove that, for any $f \in C[1,2]$,

$$\left| \int_1^2 t\overline{f(t)}\, dt \right| \leq \sqrt{\frac{7}{3}} \left(\int_1^2 |f(t)|^2\, dt \right)^{\frac{1}{2}}.$$

For which f does equality hold?

Proof. By the Cauchy-Schwarz Inequality 3.16, we have

$$\begin{aligned}
\left| \int_1^2 t\overline{f(t)}\, dt \right| &= |\langle g, f\rangle| \quad (\text{here } g(t) = t,\ t \in [1,2]) \\
&\leq \|g\|_2 \|f\|_2 \quad (\text{by the Cauchy-Schwarz inequality}) \\
&= \left(\int_1^2 t^2\, dt \right)^{\frac{1}{2}} \|f\|_2 \\
&= \left(\frac{t^3}{3}\Big|_1^2 \right)^{\frac{1}{2}} \|f\|_2 \\
&= \sqrt{\frac{7}{3}} \left(\int_1^2 |f(t)|^2\, dt \right)^{\frac{1}{2}}.
\end{aligned}$$

The equality holds if and only if $f(t) = \lambda t$ for some $\lambda \in \mathbb{C}$. □

Theorem 3.19. *Let $(V, \langle\cdot,\cdot\rangle)$ be a an inner product space. Then*

(i) $\|x\|_2 \geq 0$ *for all* $x \in V$ *and* $\|x\|_2 = 0$ *implies* $x = 0$*;*

(ii) $\|\lambda x\|_2 = |\lambda|\|x\|_2$ *for all* $x \in V$ *and* $\lambda \in \mathbb{C}$*;*

(iii) $\|x + y\|_2 \leq \|x\|_2 + \|y\|_2$ *for all* $x, y \in V$*.*

Proof. *(i)* $\|x\|_2 = \langle x, x\rangle^{\frac{1}{2}} \geq 0$. By *(iv)* of the definition of inner product, $\|x\|_2 = \langle x, x\rangle^{\frac{1}{2}} > 0$ when $x \neq 0$.

(ii) For all $x \in V$, $\lambda \in \mathbb{C}$,

$$\begin{aligned}
\|\lambda x\|_2 &= \langle \lambda x, \lambda x\rangle^{\frac{1}{2}} \\
&= (\lambda\overline{\lambda}\langle x, x\rangle)^{\frac{1}{2}} \\
&= |\lambda|\|x\|_2.
\end{aligned}$$

(iii) For all $x, y \in V$,

$$\begin{aligned}
\|x+y\|_2^2 &= \langle x+y, x+y\rangle \\
&= \langle x, x+y\rangle + \langle y, x+y\rangle \\
&= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\
&= \|x\|_2^2 + \langle x, y\rangle + \overline{\langle x, y\rangle} + \|y\|_2^2 \\
&= \|x\|_2^2 + 2\mathrm{Re}\langle x, y\rangle + \|y\|_2^2 \\
&\leq \|x\|_2^2 + 2|\langle x, y\rangle| + \|y\|_2^2 \\
&\leq \|x\|_2^2 + 2\|x\|_2\|y\|_2 + \|y\|_2^2 \quad \text{by the Cauchy-Schwarz inequality} \\
&= (\|x\|_2 + \|y\|_2)^2.
\end{aligned}$$

Thus, for all $x, y \in V$,

$$\|x+y\|_2 \leq \|x\|_2 + \|y\|_2.$$

$\square$

Remark 3.20. The theorem shows that $\|x\|_2 = \langle x, x\rangle^{\frac{1}{2}}$ does define a norm in any inner product space. Hence, roughly speaking, every inner product space is a normed space.

Theorem 3.21. (The Parallelogram Law) *Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space. Then, for any $x, y \in V$,*

$$\|x+y\|_2^2 + \|x-y\|_2^2 = 2\|x\|_2^2 + 2\|y\|_2^2.$$

(Exercise)

Example 3.22. Consider the normed space ℓ^1 of complex sequences $x = (x_n)_{n=1}^{\infty}$ such that $\sum_{n=1}^{\infty} |x_n| < \infty$ with norm $\|x\|_1 = \sum_{n=1}^{\infty} |x_n|$. Does the parallelogram law hold in ℓ^1? (Exercise).
Is the norm $\|\cdot\|_1$ induced by some inner product?

Theorem 3.23. (The Polarization Identity) *Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space. Then, for any $x, y \in V$*

$$4\langle x, y\rangle = \|x+y\|_2^2 - \|x-y\|_2^2 + i\|x+iy\|_2^2 - i\|x-iy\|_2^2,$$

here $i \in \mathbb{C}$ is such that $i^2 = -1$.

Proof. (Exercise.) $\square$

Note that the polarization identity can be written

$$\langle x, y\rangle = \frac{1}{4}\sum_{k=0}^{3} i^k \|x + i^k y\|_2^2.$$

Thus, if $\|\cdot\|_2$ is known, then the inner product $\langle\cdot,\cdot\rangle$ can be recovered.

Remark 3.24. Inner product spaces are much more special than normed spaces, and their geometry is much closer to the familiar Euclidean geometry.
As we saw, the parallelogram law does not hold in a general normed space, e.g., $(\ell^1, \|\cdot\|_1)$, $(C[0,1], \|\cdot\|_\infty)$. In fact, it holds *only* in inner product spaces.
If the parallelogram law holds in $(V, \|\cdot\|)$, then the formula

$$\langle x, y\rangle = \frac{1}{4}\sum_{k=0}^{3} i^k \|x + i^k y\|^2$$

defines a genuine inner product on V.

4 Limits

Let $(V, \|\cdot\|)$ be a normed space. The norm $\|\cdot\|$ helps us to measure the distance between two vectors $x, y \in V$ letting $d(x,y) = \|x-y\|$, we obtain a metric on V so that (V, d) becomes a metric space. Verify that all axioms for metric are satisfied (Exercise).

We have the usual metric space definition of the limit now:

Definition 4.1. Let $(V, \|\cdot\|)$ be a normed space. We say a *sequence* $(x_n)_{n=1}^{\infty}$, $x_n \in V$, *converges to* $x \in V$ as $n \to \infty$ if for every $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that for all $n \geq N$, $\|x_n - x\| < \varepsilon$, i.e. the sequence of real numbers $t_n = \|x_n - x\| \to 0$ as $n \to \infty$.

Example 4.2. Take the underlying vector space

$$C[0,1] = \{f : [0,1] \to \mathbb{C} : f \text{ is continuous on } [0,1]\}$$

with two different norms:

$$\begin{aligned} \|f\|_\infty &= \sup_{0 \leq t \leq 1} |f(t)|, \\ \|f\|_2 &= \left(\int_0^1 |f(t)|^2 \, dt \right)^{\frac{1}{2}}. \end{aligned}$$

Consider the sequence of functions f_n given by

$$f_n(t) = \begin{cases} -nt + 1, & t \in [0, \frac{1}{n}], \\ 0, & t \in [\frac{1}{n}, 1]. \end{cases}$$

The sequence converges to $g \equiv 0$ in $(C[0,1], \|\cdot\|_2)$, but not in $(C[0,1], \|\cdot\|_\infty)$.

Proof. Exercise! $\square$

Remark 4.3. The two metrics

$$\begin{aligned} d_\infty(f,g) &= \|f - g\|_\infty, \text{and} \\ d_2(f,g) &= \|f - g\|_2 \end{aligned}$$

are *inequivalent* on $C[0,1]$.

Example 4.4. Let us consider the normed space $(\ell^\infty, \|\cdot\|_\infty)$, where ℓ^∞ is the complex vector space of all bounded sequences $\underline{x} = (x_n)_{n=1}^{\infty}$ of complex numbers, with componentwise addition and scalar multiplication and

$$\|\underline{x}\|_\infty = \sup_{n \in \mathbb{N}} |x_n| \quad \text{(least upper bound)}.$$

1. Consider the sequence of vectors $(\underline{x}^m)_{m=1}^{\infty}$, where

$$\underline{x}^m = (1, \frac{1}{2}, \ldots, \frac{1}{m}, 0, \ldots)$$

and the vector

$$\underline{x} = (1, \frac{1}{2}, \ldots, \frac{1}{n}, \frac{1}{n+1}, \ldots).$$

Then $\underline{x}^m \to \underline{x}$ as $m \to \infty$ in $(\ell^\infty, \|\cdot\|_\infty)$.

Proof.

$$\begin{aligned} \|\underline{x}^m - \underline{x}\|_\infty &= \left\| (\underbrace{0, \ldots, 0}_{\text{m times}}, -\frac{1}{m+1}, -\frac{1}{m+2}, \ldots) \right\|_\infty \\ &= \sup\left\{ \left|\frac{1}{m+1}\right|, \left|\frac{1}{m+2}\right|, \ldots \right\} \\ &= \frac{1}{m+1} \to 0 \quad \text{as } m \to \infty. \end{aligned}$$

□

2. Consider the sequence of vectors $(\underline{y}^m)_{m=1}^{\infty}$, where

$$\underline{y}^m = (\underbrace{1, \ldots, 1}_{\text{m times}}, 0, 0, \ldots)$$

and the vector

$$\underline{y} = (1, 1, \ldots, 1, \ldots).$$

Then $\underline{y}^m \not\to \underline{y}$ as $m \to \infty$ in $(\ell^\infty, \|\cdot\|_\infty)$.

Proof.

$$\begin{aligned} \|\underline{y}^m - \underline{y}\|_\infty &= \|(\underbrace{0, \ldots, 0}_{\text{m times}}, -1, -1, \ldots)\|_\infty \\ &= \sup\{0, 1\} = 1 \not\to 0 \quad \text{as } m \to \infty. \end{aligned}$$

□

Definition 4.5. Let $(V, \|\cdot\|)$ be a normed space and let Y be a subset of V. We say a set Y is *dense* in $(V, \|\cdot\|)$ if for every $v \in V$ and for every $\varepsilon > 0$ there exists $y \in Y$ such that $\|v - y\| < \varepsilon$.

That is we can approximate any vector $v \in V$ by a vector $y \in Y$ as closely as we want.

Example 4.6. $\mathbb{Q}$ is dense in $(\mathbb{R}, |\cdot|)$ via decimals.

Example 4.7. Let us consider the normed space $(c_0, \|\cdot\|_\infty)$, where

$$c_0 = \{\underline{v} = (v_1, \ldots, v_i, \ldots) : v_i \in \mathbb{C}, v_i \to 0 \text{ as } i \to \infty\}$$

and

$$\|\underline{v}\|_\infty = \max_{i\in\mathbb{N}} |v_i|;$$

and c_F, the subset of c_0 consisting of those sequences $\underline{v} = (v_i)_{i=1}^\infty$ having only *finitely* many terms *different* from *zero*. Then c_F is dense in $(c_0, \|\cdot\|_\infty)$.

Proof. Given $\underline{v} \in c_0$ and $\varepsilon > 0$. Since $\underline{v} \in c_0$, we have $\underline{v} = (v_1, \ldots, v_i, \ldots)$ and $v_i \to 0$ as $i \to \infty$. Thus for $\varepsilon > 0$ $\exists N_0 \in \mathbb{N}$ such that $|v_k| < \varepsilon$ for all $k \geq N_0$. Take the element

$$\underline{w} = (v_1, \ldots, v_{N_0}, 0, \ldots) \in c_F.$$

One can see that

$$\begin{aligned}\|\underline{v} - \underline{w}\|_\infty &= \|(0, \ldots, 0, v_{N_0+1}, v_{N_0+2}, \ldots)\|_\infty \\ &= \sup\{|v_{N_0+1}|, |v_{N_0+2}|, \ldots\} < \varepsilon.\end{aligned}$$

Hence c_F is dense in $(c_0, \|\cdot\|_\infty)$. □

Definition 4.8. Let $(V, \|\cdot\|)$ be a normed space and let $\underline{v}_1, \ldots, \underline{v}_k, \cdots \in V$. We say the *series* $\sum_{k=1}^\infty \underline{v}_k$ *converges in* $(V, \|\cdot\|)$ if the sequence of partial sums

$$\underline{S}_n = \sum_{k=1}^n \underline{v}_k$$

converges in $(V, \|\cdot\|)$, that is, if there is $\underline{S} \in V$ such that $\|\underline{S}_n - \underline{S}\| \to 0$ as $n \to \infty$.

Definition 4.9. Let $(V, \|\cdot\|)$ be a normed space and let $\underline{v}_1, \ldots, \underline{v}_k, \cdots \in V$. We say the *series* $\sum_{k=1}^\infty \underline{v}_k$ *converges absolutely* in $(V, \|\cdot\|)$ if the series $\sum_{k=1}^\infty \|\underline{v}_k\|$ converges in $\mathbb{R}$, i.e. $\sum_{k=1}^\infty \|\underline{v}_k\| < \infty$.

Example 4.10. Let us consider the normed space $(\ell^2, \|\cdot\|_2)$, where

$$\ell^2 = \{(x_n)_{n=1}^\infty : x_n \in \mathbb{C}\ \forall n \in \mathbb{N}, \sum_{n=1}^\infty |x_n|^2 < \infty\}$$

and

$$\|\underline{x}\|_2 = (\sum_{k=1}^\infty |x_k|^2)^{\frac{1}{2}}.$$

(1). Fix a vector $\underline{x} \in \ell^2$, $\underline{x} = (x_1, \ldots, x_m, \ldots)$, and for $k \in \mathbb{N}$ consider the vector

$$\underline{v}^k = (0, \ldots, 0, x_k, 0, \ldots).$$

Then $\sum_{k=1}^\infty \underline{v}^k = \underline{x}$ with respect to the norm $\|\cdot\|_2$.

Proof. The partial sums

$$\begin{aligned} S_n &= \sum_{k=1}^{n} \underline{v}^k \\ &= (x_1, 0, 0, \ldots, 0, 0, \ldots) \\ &\quad + (0, x_2, 0, \ldots, 0, 0, \ldots) \\ &\quad + \ldots \\ &\quad + (0, 0, 0, \ldots, x_n, 0, \ldots) \\ &= (x_1, x_2, \ldots, x_n, 0, \ldots), \end{aligned}$$

so

$$\begin{aligned} \|S_n - \underline{x}\|_2^2 &= \|(x_1, x_2, \ldots, x_n, 0, \ldots) - (x_1, x_2, \ldots, x_n, x_{n+1}, \ldots)\|_2^2 \\ &= \|(0, \ldots, 0, -x_{n+1}, -x_{n+2}, \ldots)\|_2^2 \\ &= \sum_{k=n+1}^{\infty} |x_k|^2 \to 0 \text{ as } n \to \infty, \end{aligned}$$

since the series $\sum_{k=1}^{\infty} |x_k|^2 < \infty$. □

(2). Let $\underline{y} = (1, \frac{1}{2}, \ldots, \frac{1}{n}, \ldots)$. The series $\sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$, so $\underline{y} \in \ell^2$. We have proved that the series $\sum_{k=1}^{\infty} \underline{v}^k$, where $\underline{v}^k = (0, \ldots, 0, \frac{1}{k}, 0, \ldots)$ converges in $(\ell^2, \|\cdot\|_2)$ and $\sum_{k=1}^{\infty} \underline{v}^k = \underline{y}$. Show that $\sum_{k=1}^{\infty} \underline{v}^k$ *does not converge absolutely* in $(\ell^2, \|\cdot\|_2)$.

Proof. For $k \in \mathbb{N}$,

$$\begin{aligned} \|\underline{v}^k\|_2 &= \|(0, \ldots, 0, \frac{1}{k}, 0, \ldots)\|_2 \\ &= \left(\left|\frac{1}{k}\right|^2\right)^{\frac{1}{2}} \\ &= \frac{1}{k}. \end{aligned}$$

Thus $\sum_{k=1}^{\infty} \|\underline{v}^k\|_2 = \sum_{k=1}^{\infty} \frac{1}{k}$ does not converge. □

(3). The series $\sum_{k=1}^{\infty} \underline{w}^k$, where $\underline{w}^k = (0, 0, \ldots, 0, \frac{1}{k^2}, 0, \ldots)$, *converges absolutely* in
$(\ell^2, \|\cdot\|_2)$.

Proof.

$$
\begin{aligned}
\|\underline{w}^k\|_2 &= \|(0, 0, \ldots, 0, \frac{1}{k^2}, 0, \ldots)\|_2 \\
&= \left(\left|\frac{1}{k^2}\right|^2\right)^{\frac{1}{2}} \\
&= \frac{1}{k^2}.
\end{aligned}
$$

Thus $\sum_{k=1}^{\infty} \|\underline{w}^k\|_2 = \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty$. $\square$

5 Completeness of Normed Spaces

Definition 5.1. Let $(V, \|\cdot\|)$ be a normed space. A *sequence* $(x_n)_{n=1}^{\infty}$ in V is called a *Cauchy sequence* if, for every $\varepsilon > 0$, there exists $N_0 \in \mathbb{N}$ such that $n, m \geq N_0$ implies $\|x_n - x_m\| < \varepsilon$.

That is all terms are eventually close together. Thus it helps us to check if a sequence converges (possibly in a bigger space) without finding the limit.

Example 5.2. Let us consider the normed space $(c_F, \|\cdot\|_\infty)$, where c_F is the vector space consisting of those sequences $\underline{v} = (v_n)_{n=1}^{\infty}$ having only finitely many terms different from zero and

$$\|\underline{v}\|_\infty = \max_{i \in \mathbb{N}} |v_i|.$$

Consider the sequence $\underline{v}^k = (1, \frac{1}{2}, \ldots, \frac{1}{k}, 0, \ldots)$ in $(c_F, \|\cdot\|_\infty)$. Show that $(\underline{v}^k)_{k=1}^{\infty}$ is a *Cauchy sequence* in $(c_F, \|\cdot\|_\infty)$.

Proof. Given $\varepsilon > 0$, take $N_0 = [\frac{1}{\varepsilon}] + 1$, where $[x]$ is the integer part of x. Then, for all $m, n \geq N_0$,

$$\begin{aligned}
\|\underline{v}^n - \underline{v}^m\|_\infty &= \|(1, \frac{1}{2}, \ldots, \frac{1}{n}, 0, \ldots) - (1, \frac{1}{2}, \ldots, \frac{1}{m}, 0, \ldots)\|_\infty \\
&= \begin{cases} \|(0, \ldots, -\frac{1}{n+1}, \ldots, -\frac{1}{m}, 0, \ldots)\|_\infty & \text{if } m > n, \\ \|(0, \ldots, \frac{1}{m+1}, \ldots, \frac{1}{n}, 0, \ldots)\|_\infty & \text{if } n > m, \\ 0 & \text{if } m = n, \end{cases} \\
&= \max\left\{\frac{1}{n+1}, \frac{1}{m+1}\right\} < \varepsilon.
\end{aligned}$$

□

Remark 5.3. We have shown in Example 4.4 that $\underline{v}^k \to \underline{v} = (1, \frac{1}{2}, \ldots, \frac{1}{n}, \ldots)$ as $k \to \infty$ with respect to the norm $\|\cdot\|_\infty$. Note that $\underline{v} \in c_0$, but $\underline{v} \notin c_F$. Thus the sequence $\underline{v}^k$ is a Cauchy sequence, which does not converge in $(c_F, \|\cdot\|_\infty)$, but converges in the bigger space $(c_0, \|\cdot\|_\infty)$.

Lemma 5.4. *Let $(V, \|\cdot\|)$ be a normed space. Every convergent sequence $(x_n)_{n=1}^{\infty}$ in $(V, \|\cdot\|)$ is a Cauchy sequence.*

Proof. Given $\varepsilon > 0$. Since $x_n \to x$ as $n \to \infty$, for $\frac{\varepsilon}{2}$ there exists $N_0 \in \mathbb{N}$ such that for all $n \geq N_0$, $\|x_n - x\| < \frac{\varepsilon}{2}$. Thus by the triangle inequality

$$\begin{aligned}
\|x_n - x_m\| &= \|x_n - x + x - x_m\| \\
&\leq \|x_n - x\| + \|x - x_m\| \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon
\end{aligned}$$

for all $m, n \geq N_0$. Therefore, for $\varepsilon > 0$, there exists $N_0 \in \mathbb{N}$ such that for all $m, n \geq N_0$, $\|x_n - x_m\| < \varepsilon$. $\square$

Definition 5.5. A normed space $(V, \|\cdot\|)$ is called *complete* if every Cauchy sequence converges to a limit in V.

Definition 5.6. A complete normed space is called a *Banach space.*

Example 5.7. The normed space $(c_F, \|\cdot\|_\infty)$ is *not complete.*

Proof. We have proved that the sequence $\underline{v}^k = (1, \frac{1}{2}, \dots, \frac{1}{k}, 0, 0, \dots)$ in c_F is a Cauchy sequence, which does not converge in $(c_F, \|\cdot\|_\infty)$; see Example 5.2 and Remark 5.3. Therefore $(c_F, \|\cdot\|_\infty)$ is not complete. $\square$

Theorem 5.8. *The normed space* $(\mathbb{R}, |\cdot|)$ *is complete.*

Theorem 5.9. *The normed space* $(\mathbb{C}, |\cdot|)$ *is complete.*

You proved this in Real Analysis course.

Theorem 5.10. *A normed vector space $\mathcal{X}$ is complete if and only if every absolutely convergent series converges in norm.*

Proof. Suppose on one hand that $\mathcal{X}$ is complete and $\sum_n \|x_n\|$ converges. Writing $S_N = \sum_{i=1}^N x_n$, we have

$$\|S_N - S_M\| = \|\sum_{n=N+1}^{M} x_n\| \leq \sum_{n=N+1}^{M} \|x_n\| \leq \sum_{n=N+1} \|x_n\| \to 0$$

as N, $M \to \infty$. Thus $\{S_N\}$ is Cauchy and hence convergent.

On the other hand suppose that every absolutely convergent series converges and that $\{x_n\}$ is Cauchy. Pick for each j an index n_j so that m, $n \geq n_j$ and $\|x_m - x_n\| < 2^{-j}$. Setting $x_0 = 0$, we have

$$x_{n_k} = \sum_{j=1}^{k} (x_{n_j} - x_{n_{j-1}}).$$

By the triangle inequality the series $\sum_j (x_{n_j} - x_{n_{j-1}})$ converges absolutely. Hence x_{n_k} converges to some limit y. To see that $\{x_n\}$ converges to y we note that for $n \geq n_k$ and k large enough we have $\|x_n - x_{n_k}\| \leq 2^{-k}$ and hence

$$\|x_n - y\| \leq \|x_n - x_{n_k}\| + \|x_{n_k} - y\| \leq 2^{-k} + 2^{-k} \leq 2^{-(k-1)}$$

and therefore $x_n \to y$ as $n \to \infty$. $\square$

Example 5.11. Let X be a topological space and let $C_b(X)$ be the space of bounded continuous (complex-valued) functions on X. Define the norm $\|\cdot\|_\infty$ on $C_b(X)$ by

$$\|u\|_\infty = \sup\{|u(x)| : x \in X\}.$$

Suppose that $\sum_n \|f_n\|_\infty < \infty$. Then $f(x) := \sum_{n=1}^\infty f_n(x) \in C_b(X)$ and $\|f\|_\infty \le \sum_{n=1}^\infty \|f_n\|_\infty$ and $\sum_n^\infty f_n$ converges to f in norm (Check!). By the previous theorem we conclude that $C_b(X)$ is complete.

Example 5.12. Let $(X, \mathcal{M}, \mu)$ be a measure space. We claim that $L^1(X, \mathcal{M}, \mu)$ is a complete normed space with respect to the norm $\|f\|_1 = \int |f(x)|d\mu(x)$. (Here we identify functions that are equal a.e.)

In fact, suppose that $\sum_{n=1}^\infty \|f_n\|_1 < \infty$. Then by the Monotone Convergence Theorem,

$$\int \sum_{n=1}^\infty |f_n| d\mu = \sum_{n=1}^\infty \int |f_n| d\mu = \sum_{n=1}^\infty \|f_n\|_1 < \infty$$

so that $g = \sum_{n=1}^\infty |f_n|$ is in L^1 and hence finite for almost all x. For such x we have that $f(x) := \sum_{n=1}^\infty f_n(x)$ is finite. Since by the Monotone Convergence Theorem (MCT)

$$\int |f| d\mu = \int |\sum_{n=1}^\infty f_n| d\mu \le \int \sum_{n=1}^\infty |f_n| d\mu \le (\text{MCT}) \le \sum_{n=1}^\infty \int |f_n| d\mu = \sum_{n=1}^\infty ||f_n||_1 < \infty,$$

we have that f is in L^1. Using the MCT again,

$$\|\sum_1^N f_n - f\|_1 = \int |\sum_{N+1}^\infty f_n| d\mu \le \sum_{N+1}^\infty \int |f_n| d\mu = \sum_{N+1}^\infty \|f_n\|_1 \to 0$$

as $N \to \infty$. Now the previous theorem proves our claim.

6 Hilbert Spaces

The inner product spaces $\mathbb{C}^n$ and ℓ^2 share a further convenient property: they are *complete* with respect to the norm induced by the inner product, i.e. $\|x\| = \langle x, x\rangle^{\frac{1}{2}}$.

Definition 6.1. A *Hilbert space* is an inner product space which is complete with respect to the norm generated by the inner product.

Theorem 6.2. *The inner product space $(\ell^2, \langle\cdot,\cdot\rangle)$ is a Hilbert space.*

Proof. We shall prove the completeness of $(\ell^2, \|\cdot\|_2)$, where

$$\|\underline{x}\|_2 = (\sum_{n=1}^{\infty} |x_n|^2)^{\frac{1}{2}}$$

for $\underline{x} = (x_1, x_2, \ldots, x_n, \ldots)$.

Let $(\underline{x}^k)_{k=1}^{\infty}$ be a Cauchy sequence in $(\ell^2, \|\cdot\|_2)$ with $\underline{x}^k = (x_1^k, x_2^k, \ldots, x_n^k, \ldots)$. We need to show that $(\underline{x}^k)_{k=1}^{\infty}$ converges to a limit in $(\ell^2, \|\cdot\|_2)$. We shall do it in three steps:

1. Find a candidate limit $\underline{a}$;
2. Show that $\underline{a} \in \ell^2$;
3. Show that $\lim_{k\to\infty} \underline{x}^k = \underline{a}$ in $(\ell^2, \|\cdot\|_2)$.

Step 1. Find a candidate limit $\underline{a}$. We have to find $\underline{a} \in \ell^2$ such that

$$\begin{aligned}
&\underline{x}^1 = (x_1^1, x_2^1, \ldots, x_n^1, \ldots)\\
&\underline{x}^2 = (x_1^2, x_2^2, \ldots, x_n^2, \ldots)\\
&\underline{x}^3 = (x_1^3, x_2^3, \ldots, x_n^3, \ldots)\\
&\ldots\\
&\underline{x}^k = (x_1^k, x_2^k, \ldots, x_n^k, \ldots)\\
&\ldots
\end{aligned}$$

converges to

$$\underline{a} = (a_1, a_2, \ldots, a_n, \ldots).$$

Consider a fixed "column" n. The sequence $(x_n^k)_{k=1}^{\infty}$ is a Cauchy sequence in $(\mathbb{C}, |\cdot|)$ since

$$\begin{aligned}
|x_n^k - x_n^l|^2 &\le \sum_{i=1}^{\infty} |x_i^k - x_i^l|^2\\
&= \|\underline{x}^k - \underline{x}^l\|_2^2
\end{aligned}$$

and, by hypothesis, for any $\epsilon > 0$ there exists $N_0 \in \mathbb{N}$ such that $k, l \geq N_0$ implies $\|\underline{x}^k - \underline{x}^l\|_2 < \epsilon$, and this implies that

$$|x_n^k - x_n^l| \leq \|\underline{x}^k - \underline{x}^l\|_2 < \epsilon.$$

Recall that $(\mathbb{C}, |\cdot|)$ is complete. Thus $(x_n^k)_{k=1}^\infty$ converges to a limit $a_n \in \mathbb{C}$ as $k \to \infty$, i.e.

$$\lim_{k\to\infty} x_n^k = a_n.$$

Consider $\underline{a} = (a_1, a_2, \ldots, a_n, \ldots)$ in $\mathbb{C}^{\mathbb{N}}$.

Step 2. The candidate limit $\underline{a}$ belongs to ℓ^2. We shall show that $(\underline{x}^k - \underline{a}) \in \ell^2$ for some k. Since ℓ^2 is closed under subtraction, it follows that

$$\underline{a} = \underbrace{\underline{x}^k}_{\in \ell^2} - \underbrace{(\underline{x}^k - \underline{a})}_{\in \ell^2} \in \ell^2.$$

Let $\epsilon > 0$. By hypothesis, there exists $N_0 \in \mathbb{N}$ such that $k, l \geq N_0$ implies $\|\underline{x}^k - \underline{x}^l\|_2 < \epsilon$. Pick $N \in \mathbb{N}$. We have, for all $k, l \geq N_0$,

$$\begin{aligned} \sum_{i=1}^N |x_i^k - x_i^l|^2 &\leq \sum_{i=1}^\infty |x_i^k - x_i^l|^2 \\ &= \|\underline{x}^k - \underline{x}^l\|_2^2 < \epsilon^2. \end{aligned}$$

Let $l \to \infty$ in the finite sum on the left hand side. By the properties of limits in $\mathbb{C}$, we have

$$\begin{aligned} \lim_{l\to\infty} \sum_{i=1}^N |x_i^k - x_i^l|^2 &= \sum_{i=1}^N |x_i^k - \lim_{l\to\infty} x_i^l|^2 \\ &= \sum_{i=1}^N |x_i^k - a_i|^2 \end{aligned}$$

and

$$\lim_{l\to\infty} \sum_{i=1}^N |x_i^k - x_i^l|^2 \leq \epsilon^2.$$

Thus $\sum_{i=1}^N |x_i^k - a_i|^2 \leq \epsilon^2$. Since this holds for all $N \in \mathbb{N}$, we have, on letting $N \to \infty$,

$$\sum_{i=1}^\infty |x_i^k - a_i|^2 \leq \epsilon^2.$$

Hence $\underline{x}^k - \underline{a} \in \ell^2$ and

$$\|\underline{x}^k - \underline{a}\|_2 = \left(\sum_{i=1}^{\infty} |x_i^k - a_i|^2 \right)^{\frac{1}{2}} \leq \epsilon$$

for all $k \geq N_0$.

Step 3. $\lim_{k\to\infty} x^k = a$. We have shown in Step 2 that, for any $\epsilon > 0$, there exists $N_0 \in \mathbb{N}$ such that $k \geq N_0$ implies $\|\underline{x}^k - \underline{a}\|_2 \leq \epsilon$. Hence $\lim_{k\to\infty} \underline{x}^k = \underline{a}$.

Thus every Cauchy sequence in ℓ^2 converges to a limit in ℓ^2, and so ℓ^2 is complete. $\square$

Example 6.3. Let ℓ^2_F be the subspace of ℓ^2 consisting of those sequences $\underline{x} = (x_n)_{n=1}^{\infty}$ having only *finitely* many *terms different from zero.* Show that ℓ^2_F is not a complete normed space with respect to the norm

$$\|\underline{x}\|_2 = \langle \underline{x}, \underline{x} \rangle^{\frac{1}{2}} = \left(\sum_{n=1}^{\infty} |x_i|^2 \right)^{\frac{1}{2}},$$

therefore ℓ^2_F is not a Hilbert space.

Proof. 1. Consider the sequence $(\underline{x}^k)_{k=1}^{\infty}$ in ℓ^2_F with

$$\begin{aligned} &\underline{x}^1 = (1, 0, \ldots, 0, 0, \ldots) \\ &\ldots \\ &\underline{x}^k = (1, \frac{1}{2}, \ldots, \frac{1}{k}, 0, \ldots) \\ &\ldots. \end{aligned}$$

2. We can see that

$$\begin{aligned} \|\underline{x}^k - \underline{x}^l\|_2 &= \left(\sum_{i=1}^{\infty} |x_i^k - x_i^l|^2 \right)^{\frac{1}{2}} \\ &= \left(\sum_{i=\min\{k,l\}+1}^{\max\{k,l\}} \left(\frac{1}{i}\right)^2 \right)^{\frac{1}{2}} \\ &= |S_{\max\{k,l\}} - S_{\min\{k,l\}}|^{\frac{1}{2}} \to 0 \text{ as } k, l \to \infty, \end{aligned}$$

where $S_N = \sum_{i=1}^{N} \frac{1}{i^2}$, since the series $\sum_{i=1}^{\infty} \frac{1}{i^2}$ converges in $(\mathbb{R}, |\cdot|)$. Thus the sequence $(\underline{x}^k)_{k=1}^{\infty}$ is a *Cauchy sequence* in $(\ell^2_F, \|\cdot\|_2)$.

3. We can also see that, for $\underline{y} = (1, \frac{1}{2}, \ldots, \frac{1}{n}, \ldots) \in \ell^2$,

$$\begin{aligned}\|\underline{x}^k - \underline{y}\|_2 &= \left\|(0, \ldots, 0, -\frac{1}{k+1}, -\frac{1}{k+2}, \ldots)\right\|_2 \\ &= \left(\sum_{i=k+1}^{\infty} \frac{1}{i^2}\right)^{\frac{1}{2}} \to 0 \text{ as } k \to \infty.\end{aligned}$$

Hence $\lim_{k\to\infty} \underline{x}^k = \underline{y}$ in ℓ^2.

4. Note if a sequence in a normed space tends to a limit, then the limit is *unique*. Suppose $(\underline{x}^k)_{k=1}^{\infty}$ converges to $\underline{w} \in \ell^2_F \subset \ell^2$, then in $(\ell^2, \|\cdot\|_2)$, $\lim_{k\to\infty} \underline{x}^k = \underline{y}$ and $\lim_{k\to\infty} \underline{x}^k = \underline{w}$. Thus $\underline{w} = \underline{y}$, but $\underline{y} \notin \ell^2_F$. Hence the sequence $(\underline{x}^k)_{k=1}^{\infty}$ does not converge to a limit in ℓ^2_F. Therefore ℓ^2_F is *not complete*, and so ℓ^2_F is *not a Hilbert space.* $\square$

Remark 6.4. $(\ell^2_F, \langle \cdot, \cdot \rangle)$ is an *inner product* which *is not a Hilbert space.*

Remark 6.5. Every *Hilbert space* with $\|\cdot\|_2 = \langle \cdot, \cdot \rangle^{\frac{1}{2}}$ is a *Banach space*, i.e., a complete normed space.

Example 6.6. $(C[0,1], \|\cdot\|_\infty)$ is a *Banach space* which is *not* a *Hilbert space*, because it is impossible to define an inner product on $C[0,1]$ which induces the norm $\|\cdot\|_\infty$. Why?

Example 6.7. The complex vector space $C[0,1]$ with the inner product $\langle f, g \rangle = \int_0^1 f(t)\overline{g(t)}\, dt$, is an *inner product space* which is *not complete* with respect to the norm

$$\|f\|_2 = \langle f, f \rangle^{\frac{1}{2}} = \left(\int_0^1 |f(t)|^2\, dt\right)^{\frac{1}{2}},$$

and so is *not a Hilbert space.*

Proof. In Section 3 we proved that $(C[0,1], \langle \cdot, \cdot \rangle)$ is an inner product space. Now we have to show that $(C[0,1], \|\cdot\|_2)$ is not complete.

1) Consider the functions $f_n \in C[0,1]$, where

$$f_n(t) = \begin{cases} 0, & t \in [0, \frac{1}{2} - \frac{1}{n}], \\ n(t - \frac{1}{2}) + 1, & t \in [\frac{1}{2} - \frac{1}{n}, \frac{1}{2}], \\ 1, & t \in [\frac{1}{2}, 1]. \end{cases}$$

Show that the sequence $(f_n)_{n=1}^{\infty}$ is a *Cauchy sequence* in $(C[0,1], \|\cdot\|_2)$ but $\{f_n\}$ has no limit in $C([0,1])$ and complete the proof.

$\square$

7 Closed Linear Subspaces

Definition 7.1. Let $(V, \|\cdot\|)$ be a normed space and let X be a subset of V. We say X is *closed* if, for every convergent sequence of vectors of X, the limit of it is in X.

Example 7.2. 1. $(\mathbb{R}, |\cdot|)$ and $X = (0,1)$. X is *not closed*, for example, $(\frac{1}{n+1})_{n=1}^{\infty} \subset (0,1)$, but
$$\lim_{n\to\infty} \frac{1}{n+1} = 0 \notin (0,1).$$

2. $(\mathbb{R}, |\cdot|)$ and $X = [0,1]$. X is *closed*.

Definition 7.3. Let $(V, \|\cdot\|)$ be a normed space. We say W is a *closed linear subspace* of V if W is a linear subspace and W is closed with respect to $\|\cdot\|$.

Example 7.4. c_F is a linear subspace of c_0, but c_F is *not closed* with respect to the norm $\|\cdot\|_\infty$. We have proved that, for $\underline{v}^k = (1, \ldots, \frac{1}{k}, 0, \ldots)$ from c_F, $\lim_{k\to\infty} \underline{v}^k = \underline{v} = (1, \ldots, \frac{1}{k}, \frac{1}{k+1}, \ldots)$, so $\underline{v} \notin c_F$.

Example 7.5. Show that, $W = \{f \in C[0,1] : f(t_0) = 0\}$, where $t_0 \in [0,1]$, is a closed linear subspace of $C[0,1]$ with respect to the norm
$$\|f\|_\infty = \sup_{t\in[0,1]} |f(t)|.$$

Proof. 1. It is easy to show that W is a linear subspace.

2. Let $(f_n)_{n=1}^\infty$ be a convergent sequence of functions from W, that is, $f_n \in W$ for all $n \in \mathbb{N}$ and there exists $f \in C[0,1]$ such that $\|f_n - f\|_\infty \to 0$ as $n \to \infty$. Note that, for $t_0 \in [0,1]$,
$$\begin{aligned} |f_n(t_0) - f(t_0)| &\le \sup_{t\in[0,1]} |f_n(t) - f(t)| \\ &= \|f_n - f\|_\infty \to 0 \text{ as } n \to \infty. \end{aligned}$$

Thus $\lim_{n\to\infty} f_n(t_0) = f(t_0)$. By assumption, $f_n \in W$, that is, $f_n(t_0) = 0$. Hence $f(t_0) = 0$, and so $f \in W$. Thus W is closed with respect to $\|\cdot\|_\infty$. □

8 Banach Spaces

Recall that a complete normed space is called a *Banach space*.

Examples of Banach Spaces 8.1. 1. $(\mathbb{C}, |\cdot|)$.

2. $(\mathbb{C}^n, \|\cdot\|_2)$, where $\|x\|_2 = (\sum_{i=1}^n |x_i|^2)^{\frac{1}{2}}$ for $x = (x_i)_{i=1}^n$.

3. $(\ell^2, \|\cdot\|_2)$, where $\|x\|_2 = (\sum_{i=1}^\infty |x_i|^2)^{\frac{1}{2}}$ for $x = (x_i)_{i=1}^\infty$.

4. $(\ell^1, \|\cdot\|_1)$, where $\|x\|_1 = \sum_{i=1}^\infty |x_i|$ for $x = (x_i)_{i=1}^\infty$.

5. $(\ell^\infty, \|\cdot\|_\infty)$, where $\|x\|_\infty = \sup_{i\in\mathbb{N}} |x_i|$ for $x = (x_i)_{i=1}^\infty$.

6. $(c_0, \|\cdot\|_\infty)$, where $\|x\|_\infty = \max_{i\in\mathbb{N}} |x_i|$ for $x = (x_i)_{i=1}^\infty$.

7. $(C[0,1], \|\cdot\|_\infty)$, where $\|f\|_\infty = \sup_{t\in[0,1]} |f(t)|$.

Proposition 8.2. *A closed linear subspace Y of a Banach space $(X, \|\cdot\|)$ is a Banach space.*

Proof. Let $(y_n)_{n=1}^\infty$ be a Cauchy sequence in $(Y, \|\cdot\|)$, and so in $(X, \|\cdot\|)$. Since $(X, \|\cdot\|)$ is a Banach space, the sequence $(y_n)_{n=1}^\infty$ coverges to a limit $y \in X$. By the assumption, Y is closed. Therefore the limit of the convergent sequence $(y_n)_{n=1}^\infty$ from Y belongs to Y, that is,

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

Thus $(Y, \|\cdot\|)$ is complete, and $(Y, \|\cdot\|)$ is a Banach space. $\square$

Example 8.3. *The Hilbert Space $L^2(a,b)$.* Let $-\infty \le a < b \le \infty$. $L^2(a,b)$ is the linear space of Lebesgue measurable functions $f : (a,b) \to \mathbb{C}$ which are square-integrable, in the sense that

$$\int_a^b |f(t)|^2 \, dt < \infty \quad \text{(Lebesgue integral)},$$

with pointwise operations and inner product

$$\langle f, g\rangle = \int_a^b f(t)\overline{g(t)} \, dt.$$

Remark 8.4. Functions f and g are regarded as equal in $L^2(a,b)$ if they are equal "almost everywhere", that is, if $\{t : f(t) \neq g(t)\}$ is a null set. Strictly speaking we should define the *elements* of $L^2(a,b)$ to be not functions but *equivalence classes* of functions equal almost everywhere. For example, $f = g$ in $L^2(a,b)$ where $f(t) = 1$ and

$$g(t) = \begin{cases} 0, & t = \frac{1}{n}, n \in \mathbf{N}, \\ 1, & \text{otherwise.} \end{cases}$$

Remark 8.5. For *finite* a and $b \in \mathbb{R}$, the linear space $C[a,b]$ of continuous complex valued functions on $[a,b]$ is a dense subspace of $L^2(a,b)$ with respect to the norm

$$\|f\|_2 = \left\{ \int_a^b |f(t)|^2 \, dt \right\}^{\frac{1}{2}}.$$

Remark 8.6. $L^2(a,b)$ is *complete* with respect to the norm $\|\cdot\|_2$. The norm $\|\cdot\|_2$ is induced by the inner product

$$\langle f, g \rangle = \int_a^b f(t)\overline{g(t)} \, dt,$$

so $L^2(a,b)$ is a *Hilbert space.*

8.1 Completion of a Normed space

If a normed space is not complete one can always imbed it into a complete normed space.

Let $(X, \|\cdot\|_X)$ be a normed space. A Banach space $(\tilde{X}, \|\cdot\|_{\tilde{X}})$ is called a completion of X if

- X is a subspace of $\tilde{X}$
- $\|x\|_X = \|x\|_{\tilde{X}}$ for all $x \in X$;
- X is dense in $\tilde{X}$,

Theorem 8.7. *Every normed linear space has a unique completion (up to isomorphism that fixes points of X).*

We will just give an idea of the proof. It consists of several steps:

Step 1. Define the set $\tilde{X}$ (= a space of equivalence classes of Cauchy sequences in X): we say that two Cauchy sequences $\{x_n\}$ and $\{y_n\}$ are equivalent (write $\{x_n\} \sim \{y_n\}$) if $\|x_n - y_n\| \to 0$ as $n \to \infty$. It is not difficult to check that $\sim$ is an equivalence relation. Therefore the space of all such sequences can be decomposed into equivalence classes. We define $\tilde{X}$ to be the set of all such equivalence classes. Denote by $[\{x_n\}]$ the equivalence class containing $\{x_n\}$.

Step 2. Define a vector space structure on $\tilde{X}$: we let

$$[\{x_n\}] + [\{y_n\}] := [\{x_n + y_n\}] \text{ and } \lambda[\{x_n\}] := [\{\lambda x_n\}].$$

One checks that the operations are well defined, i.e. do not depend on particular choice of sequences in the equivalence classes.

Step 3. Define a norm on $\tilde{X}$: we let

$$\|[\{x_n\}]\|_{\tilde{X}} := \lim \|x_n\|_X.$$

One checks that it is well-defined and satisfies the axioms for norms. Observe that $\|[\{x_n\}]\|_{\tilde{X}} = 0$ iff $x_n \to 0$ as $n \to \infty$, i.e. $\{x_n\}$ is equivalent to the the constant sequence consisting of zero's.

Step 4. X as a subspace of $\tilde{X}$: To each $x \in X$ we associate the equivalence class consisting of Cauchy sequences which converges to x, i.e. the equivalence class of the constant sequence $x^* = \{x, x, \ldots\}$; $x \leftrightarrow [x^*]$. This map is injective and $\|[x^*]\|_{\tilde{X}} = \|x\|_X$.

Step 5. X is dense in $\tilde{X}$: let $\{x_n\}$ be a Cauchy sequence in X. Then given $\epsilon > 0$ there exists $N = N(\epsilon) > 0$ such that $\|x_N - x_n\|_X < \epsilon$ for all $n \geq N$ and hence $\|[\{x_n\}] - [x_N^*]\|_{\tilde{X}} = \lim_n \|x_n - x_N\|_X \leq \epsilon$ giving the density.

Step 6. $\tilde{X}$ is complete: let $\tilde{x}_n = [\{x_k^n\}_k]$, $n = 1, 2, \ldots$ be a Cauchy sequence in $\tilde{X}$. For each n take $y_n \in X$ such that $\|\tilde{x}_n - [y_n^*]\|_{\tilde{X}} < 1/n$ which exists by Step 5. Then one shows that sequence $(y_1, y_2, \ldots)$ is Cauchy and that $\tilde{x}_n$ converges to the equivalence class of this Cauchy sequence.

Step 7. Uniqueness of the completion.

In a similar way one defines a completion of an inner product space to a Hilbert space.

9 L^p-spaces

Let (X, M, μ) be a measure space. For measurable $f : X \to \mathbb{C}$ and $0 < p < \infty$ we define

$$||f||_p = \left(\int |f|^p d\mu \right)^{1/p}$$

(allowing the possibility that $||f||_p = \infty$).

Set

$$L^p(X, M, \mu) = \{f : X \to \mathbb{C} : f \text{ is measurable and } ||f||_p < \infty\}.$$

We denote $L^p(X, M, \mu)$ by $L^p(\mu)$ or $L^p(X)$ or simply L^p when no confusion arise.

If X is non-empty and $M = \mathcal{P}(M)$, μ is the counting measure we denote $L^p(X, M, \mu)$ by $\ell^p(X)$. We denote $\ell^p(\mathbb{N})$ simply by ℓ^p.

- L^p is a vector space with respect to the usual addition of functions and multiplication by scalar. In fact, if f, $g \in L^p$ then

$$|f + g|^p \leq (2\max(|f|, |g|))^p \leq 2^p(|f|^p + |g|^p)$$

and hence

$$\int |f + g|^p d\mu \leq 2^p(\int |f|^p d\mu + \int |g|^p d\mu) < \infty$$

giving $f + g \in L^p$. That $f \in L^p$ implies $\lambda f \in L^p$ for $\lambda \in \mathbb{C}$ is trivial.

- L^p, $1 \leq p < \infty$, is a normed space with respect to $|| \cdot ||_p$ if we identify functions which are equal almost everywhere.

Lemma 9.1. *If a, $b \geq 0$ and $0 < \lambda < 1$ then*

$$a^\lambda b^{1-\lambda} \leq \lambda a + (1 - \lambda)b \tag{9.3}$$

with equality iff $a = b$.

Proof. The inequality clearly holds for $a = 0$ or $b = 0$. Assume $a, b \neq 0$ and consider $g(x) = a^x b^{1-x}$. Then $g : [0, 1] \to \mathbb{R}$ is convex on $[0, 1]$ as $g''(x) = a^x b^{1-x}(\ln a - \ln b)^2 \geq 0$ implying $g(\lambda \cdot 1 + (1 - \lambda) \cdot 0) \leq \lambda g(1) + (1 - \lambda)g(0)$ that is equivalent to

$$a^\lambda b^{1-\lambda} \leq \lambda a + (1 - \lambda)b.$$

The inequality is strict if $a \neq b$. $\square$

Theorem 9.2. (Hölder's Inequality) *Suppose* $1 < p < \infty$ *and* $p^{-1}+q^{-1} = 1$. *If* f, g *are measurable functions on* X *then*

$$||fg||_1 \leq ||f||_p||g||_q. \tag{9.4}$$

In particular, if $f \in L^p$, $g \in L^q$ *then* $fg \in L^1$. *In this case the equality in (9.4) holds iff* $\alpha|f|^p = \beta|g|^q$ *a.e. for some* α, $\beta \geq 0$, $(\alpha, \beta) \neq (0, 0)$.

Proof. If $||f||_p = 0$ or $||g||_q = 0$ the result is trivial.
In inequality (9.3) let

$$a = \frac{|f|^p}{||f||_p^p}, \quad b = \frac{|g|^q}{||g||_q^q} \text{ and } \lambda = \frac{1}{p}.$$

Then

$$\frac{|f||g|}{\|f\|_p\|g\|_q} \leq \frac{1}{p}\frac{|f|^p}{\|f\|_p^p} + \frac{1}{q}\frac{|g|^q}{\|g\|_q^q} \tag{9.5}$$

and hence

$$\frac{1}{\|f\|_p\|g\|_q}\int |f||g|d\mu \leq \frac{1}{p} + \frac{1}{q},$$

giving the first statement. The equality holds if we have the equality in (9.5) which happens iff $a = b$. □

The number q such that $p^{-1} + q^{-1} = 1$ is called the *conjugate exponent to* p.

Theorem 9.3. (Minkowski's Inequality) *If* $1 \leq p < \infty$ *and* f, $g \in L^p$ *then*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Proof. The result is obvious if $p = 1$ or if $\|f + g\|_p = 0$.
Assume $p > 1$ and $\|f + g\|_p \neq 0$. We have

$$|f + g|^p \leq (|f| + |g|)(|f + g|)^{p-1} = |f||f + g|^{p-1} + |g||f + g|^{p-1}.$$

After applying the Hölder inequality we obtain

$$\begin{aligned}
\int |f+g|^p d\mu &\leq \left(\int |f|^p d\mu\right)^{1/p}\left(\int |f+g|^{(p-1)q}d\mu\right)^{1/q} \\
&+ \left(\int |g|^p d\mu\right)^{1/p}\left(\int |f+g|^{(p-1)q}d\mu\right)^{1/q} \\
&= \|f\|_p\left(\int |f+g|^p d\mu\right)^{1/q} + \|q\|_p\left(\int |f+g|^p d\mu\right)^{1/q}
\end{aligned}$$

and

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p.$$

$\square$

Clearly $\|f\|_p \geq 0$ and $\|\lambda f\|_p = |\lambda|\|f\|_p$. Hence together with the Minskowski Inequality this shows that L^p (where we identify functions that are equal almost everywhere: $\|f\|_p = 0$ iff $f = 0$ a.e.) is a normed space.

Remark 9.4. We remark that the triangular inequality fails for $p < 1$. Suppose $a > 0$, $b > 0$ and $0 < p < 1$. Then for $t > 0$ we have $t^{p-1} < (a+t)^{p-1}$, and by integrating from 0 to b we obtain $a^p + b^p > (a+b)^p$. If now E and F are disjoint sets of positive finite measure in X then letting $a = \mu(E)^{1/p}$ and $b = \mu(F)^{1/p}$ we get

$$\begin{aligned} \|\chi_E + \chi_F\|_p^p &= \int (\chi_E(x) + \chi_F(x))^p d\mu = \int_E d\mu + \int_F d\mu = a^p + b^p \\ &> (a+b)^p = (\|\chi_E\|_p + \|\chi_F\|_p)^p. \end{aligned}$$

- L^p, $1 \leq p < \infty$ is a Banach space.

Proof. We must prove completness of the space. By Theorem 14.10 it is enough to see that any absolutely convergent series in L^p is convergent. Let $\{f_n\}_n \subset L^p$ and $\sum_{n=1}^{\infty} \|f_n\|_p =: B < \infty$.
Let $G_n = \sum_{k=1}^{n} |f_k|$ and $G = \sum_{k=1}^{\infty} |f_k|$.
By Minkowski's inequality, $\|G_n\|_p \leq \sum_{k=1}^{n} \|f_k\|_p \leq B$ for any n. Hence by the monotone convergence theorem

$$\int G^p d\mu = \int \lim G_n^p d\mu \leq B^p.$$

Hence $G \in L^p$ and, in particular $G(x)$ is finite for almost all x. This implies that the series $\sum_{k=1}^{\infty} f_k(x)$ converges for such x. Let $F(x) = \sum_{k=1}^{\infty} f_k(x)$. We have $|F(x)| \leq G(x)$ a.e. and hence $F \in L^p$.
To see that F is the limit of $\sum_{k=1}^{\infty} f_k$ in L^p we observe that

$$|F - \sum_{k=1}^{n} f_k|^p \leq \left(\sum_{k=n+1}^{\infty} |f_k| \right)^p \leq G^p \text{ a.e.}$$

By the dominated convergence theorem,

$$\|F - \sum_{k=1}^{n} f_k\|_p^p = \int |F - \sum_{k=1}^{n} f_k|^p d\mu \to 0.$$

$\square$

- Some other properties of L^p.

Proposition 9.5. *For $1 \le p < \infty$, the set of simple functions $f = \sum_{j=1}^n a_j \chi_{E_j}$, where E_j is measurable and $\mu(E_j) < \infty$, $a_j \in \mathbb{C}$, is dense in L^p.*

Proof. Clearly $\sum_{j=1}^n a_j \chi_{E_j} \in L^p$: we have $\int |\chi_{E_j}|^p d\mu = \mu(E_j) < \infty$; as L^p is a vector space $\sum_{j=1}^n a_j \chi_{E_j} \in L^p$.

Fix an arbitrary $f \in L^p$. It is a fundamental fact of Integration theory that one can find a sequence of simple functions $\{f_n\}_n$ such that $f_n \to f$ a.e. and $|f_n| \le |f|$. Then $f_n \in L^p$ and $|f_n - f|^p \le 2^p |f|^p \in L^1$. Using the dominated convergence theorem we have

$$\lim_n \int |f_n - f|^p d\mu = \int \lim_n |f_n - f| d\mu = 0,$$

i.e. $\|f_n - f\|_p \to 0$. □

- Separability of L^p.

Definition 9.6. A measure μ in a measurable space $(X, \mathcal{M})$ is called separable if there exists a countable collection of measurable sets $\mathfrak{A} = \{E_i\}_i$ such that given $\varepsilon > 0$ and $A \in \mathcal{M}$ there exists $E_i \in \mathfrak{A}$ such that $\mu((E_i \setminus A) \cup (A \setminus E_i)) < \varepsilon$.

The Lebesgue measure on $[a, b)$ is separable. For this measure, one can set, e.g.,

$$\mathfrak{A} = \{\cup_{j=1}^n [\alpha_j, \beta_j) : \alpha_j, \beta_j \in \mathbb{Q} \cap [a, b), n \in \mathbb{N}\}.$$

Proposition 9.7. *$L^p(\mu)$, $1 \le p < \infty$, is separable if the measure μ is separable.*

Proof = Exercise.

- The space L^∞.

Let $(X, \mathcal{M}, \mu)$ be a measure space. For measurable function f on X we define

$$\|f\|_\infty = \inf\{a \ge 0 : \mu(\{x : |f(x)| > a\}) = 0\}$$

with the convention $\inf \emptyset = \infty$.

Note that if $\|f\|_\infty < \infty$, $\mu(\{x : |f(x)| > \|f\|_\infty\}) = 0$. $\|f\|_\infty$ is called the *essential supremum* of f.

Observe that if $f = g$ a.e. then $\|f\|_\infty = \|g\|_\infty$. We define

$$L^\infty = L^\infty(X, \mathcal{M}, \mu) = \{f : X \to \mathbb{C} : f \text{ measurable and } \|f\|_\infty < \infty\}$$

identifying functions which are equal a.e. We have the following hold.

Theorem 9.8. *1.* **Hölder's inequality**. $\|fg\|_1 \le \|f\|_1 \|g|_\infty$ $\forall$ *measurable* f, g;

2. **Minkowski's inequality**. $\|f+g\|_\infty \le \|f\|_\infty + \|g\|_\infty$;

3. $\|f_n - f\|_\infty \Rightarrow \exists E \in \mathcal{M}$, $\mu(E^c) = 0$, *such that*

$$\sup_{x \in E} |f_n(x) - f(x)| \to 0 \text{ as } n \to \infty;$$

4. the family of simple functions is dense in L^∞;

5. L^∞ *is a Banach space.*

Proof. (1) follows from $\int |f(x)g(x)| d\mu \le \|g\|_\infty \int |f(x)| d\mu = \|f\|_1 \|g\|_\infty$.

(2) We have $|f(x)+g(x)| \le |f(x)| + |g(x)| \le \|f\|_\infty + \|g\|_\infty$ for all x in some E with $\mu(E^c) = 0$ and hence $\|f+g\|_\infty \le \|f\|_\infty + \|g\|_\infty$.

(3) Let $E = \cap_n \{x : |f_n(x) - f(x)| \le \|f_n - f\|_\infty\}$. Then $E^c = \cup_n \{x : |f_n(x) - f(x)| > \|f_n - f\|_\infty\}$ and hence $\mu(E^c) = 0$. Moreover, for all $x \in E$ we have

$$|f_n(x) - f(x)| \le \|f_n - f\|_\infty \to 0.$$

(4) Use Theorem 2.10 in Folland.

(5) Use arguments similar to ones used to prove the completness of L^p-spaces, $1 \le p < \infty$.

$\square$

If $X = N$ and μ is the counting measure on $\mathcal{M} = \mathcal{P}(X)$ then $L^\infty(X, \mathcal{M}, \mu)$ is denoted by ℓ^∞. We have

$$\ell^\infty = \{x = (x_1, x_2, \dots) : \sup_i |x_i| < \infty\}.$$

- Relations between different L^p spaces.

  1. If $0 < p < q < r \le \infty$ then $L^q \subset L^p + L^r$, i.e.,

  $$\forall f \in L^q \quad \exists g \in L^p, h \in L^r \text{ such that } f = g + h.$$

  2. If $0 < p < q < r \le \infty$ then $L^p \cap L^r \subset L^q$ and

  $$\|f\|_q \le \|f\|_p^\lambda \|f\|_r^{1-\lambda},$$

  where $\lambda \in (0,1)$ is such that $q^{-1} = \lambda p^{-1} + (1-\lambda) r^{-1}$.

3. If $0 < p < q \leq \infty$ then $\ell^p(A) \subset \ell^q(A)$ and $\|f\|_q \leq \|f\|_p$ for any $f \in \ell^p(A)$.
4. If $\mu(X) < \infty$, $0 < p < q \leq \infty$ then $L^p(\mu) \supset L^q(\mu)$ and $\|f\|_p \leq \|f\|_q \mu(X)^{(q-p)/pq}$.

Proof. 1. Let $f \in L^q$ and let $E = \{x : |f(x)| > 1\}$. Take $g = f\chi_E$ and $h = f\chi_{E^c}$. Then

$$|g|^p = |f|^p\chi_E \leq |f|^q\chi_E$$

giving $g \in L^p$. Similarly

$$|h|^r = |f|^r\chi_{E^c} \leq |f|^q\chi_{E^c}$$

and therefore $h \in L^r$. The statement is proved since $f = g + h$.

2. Use Hölder's inequality.

3. One proves first that $\ell^p(A) \subset \ell^\infty(A)$ and then use the second statement.

4. If $q = \infty$

$$\|f\|_p^p = \int |f|^p d\mu \leq \int \|f\|_\infty d\mu = \|f\|_\infty \mu(X).$$

If $q < \infty$ the Hölder inequality gives

$$\|f\|_p^p = \int |f|^p \cdot 1 d\mu \leq \left(\int (|f|^p)^{q/p} d\mu\right)^{p/q} \left(\int 1 d\mu\right)^{q-p/q} = \|f\|_q^p \mu(X)^{q-p/q}.$$

□

We remark that if $\mu(X) = \infty$ there is no relation in general between L^p and L^q.

10 Linear Operators

The motive for introducing and studying Banach and Hilbert spaces was to aid the investigation of *linear differential* and *integral equations* arising in Physics. Such equations can be written

$$Ay = w$$

where y, w denote elements of some linear space of functions (e.g. $L^2(a,b)$ or a space of differentiable functions), and A is a linear transformation, e.g. $A = \frac{d}{dx}$ or $A = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$, or $Ax(t) = \int_a^b K(t,s)x(s)\,ds$. In attacking these problems we seek inspiration from *linear algebra*.

Definition 10.1. If E, F are vector spaces over $\mathbb{C}$ a *linear operator* from E to F is a mapping

$$T : E \to F$$

satisfying

$$T(\lambda x + \mu y) = \lambda Tx + \mu Ty$$

for all $x, y \in E$ and $\lambda, \mu \in \mathbb{C}$.

Definition 10.2. If $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ are normed spaces a linear operator $T : E \to F$ is said to be *bounded* if there exists $M \geq 0$ such that

$$\|Tx\|_F \leq M\|x\|_E \quad \text{for all } x \in E.$$

Definition 10.3. If $T : E \to F$ is a bounded linear operator from $(E, \|\cdot\|_E)$ to $(F, \|\cdot\|_F)$ then the *norm* (or *operator norm*) of T is the nonnegative real number

$$\sup\{\|Tx\|_F : x \in E, \|x\|_E \leq 1\}$$

and is denoted by $\|T\|$, so

$$\|T\| = \sup_{\|x\|_E \leq 1} \|Tx\|_F.$$

Lemma 10.4. *Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed spaces and let $T : E \to F$ be a bounded linear operator. Then $\|Tx\|_F \leq \|T\|\|x\|_E$ for all $x \in E$.*

Proof. For $x \in E$,

$$\left\| \frac{x}{\|x\|_E} \right\|_E = \frac{1}{\|x\|_E}\|x\|_E = 1.$$

Since T is a linear operator,

$$T\left(\frac{x}{\|x\|_E}\right) = \frac{1}{\|x\|_E}T(x),$$

and

$$\left\|T\left(\frac{x}{\|x\|_E}\right)\right\|_F = \left\|\frac{1}{\|x\|_E}Tx\right\|_F = \frac{1}{\|x\|_E}\|Tx\|_F.$$

Therefore

$$\frac{1}{\|x\|_E}\|Tx\|_F = \left\|T\left(\frac{x}{\|x\|_E}\right)\right\|_F \leq \|T\|.$$

Thus $\|Tx\|_F \leq \|T\|\|x\|_E$ for all $x \in E$. □

Example 10.5. Any *matrix* $A \in M_{m\times n}(\mathbb{C})$ gives a *linear operator* from $\mathbb{C}^n$ to $\mathbb{C}^m$. For any vector $\underline{v} = (x_1, \ldots, x_n) \in \mathbb{C}^n$,

$$A\underline{v} = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \sum_{j=1}^n a_{2j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix}$$

belongs to $\mathbb{C}^m$. Thus consider the mapping $T : \mathbb{C}^n \to \mathbb{C}^m$ such that $T\underline{v} = A\underline{v}$ for $\underline{v} \in \mathbb{C}^n$, where $A \in M_{m\times n}(\mathbb{C})$, $A = (a_{ij})_{i,j=1}^{m,n}$.

1. Show that T is a *linear operator*.

Proof. For any $\underline{v}_1 = (x_1^1, x_2^1, \ldots, x_n^1)$, $\underline{v}_2 = (x_1^2, x_2^2, \ldots, x_n^2) \in \mathbb{C}^n$ and $\lambda, \mu \in \mathbb{C}$, we have

$$\begin{aligned}
T(\lambda \underline{v}_1 + \mu \underline{v}_2) &= A(\lambda \underline{v}_1 + \mu \underline{v}_2) \\
&= \begin{pmatrix} \sum_{j=1}^n a_{1j}(\lambda x_j^1 + \mu x_j^2) \\ \sum_{j=1}^n a_{2j}(\lambda x_j^1 + \mu x_j^2) \\ \vdots \\ \sum_{j=1}^n a_{mj}(\lambda x_j^1 + \mu x_j^2) \end{pmatrix} \\
&= \begin{pmatrix} \lambda \sum_{j=1}^n a_{1j} x_j^1 + \mu \sum_{j=1}^n a_{1j} x_j^2 \\ \lambda \sum_{j=1}^n a_{2j} x_j^1 + \mu \sum_{j=1}^n a_{2j} x_j^2 \\ \vdots \\ \lambda \sum_{j=1}^n a_{mj} x_j^1 + \mu \sum_{j=1}^n a_{mj} x_j^2 \end{pmatrix} \\
&= \lambda \begin{pmatrix} \sum_{j=1}^n a_{1j} x_j^1 \\ \sum_{j=1}^n a_{2j} x_j^1 \\ \vdots \\ \sum_{j=1}^n a_{mj} x_j^1 \end{pmatrix} + \mu \begin{pmatrix} \sum_{j=1}^n a_{1j} x_j^2 \\ \sum_{j=1}^n a_{2j} x_j^2 \\ \vdots \\ \sum_{j=1}^n a_{mj} x_j^2 \end{pmatrix} \\
&= \lambda A \underline{v}_1 + \mu A \underline{v}_2 \\
&= \lambda T \underline{v}_1 + \mu T \underline{v}_2.
\end{aligned}$$

$\square$

2. Show that

$$T : (\mathbb{C}^n, \|\cdot\|_2) \to (\mathbb{C}^m, \|\cdot\|_2) : \underline{v} \mapsto A\underline{v}$$

is bounded.

Proof.

$$\begin{aligned}
\|T\underline{v}\|_2 &= \left\{ \sum_{k=1}^m |(T\underline{v})_k|^2 \right\}^{\frac{1}{2}} \\
&= \left\{ \sum_{k=1}^m \left| \sum_{j=1}^n a_{kj} x_j \right|^2 \right\}^{\frac{1}{2}}.
\end{aligned}$$

By the Cauchy-Schwarz inequality,

$$\begin{aligned}
\left| \sum_{j=1}^n a_{kj} x_j \right| &\le \left\{ \sum_{j=1}^n |x_j|^2 \right\}^{\frac{1}{2}} \left\{ \sum_{j=1}^n |a_{kj}|^2 \right\}^{\frac{1}{2}} \\
&= \|\underline{v}\|_2 \left\{ \sum_{j=1}^n |a_{kj}|^2 \right\}^{\frac{1}{2}}.
\end{aligned}$$

Thus

$$\begin{aligned}
\|T\underline{v}\|_2 &\leq \left\{ \sum_{k=1}^{m} \|\underline{v}\|_2^2 \sum_{j=1}^{n} |a_{kj}|^2 \right\}^{\frac{1}{2}} \\
&= \left\{ \sum_{k=1}^{m} \sum_{j=1}^{n} |a_{kj}|^2 \right\}^{\frac{1}{2}} \|\underline{v}\|_2.
\end{aligned}$$

Thus T is bounded and

$$\begin{aligned}
\|T\| &= \sup_{\|\underline{v}\|_2 \leq 1} \|T\underline{v}\|_2 \\
&\leq \sup_{\|\underline{v}\|_2 \leq 1} \left\{ \sum_{k=1}^{m} \sum_{j=1}^{n} |a_{kj}|^2 \right\}^{\frac{1}{2}} \|\underline{v}\|_2 \\
&= \left\{ \sum_{k=1}^{m} \sum_{j=1}^{n} |a_{kj}|^2 \right\}^{\frac{1}{2}}.
\end{aligned}$$

$\square$

Exercise 10.6. *What will be the norm of $T \in M_n(\mathbb{C})$? Show that $\|T\| = max\{|\lambda|^{1/2} : \lambda$ is an eigenvalue of $T^*T\}$. Hint: Show that $\|A^*A\| = \|A\|^2$ and that the norm of a selfadjoint matrix is equal to the absolute value of its eigenvalue.*

Example 10.7. Consider the normed space $(C[0,1], \|\cdot\|_\infty)$. Define

$$T : C[0,1] \to C[0,1]$$

by the formula

$$(Tx)(t) = f(t)x(t)$$

where $f \in C[0,1]$ and $t \in [0,1]$. Show that T is a bounded linear operator. Find $\|T\|$.

Proof. 1. *Let us show that T is a linear operator.* For all $x, y \in C[0,1]$, $\lambda, \mu \in \mathbb{C}$,

$$\begin{aligned}
[T(\lambda x + \mu y)](t) &= f(t)(\lambda x + \mu y)(t) \\
&= f(t)(\lambda x(t) + \mu y(t)) \\
&= \lambda f(t)x(t) + \mu f(t)y(t) \\
&= \lambda(Tx)(t) + \mu(Ty)(t) \\
&= (\lambda Tx + \mu Ty)(t), \quad t \in [0,1].
\end{aligned}$$

Thus

$$T(\lambda x + \mu y) = \lambda Tx + \mu Ty,$$

so T is a linear operator.

2. *Let us show that T is bounded.* For all $x \in C[0,1]$,

$$\begin{aligned}\|Tx\|_\infty &= \sup_{t\in[0,1]} |(Tx)(t)| \\ &= \sup_{t\in[0,1]} |f(t)x(t)| \\ &\leq \sup_{t\in[0,1]} |f(t)| \sup_{t\in[0,1]} |x(t)| \\ &= \|f\|_\infty \|x\|_\infty.\end{aligned}$$

Thus T is bounded.

3. *Find $\|T\|$.*

$$\begin{aligned}\|T\| &= \sup_{\|x\|_\infty \leq 1} \|Tx\|_\infty \\ &\leq \sup_{\|x\|_\infty \leq 1} \|f\|_\infty \|x\|_\infty \\ &= \|f\|_\infty.\end{aligned}$$

For $x_0(t) = 1,\ t \in [0,1]$,

$$\|x_0\|_\infty = \sup_{t\in[0,1]} |x_0(t)| = \sup_{t\in[0,1]} 1 = 1,$$

and $(Tx_0)(t) = f(t)x_0(t) = f(t),\ t \in [0,1]$. Hence

$$\|f\|_\infty = \|Tx_0\|_\infty \leq \|T\| \leq \|f\|_\infty.$$

Therefore, $\|T\| = \|f\|_\infty$. □

Definition 10.8. Let $T : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ be a bounded linear operator. A vector $x \in X$ is called a *maximising vector* for T if $\|x\|_X = 1$ and $\|Tx\|_Y = \|T\|$.

Example 10.9. Consider the multiplication operator

$$\begin{aligned}T : (C[0,1], \|\cdot\|_\infty) &\to (C[0,1], \|\cdot\|_\infty) \\ x &\mapsto f \cdot x\end{aligned}$$

where $f \in C[0,1]$. We have proved that, for $x_0(t) = 1, t \in [0,1]$, $\|x_0\|_\infty = 1$ and $\|Tx_0\|_\infty = \|T\| = \|f\|_\infty$. Thus x_0 is a maximising vector for T.

Example 10.10. Define a mapping

$$\begin{aligned} R &: C[0,1] \to \mathbf{C} \\ Rf &= \int_0^1 xf(x)\,dx. \end{aligned}$$

1. *Show that R is a linear operator.*

Proof. For all $f, g \in C[0,1]$ and $\lambda, \mu \in \mathbb{C}$,

$$\begin{aligned} R(\lambda f + \mu g) &= \int_0^1 x(\lambda f + \mu g)(x)\,dx \\ &= \int_0^1 x(\lambda f(x) + \mu g(x))\,dx \\ &= \int_0^1 \lambda x f(x) + \mu x g(x)\,dx \\ &= \lambda \int_0^1 xf(x)\,dx + \mu \int_0^1 xg(x)\,dx \quad \text{by the linearity of integrals} \\ &= \lambda Rf + \mu Rg. \end{aligned}$$

Thus R is a linear operator. □

2. *Consider $C[0,1]$ with the norm $\|f\|_1 = \int_0^1 |f(t)|\,dt$. Show that R is bounded.*

Proof.

$$\begin{aligned} |Rf| &= |\int_0^1 xf(x)\,dx| \\ &\le \int_0^1 |xf(x)|\,dx \\ &\le \int_0^1 |f(x)|\,dx \\ &= \|f\|_1 \end{aligned}$$

for all $f \in C[0,1]$. Thus R is bounded and

$$\|R\| = \sup_{\|f\|_1 \le 1} |Rf| \le \sup_{\|f\|_1 \le 1} \|f\|_1 = 1.$$

□

3. *Show that $\|R\| = 1$.*

Proof. Consider the sequence of continuous functions $f_n(t) = (n+1)t^n$. We can see that

$$\begin{aligned} \|f_n\|_1 &= \int_0^1 |f_n(t)|\, dt \\ &= \int_0^1 (n+1)t^n\, dt \\ &= (n+1)\frac{t^{n+1}}{n+1}\bigg|_0^1 = 1 \end{aligned}$$

and

$$\begin{aligned} |Rf_n| &= \left|\int_0^1 x f_n(x)\, dx\right| \\ &= \left|\int_0^1 x(n+1)x^n\, dx\right| \\ &= \left|(n+1)\int_0^1 x^{n+1}\, dx\right| \\ &= (n+1)\frac{x^{n+2}}{n+2}\bigg|_0^1 \\ &= \frac{n+1}{n+2}. \end{aligned}$$

Therefore

$$\|R\| = \sup_{\|f\|_1 \le 1} |Rf| \ge |Rf_n| = \frac{n+1}{n+2}$$

for all n. Hence

$$\|R\| \ge \lim_{n\to\infty} \frac{n+1}{n+2} = 1.$$

Thus $1 \ge \|R\| \ge 1$, so $\|R\| = 1$. □

Definition 10.11. Let $T : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ be a mapping between normed spaces. We say T is *continuous at* $x \in X$ if for every $\varepsilon > 0$ there exists $\delta > 0$ such that, for every $x' \in X$, $\|x - x'\|_X < \delta$ implies

$$\|Tx - Tx'\|_Y < \varepsilon.$$

The next statement is Proposition 5.2 in Folland's book.

Theorem 10.12. *Let $T : (X, \|\cdot\|_X) \to (Y, \|\cdot\|_Y)$ be a linear operator between normed spaces. Then the following are equivalent:*

(i) T is bounded;

(ii) T is continuous on X, that is, T is continuous at all $x \in X$;

(iii) T is continuous at $\underline{0} \in X$.

Proof. We are going to show that $(i) \Rightarrow (ii) \Rightarrow (iii) \Rightarrow (i)$.

$(ii) \Rightarrow (iii)$. This is trivial, since $\underline{0} \in X$.

$(i) \Rightarrow (ii)$. Suppose that T is bounded and so, for all $x \in X$,

$$\|Tx\|_Y \leq \|T\|\|x\|_X.$$

Given $x \in X$ and $\varepsilon > 0$, set $\delta = \frac{\varepsilon}{\|T\|}$. Then, for all $x' \in X$, $\|x - x'\|_X < \delta$ implies

$$\begin{aligned} \|Tx - Tx'\|_Y &= \|T(x - x')\|_Y \quad && T \text{ is a linear operator} \\ &\leq \|T\|\|x - x'\|_X && T \text{ is a bounded operator} \\ &< \|T\|\delta \\ &= \|T\|\frac{\varepsilon}{\|T\|} = \varepsilon. \end{aligned}$$

Thus T is continuous at all $x \in X$.

$(iii) \Rightarrow (i)$. Suppose that T is continuous at $\underline{0} \in X$. Thus, for $\varepsilon = 1$, there exists $\delta > 0$ such that, for every $x' \in X$, $\|\underline{0} - x'\|_X < \delta$ implies $\|T\underline{0} - Tx'\|_Y < 1$. For any $x \in X$, $x \neq \underline{0}$, set $x' = \frac{\delta}{2}\frac{x}{\|x\|_X} \in X$. Note that

$$\begin{aligned} \|x'\|_X &= \left\| \frac{\delta}{2}\frac{x}{\|x\|_X} \right\|_X \\ &= \frac{\delta}{2}\frac{1}{\|x\|_X}\|x\|_X \quad && \text{by the definition of norm} \\ &= \frac{\delta}{2} < \delta. \end{aligned}$$

Therefore, $\|Tx'\|_Y < 1$, and

$$\begin{aligned} \|Tx'\|_Y &= \left\| T\left(\frac{\delta}{2}\frac{x}{\|x\|_X}\right) \right\|_Y \\ &= \left\| \frac{\delta}{2}\frac{1}{\|x\|_X} Tx \right\|_Y \quad && T \text{ is a linear operator} \\ &= \frac{\delta}{2}\frac{1}{\|x\|_X}\|Tx\|_Y && \text{by the definition of norm.} \end{aligned}$$

Thus

$$\frac{\delta}{2}\frac{1}{\|x\|_X}\|Tx\|_Y < 1,$$

so

$$\|Tx\|_Y \leq \frac{2}{\delta}\|x\|_X$$

for all $x \in X$. Hence T is bounded and $\|T\| \leq \frac{2}{\delta}$. □

Theorem 10.13. *Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed spaces. Then the set*

$$L(X,Y) = \{T : X \to Y : T \textit{ is a bounded linear operator}\}$$

is a normed space with respect to the pointwise operations:

$$(T+S)(x) = Tx + Sx \textit{ and}$$
$$(\lambda T)(x) = \lambda Tx,$$

where $T, S \in L(X,Y), \lambda \in \mathbb{C}, x \in X$, and the operator norm. If $(Y, \|\cdot\|_Y)$ is a Banach space, then so is $L(X,Y)$.

Proof. It is easy to verify that $L(X,Y)$ is a normed space with respect to the operator norm (**Check!**). Let us show that $L(X,Y)$ is complete whenever so is $(Y, \|\cdot\|_Y)$. Take a Cauchy sequence $\{T_n\}$ in $L(X,Y)$. Since for $x \in X$, $\|T_n x - T_m x\| \leq \|T_n - T_m\|\|x\|$, the sequence $\{T_n x\}$ is Cauchy in Y and hence has a limit in Y. We define now $T : X \to Y$ by $Tx = \lim T_n x$, $x \in X$. It is easy to see that T is a linear operator. Let us show that it is bounded. For this we observe first that by the triangle inequality $|\|T_n x\| - \|Tx\|| \leq \|T_n x - Tx\|$ and hence $\|Tx\| = \lim \|T_n x\|$. Applying the triangular inequality to the operator norm we obtain

$$|\|T_n\| - \|T_m\|| \leq \|T_n - T_m\|$$

showing that the sequence of real numbers $\{\|T_n\|\}$ is Cauchy and hence has a limit, K. Therefore, for $x \in X$,

$$\|Tx\| = \lim \|T_n x\| \leq \lim \|T_n\|\|x\| \leq K\|x\|$$

and hence T is bounded.

What is left to show is that $\|T_n - T\| \to 0$, as $n \to \infty$.

As $\{T_n\}$ is Cauchy, given ϵ there exists N such that for any m, $n \geq N$, $\|T_n - T_m\| < \epsilon$ and hence for any $x \in X$ $\|T_n x - T_m x\| \leq \|T_n - T_m\|\|x\| < \epsilon\|x\|$. Letting m go to infinity we obtain $\|T_n x - Tx\| \leq \epsilon\|x\|$ and hence $\|T_n - T\| < \epsilon$, giving the statement.

□

Corollary 10.14. *Let $(X, \|\cdot\|_X)$ be a normed space. Then $L(X, \mathbb{C})$ is a Banach space.*

Proof. It follows from the previous theorem, since $\mathbb{C}$ is a Banach space. □

Remark 10.15. The Banach space $X^* = B(X, \mathbb{C})$ is called the *dual space* of X. The vectors $F \in X^*$ are called *continuous linear functionals*.

11 Dual Spaces

Definition 11.1. A *linear functional* on a complex vector space V is a mapping $F : V \to \mathbb{C}$ which satisfies

$$F(\lambda x + \mu y) = \lambda F(x) + \mu F(y)$$

for all $x, y \in V$ and $\lambda, \mu \in \mathbb{C}$.

Definition 11.2. A linear functional on a normed space $(X, \|\cdot\|_x)$ is called *bounded* if there exists $K > 0$ such that

$$|F(x)| \leq K\|x\|_X$$

for all $x \in X$.

Definition 11.3. A linear functional $F : X \to \mathbb{C}$ on a normed space $(X, \|\cdot\|_x)$ is *continuous at* $x \in X$ if for every $\epsilon > 0$ there exists $\delta > 0$ such that, for every $y \in X$, $\|x - y\|_X < \delta$ implies

$$|F(x) - F(y)| < \epsilon.$$

Theorem 11.4. *Let $(X, \|\cdot\|_X)$ be a normed space and let F be a linear functional on X. Then the following are equivalent:*

(*i*) *F is bounded;*

(*ii*) *F is continuous on X, that is, F is continuous at all $x \in X$;*

(*iii*) *F is continuous at $\underline{0} \in X$.*

Proof. It follows from the similar theorem for linear operators. $\square$

Definition 11.5. The *norm* of a bounded linear functional $F : X \to \mathbb{C}$ on a normed space $(X, \|\cdot\|_X)$ is

$$\|F\| = \sup_{\|x\|_X \leq 1} |F(x)|.$$

Exercise 11.6. *Let $C^1[0,1]$ be the complex vector space of continuously differentiable complex valued functions on $[0,1]$ with the supremum norm $\|\cdot\|_\infty$. Define a mapping*

$$F : C^1[0,1] \to \mathbb{C}$$

by

$$F(f) = f'(1).$$

Show that F is a linear functional. Show that F is unbounded. Is F continuous?

Theorem 11.7. *Let $(V, \langle \cdot, \cdot \rangle)$ be an inner product space. Then, for every $y \in V$, the mapping*

$$\phi_y : V \to \mathbb{C}$$

defined by

$$\phi_y(x) = \langle x, y \rangle, \ x \in V,$$

is a continuous linear functional with respect to the norm $\|x\|_2 = \langle x, x \rangle^{\frac{1}{2}}$. Moreover,

$$\|\phi_y\| = \|y\|_2.$$

Proof. For all $x_1, x_2 \in V$ and $\lambda, \mu \in \mathbb{C}$,

$$\begin{aligned} \phi_y(\lambda x_1 + \mu x_2) &= \langle \lambda x_1 + \mu x_2, y \rangle \\ &= \lambda \langle x_1, y \rangle + \mu \langle x_2, y \rangle \quad \text{by the definition of inner product} \\ &= \lambda \phi_y(x_1) + \mu \phi_y(x_2). \end{aligned}$$

Thus ϕ_y is a linear functional.

(*ii*) By the Cauchy-Schwarz inequality,

$$|\phi_y(x)| = |\langle x, y \rangle| \leq \|x\|_2 \|y\|_2$$

for all $x \in V$. Thus ϕ_y is bounded. This implies that ϕ_y is continuous on V.

(*iii*) By the definition,

$$\begin{aligned} \|\phi_y\| &= \sup_{\|x\|_2 \leq 1} |\langle x, y \rangle| \\ &\leq \sup_{\|x\|_2 \leq 1} \|x\|_2 \|y\|_2 \quad \text{by the Cauchy-Schwarz inequality} \\ &= \|y\|_2. \end{aligned}$$

For $x_0 = \frac{y}{\|y\|_2}$, we have

$$\|x_0\|_2 = \left\| \frac{y}{\|y\|_2} \right\| = \frac{1}{\|y\|_2} \|y\|_2 = 1$$

and

$$\begin{aligned} \phi_y(x_0) &= \langle x_0, y \rangle \\ &= \left\langle \frac{y}{\|y\|_2}, y \right\rangle \\ &= \frac{1}{\|y\|_2} \langle y, y \rangle \\ &= \frac{1}{\|y\|_2} \|y\|_2^2 = \|y\|_2. \end{aligned}$$

Hence

$$\|y\|_2 = |\phi_y(x_0)| \le \|\phi_y\| \le \|y\|_2,$$

so $\|\phi_y\| = \|y\|_2$. $\square$

Example 11.8. Define a mapping

$$T : C[0,1] \to \mathbb{C}$$

by the formula

$$T(f) = 5i \int_0^1 f(t)e^{-2t}\,dt, \quad i^2 = -1.$$

Is T a bounded linear functional with respect to the norm $\|\cdot\|_2$? Recall $\|f\|_2 = \{\int_0^1 |f(t)|^2\,dt\}^{\frac{1}{2}}$. Find $\|T\|$.

Solution. 1) Consider $C[0,1]$ with the inner product

$$\langle f, g\rangle = \int_0^1 f(t)\overline{g(t)}\,dt.$$

one can see that

$$\|f\|_2 = \left\{\int_0^1 |f(t)|^2\,dt\right\}^{\frac{1}{2}} = \langle f, f\rangle^{\frac{1}{2}}.$$

2) For every $f \in C[0,1]$,

$$T(f) = 5i \int_0^1 f(t)e^{-2t}\,dt = \langle f, g\rangle,$$

where $g(t) = -5ie^{-2t}$.

3) By the above theorem, T is a bounded linear functional on $(C[0,1], \langle\cdot,\cdot\rangle)$ and

$\|T\| = \|g\|_2$, where

$$\begin{aligned}
\|g\|_2^2 &= \int_0^1 |g(t)|^2 \, dt \\
&= \int_0^1 |-5ie^{-2t}|^2 \, dt \\
&= \int_0^1 25e^{-4t} \, dt \\
&= 25 \int_0^1 e^{-4t} \, dt \\
&= 25 \times \frac{1}{-4} e^{-4t}|_0^1 \\
&= \frac{25}{4}(e^0 - e^{-4}) \\
&= \frac{25}{4}(1 - e^{-4}).
\end{aligned}$$

Thus $\|T\| = \frac{5}{2}\sqrt{1 - e^{-4}}$. □

The following theorem will be proved later.

Theorem 11.9. (Riesz-Fréchet) *Let H be a Hilbert space and let F be a continuous linear functional on H. There exists a unique $y \in H$ such that*

$$F(x) = \langle x, y \rangle$$

for all $x \in H$. Moreover, $\|F\| = \|y\|_2$.

Remark 11.10. The theorem does not hold for an arbitrary inner product space.

Example 11.11. Let ℓ_F^2 be the linear subspace of ℓ^2 consisting of those sequences having only finitely many terms different from zero. We know that ℓ_F^2 is an inner product space which is not a Hilbert space (see section 6, Hilbert spaces).

Define a mapping

$$F : \ell_F^2 \to \mathbb{C}$$

by

$$F((x_n)_{n=1}^{\infty}) = \sum_{n=1}^{\infty} \frac{1}{n} x_n.$$

1. We can see that $F(x) = \langle x, y \rangle$ where $y = (1, \frac{1}{2}, \ldots, \frac{1}{n}, \ldots) \in \ell^2$; note that $y \notin \ell_F^2$.

2. However, F is a *bounded linear functional* on ℓ^2_F and

$$\|F\| = \sup_{x \in \ell^2_F, \|x\|_2 \le 1} |\langle x, y \rangle| \le \|y\|_2.$$

3. Let us show that F is *not equal* to $\langle \cdot, z \rangle$ for any $z \in \ell^2_F$. Let there be $z \in \ell^2_F$ with $F(x) = \langle x, z \rangle$ for all $x \in \ell^2_F$. Then $\langle x, y \rangle = \langle x, z \rangle$ for all $x \in \ell^2_F$. Therefore $\langle e_n, y \rangle = \langle e_n, z \rangle$ for all $n \in \mathbb{N}$, where

$$e_n = (0, \dots, 1, 0, \dots) \quad [1 \text{ in the } n\text{th place}].$$

Hence $y_n = z_n$ for all $n \in \mathbb{N}$, so $y = z$ and $z \notin \ell^2_F$. Thus F is a continuous linear functional on an inner product space $(\ell^2_F, \langle \cdot, \cdot \rangle)$ such that F is not equal to $\langle \cdot, z \rangle$ for any $z \in \ell^2_F$.

Definition 11.12. Let $(X, \|\cdot\|)$ be a normed space. The Banach space $X^* = L(X, \mathbb{C})$ of all continuous linear functionals on X is called the *dual space of* X.

11.1 Mappings

Definition 11.13. Let X, Y be sets.

(*i*) A mapping $f : X \to Y$ is *injective* if, for any $x_1, x_2 \in X$, $x_1 \neq x_2$ implies $f(x_1) \neq f(x_2)$.

(*ii*) The subset $\operatorname{Im} f = \{f(x) : x \in X\}$ is called the *image* of X under f.

(*iii*) A mapping $f : X \to Y$ is *surjective* if $\operatorname{Im} f = Y$.

(*iv*) A mapping $f : X \to Y$ that is both injective and surjective is called *bijective*.

Definition 11.14. Let E, F be vector spaces, $T : E \to F$ be a linear operator. The *kernel* of T is

$$\operatorname{Ker} T = \{x \in E : Tx = \underline{0}\}.$$

Proposition 11.15. *Let E, F be vector spaces, $T : E \to F$ be a linear operator. Then T is injective if and only if* $\operatorname{Ker} T = \underline{0}$.

Proof. Let $x \in \operatorname{Ker} T$, then

$$Tx = T\underline{0} = \underline{0}.$$

($\Rightarrow$) Suppose T *is injective.* Therefore $x \neq \underline{0}$ implies $Tx \neq T\underline{0}$, so $x \notin \operatorname{Ker} T$, only $\underline{0} \in \operatorname{Ker} T$.

($\Leftarrow$) Let $\operatorname{Ker} T = \underline{0}$. For any x_1, x_2 from E such that $x_1 \neq x_2$, we have

$$x_1 - x_2 \neq \underline{0} \Rightarrow (x_1 - x_2) \notin \operatorname{Ker} T.$$

Thus

$$T(x_1) - T(x_2) = T(x_1 - x_2) \neq \underline{0}.$$

Therefore $T(x_1) \neq T(x_2)$, so T is injective. $\square$

Theorem 11.16. *Let $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ be normed spaces and let $T : E \to F$ be a bounded linear operator. Then* $\operatorname{Ker} T$ *is a closed linear subspace of $(E, \|\cdot\|_E)$.*

Proof. (Exercise.) $\square$

Exercise 11.17. *Let E be a normed space and let $T : E \to \mathbb{C}$ be a linear functional. Show that f is bounded if and only if* $\operatorname{Ker} T$ *is a closed subspace of E.*

Definition 11.18. Let E, F be vector spaces over $\mathbb{C}$. A linear operator $T : E \to F$ which is bijective is called an *isomorphism of vector spaces.*

Definition 11.19. Let $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ be normed spaces. A linear operator $T : E \to F$ is called an *isometry* (or a *norm-preserving mapping*) if, for all $y \in E$,

$$\|T(y)\|_F = \|y\|_E.$$

Remark 11.20. The Riesz-Fréchet theorem shows that, for any Hilbert space H, there is a mapping

$$T : H \to H^*$$

given by $T(y) = \phi_y$ where $\phi_y(x) = \langle x, y \rangle$ which is

(i) bijective;

(ii) norm-preserving, that is, $\|T(y)\| = \|y\|$; and

(iii) conjugate-linear, that is, $T(\lambda y + \mu z) = \overline{\lambda} T y + \overline{\mu} T z$.

Thus a Hilbert space can be identified with its own dual space. Hilbert spaces are sometimes said to be "*self-dual*".

Example 11.21. Let us show that $(\ell^1)^*$ can be identified with ℓ^∞.

Proof. 1. There is a mapping $T : \ell^\infty \to (\ell^1)^*$ given by $T(c) = \phi_c$ where

$$\phi_c(x) = \sum_{n=1}^{\infty} c_n x_n,$$

$c = (c_n)_{n=1}^\infty \in \ell^\infty, x = (x_n)_{n=1}^\infty \in \ell^1$. To say $(c_n)_{n=1}^\infty \in \ell^\infty$ simply means that the c_n are bounded, say, by M_c. For $(x_n)_{n=1}^\infty \in \ell^1$, $\sum_{n=1}^\infty |x_n|$ is finite, and hence $\sum_{n=1}^\infty c_n x_n$ is an absolutely convergent series:

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

Thus ϕ_c is well defined on ℓ^1.

We have to show that, for all $c \in \ell^\infty$, the mapping $\phi_c \in (\ell^1)^*$, that is, ϕ_c is a bounded linear functional on ℓ^1.

(a) For all $x = (x_n)_{n=1}^\infty, y = (y_n)_{n=1}^\infty$ from ℓ^1 and $\lambda, \mu \in \mathbb{C}$,

$$\begin{aligned}
\phi_c(\lambda x + \mu y) &= \sum_{n=1}^{\infty} c_n(\lambda x_n + \mu y_n) \\
&= \lambda \sum_{n=1}^{\infty} c_n x_n + \mu \sum_{n=1}^{\infty} c_n y_n \\
&= \lambda \phi_c(x) + \mu \phi_c(y).
\end{aligned}$$

Thus ϕ_c is a linear functional.

(b) For all $x = (x_n)_{n=1}^\infty \in \ell^1$,

$$\begin{aligned}
|\phi_c(x)| &= \left| \sum_{n=1}^{\infty} c_n x_n \right| \\
&\le \sum_{n=1}^{\infty} |c_n x_n| \\
&\le \sup_{n \in \mathbb{N}} |c_n| \sum_{n=1}^{\infty} |x_n| \\
&= \|c\|_\infty \|x\|_1.
\end{aligned}$$

Thus ϕ_c is bounded. Therefore the mapping

$$T : \ell^\infty \to (\ell^1)^* : c \mapsto \phi_c$$

is well defined.

2. *Let us show that T is linear.* For all $\lambda, \mu \in \mathbb{C}$; $c = (c_n)_{n=1}^\infty, d = (d_n)_{n=1}^\infty \in \ell^\infty$ and $x = (x_n)_{n=1}^\infty \in \ell^1$,

$$\begin{aligned}
[T(\lambda c + \mu d)](x) &= \phi_{\lambda c + \mu d}(x) \\
&= \sum_{n=1}^\infty (\lambda c + \mu d)_n x_n \\
&= \sum_{n=1}^\infty (\lambda c_n + \mu d_n) x_n \\
&= \lambda \sum_{n=1}^\infty c_n x_n + \mu \sum_{n=1}^\infty d_n x_n \\
&= \lambda \phi_c(x) + \mu \phi_d(x) \\
&= [\lambda T(c) + \mu T(d)](x).
\end{aligned}$$

3. *Let us show that T is surjective.* Pick any $g \in (\ell^1)^*$: we must find $c \in \ell^\infty$ such that $T(c) = g$; that means $\phi_c = g$.

(a) Let $c_n = g(e_n)$, where

$$e_n = (0, \ldots, 0, 1, 0, \ldots) \quad (1 \text{ in the } n\text{th position}).$$

This is the only possible candidate for c since $\phi_c(e_n) = c_n$. Note that, for all $n \in \mathbb{N}$,

$$\begin{aligned}
|c_n| &= |g(e_n)| \\
&\leq \|g\| \|e_n\|_1 \quad (\text{as } g \text{ is bounded}) \\
&= \|g\|.
\end{aligned}$$

Thus c is a bounded sequence, that is, $c \in \ell^\infty$, and

$$\|c\|_\infty = \sup_{n \in \mathbb{N}} |c_n| \leq \|g\|.$$

(b) To show that $\phi_c = g$, consider any $x = (x_n)_{n=1}^\infty \in \ell^1$ and write $S^k = \sum_{n=1}^k x_n e_n$. Note that

$$\begin{aligned}
\|x - S^k\|_1 &= \|(0, \ldots, 0, x_{k+1}, x_{k+2}, \ldots)\|_1 \\
&= \sum_{n=k+1}^\infty |x_n| \to 0 \text{ as } k \to \infty.
\end{aligned}$$

Therefore, since g is continuous,

$$\begin{aligned}
g(x) &= g(\lim_{k \to \infty} S^k) \\
&= \lim_{k \to \infty} g(S^k).
\end{aligned}$$

By the linearity of g,

$$
\begin{aligned}
g(S^k) &= g\left(\sum_{n=1}^{k} x_n e_n\right) \\
&= \sum_{n=1}^{k} x_n g(e_n) \\
&= \sum_{n=1}^{k} x_n c_n.
\end{aligned}
$$

Hence, for any $x \in \ell^1$,

$$
g(x) = \sum_{k=1}^{\infty} x_n c_n = \phi_c(x).
$$

4. *Let us show that T is injective.* Suppose $c \neq d, c = (c_n)_{n=1}^{\infty}, d = (d_n)_{n=1}^{\infty} \in \ell^{\infty}$. Thus there exists $n_0 \in \mathbb{N}$ such that

$$
c_{n_0} \neq d_{n_0}.
$$

Therefore,

$$
\phi_c(e_{n_0}) = c_{n_0} \neq d_{n_0} = \phi_d(e_{n_0}),
$$

so $\phi_c \neq \phi_d$.

5. *It remains to show that T is a norm-preserving mapping.* We have already proved that

$$
\|c\|_{\infty} \leq \|g\| = \|\phi_c\|
$$

for all $c \in \ell^{\infty}$, and that, for all $x \in \ell^1$,

$$
|\phi_c(X)| \leq \|c\|_{\infty} \|x\|_1.
$$

This implies

$$
\|\phi_c\| = \sup_{\|x\|_1 \leq 1} |\phi_c(x)| \leq \|c\|_{\infty}.
$$

Thus $\|\phi_c\| = \|c\|_{\infty}$, so T preserves norms:

$$
\|T(c)\|_{(\ell^1)^*} = \|\phi_c\| = \|c\|_{\infty}.
$$

Therefore, T is an *isomorphism* of vector spaces and an *isometry*. □

11.2 The dual of L^p

Let $(X, \mathcal{M}, \mu)$ be a measure space. Consider Banach spaces $L^p(X, \mathcal{M}, \mu)$, $1 \leq p \leq \infty$ which we simply denote by L^p if no confusion arise. Suppose that p and q are conjugate exponents, i.e. $\frac{1}{p} + \frac{1}{q} = 1$, if $1 < p < \infty$ or $(p, q) = (1, \infty)$ or $(\infty, 1)$. The aim of this section is to identify the dual spaces $(L^p)^*$.

For $g \in L^q$ let

$$\phi_g(f) = \int fg d\mu, \quad f \in L^p.$$

It follows from (extended) Hölder's inequality that $\phi_g(f)$ if finite for all $f \in L^p$ and

$$|\phi_g(f)| = |\int fg d\mu| \leq \|f\|_p \|g\|_q, \quad 1 \leq p \leq \infty.$$

Moreover, ϕ_g is linear and hence defines a bounded linear functional on L^p and $\|\phi_g\| \leq \|g\|_p$.

Hence the mapping $T : L^q \to (L^p)^*$, $T(g) = \phi_g$ is a well-defined linear (easy) bounded map. Let us clarify when it is injective, surjective and isometric.

Next statement shows that T is almost always an isometry and hence injective.

Proposition 11.22. *Let $1 \leq q < \infty$. Then $\|\phi_g\| = \|g\|_q$. If the measure μ is semifinite, the result holds even for $q = \infty$.*

Proof. Let $q < \infty$. In order to show the equality for norms it is enough to see that $\|\phi_g\| \geq \|g\|_q$. If $g = 0$ a.e. the statement is trivial. Hence assume $\|g\|_q \neq 0$ and consider

$$f = \frac{|g|^{q-1}\overline{\text{sgn}g}}{\|g\|^{q-1}},$$

where $\text{sgn}z = z/|z|$ if $z \neq 0$ and 0 if $z = 0$.

Then for $1 < p < \infty$

$$\|f\|_p^p = \frac{\int |g|^{(q-1)p}}{\|g\|_q^{(q-1)p}} = \frac{\int |g|^q}{\int |g|^q} = 1$$

so

$$\|\phi_g\| \geq \int fg = \frac{\int |g|^q}{\|g\|_q^{q-1}} = \|g\|_q.$$

If $p = \infty$, $q = 1$, $f = \overline{\text{sgn}g}$, $\|f\|_\infty = 1$ and $\|\phi_g\| \geq \int fg = \|g\|_1$.

Let now $q = \infty$ and μ is semifinite, i.e. $\forall A \in \mathcal{M}$, $\mu(A) = \infty$, $\exists B \subset A$, $B \in \mathcal{M}$, such that $0 < \mu(B) < \infty$. Then for $\epsilon > 0$ and $A = \{x : |g(x)| > \|g\|_\infty - \epsilon\}$, we have $\mu(A) > 0$. Let $B \subset A$ such that $0 < \mu(B) < \infty$ and $f = \mu(B)^{-1}\chi_B\overline{\text{sgn}g}$. Then $\|f\|_1 = 1$ and

$$\|\phi_g\| \geq \int fg = \frac{1}{\mu(B)} \int_B |g| \geq \|g\|_\infty - \epsilon.$$

Since ϵ is arbitrary, $\|\phi_g\| \geq \|g\|_\infty$. $\square$

Next theorem shows that in "almost all" cases T is surjective.

Theorem 11.23. *Let μ be σ-finite and $1 \leq p < \infty$. Then for any $\phi \in (L^p)^*$ there exists $g \in L^q$ such that $\phi = \phi_g$ and hence L^q is isometrically isomorphic to $(L^p)^*$.*

Remark 11.24. If $1 < p < \infty$ then the result holds without restriction on μ (see proof in Folland).

Proof. Fix $\phi \in (L^p)^*$ and assume first that μ is finite. Then $\chi_E \in L^p$ for all measurable sets E. Define

$$\nu(E) = \phi(\chi_E), \quad E \in \mathcal{M}.$$

Then ν is a complex measure absolutely continuous with respect to μ. In fact,

- ν is σ-additive: For any disjoint sequence of measurable subsets $\{E_j\}$ and $E = \cup_1^\infty E_j$ we have $\chi_E = \sum_1^\infty \chi_{E_j}$, where the series converges in the L^p-norm:

  $$\|\chi_E - \sum_1^N \chi_{E_i}\| = \| \sum_{N+1}^\infty \chi_{E_i}\| = \mu(\cup_{N+1}^\infty E_i)^{1/p} \to 0, \text{ as } N \to \infty.$$

  Since ϕ is continuous,

  $$\nu(E) = \phi(\chi_E) = \sum_1^\infty \phi(\chi_{E_i}) = \sum_1^\infty \nu(E_i).$$

- $\nu \ll \mu$: If $\mu(A) = 0$ then $\chi_A = 0$ a.e. and hence $\nu(A) = \phi(\chi_A) = 0$.

By the Radon-Nikodym theorem there exists $g \in L^1(\mu)$ such that for any $E \in \mathcal{M}$,

$$\phi(\chi_E) = \nu(E) = \int_E g d\mu = \int \chi_E g d\mu$$

and hence $\phi(f) = \phi_g(f) = \int fg$ for any simple function f so that we have the desired representation on the set of simple functions.

Note however that we have not shown yet that $g \in L^q$, only that it is in L^1. We postpone the proof of this statement until the end of the proof.

Once we know this, we can prove the equality $\phi(f) = \phi_g(f)$ for any $f \in L^p$. In fact, since simple functions are dense in L^p, given $f \in L^p$, there exists a sequence of simple functions f_n such that $\|f_n - f\|_p \to 0$. As ϕ is continuous, $\phi(f_n) \to \phi(f)$. On the other hand, $\phi(f_n) = \int f_n g \to \int fg$: by Hölder's inequality,

$$\|\phi(f_n) - \int fg\| \leq |\int (f_n - f)g| \leq \|f_n - f\|_p \|g\|_q \to 0.$$

Hence $\phi(f) = \int fg$, $f \in L^p$.

Assume now that μ is σ-finite. Then there exists an increasing sequence of sets $\{E_n\}$ such that $0 < \mu(E_n) < \infty$ and $X = \cup_1^\infty E_n$. In what follows we will identify $L^p(E_n)$, $L^q(E_n)$ with the corresponding subspaces of $L^p(X)$ and $L^q(X)$ respectively. By the preceding arguments there exists $g_n \in L^q(E_n)$ such that $\phi(f) = \int fg_n$ for any $f \in L^p(E_n)$ and

$$\|g_n\|_q = \|\phi|_{L^p(E_n)}\| \leq \|\phi\|,$$

where $\phi|_{L^p(E_n)}$ is the restriction of the functional ϕ to the subspace $L^p(E_n)$. Clearly that g_n is unique modulo alternation on null set and hence $g_n = g_m$ a.e. on E_n for all $m \geq n$. Hence we can define g on the whole space X by letting $g = g_n$ on E_n. By the monotone convergence theorem we obtain

$$\|g\|_q^q = \lim_{n\to\infty} \|g_n\|_q^q \leq \|\phi\|^q$$

and therefore $g \in L^q$.

Now let $f \in L^p$ Then $f\chi_{E_n} \in L^p(E_n)$ and by the dominated convergence theorem $f\chi_{E_n} \to f$ in L^p. Hence

$$\phi(f) = \lim_{n\to\infty} \phi(f\chi_{E_n}) = \lim_{n\to\infty} \int fg_n = \lim_{n\to\infty} \int_{E_n} fg = \int fg$$

as desired.

Now let us fill the gap in the proof and show that $g \in L^q$. We may assume that $g \neq 0$ a.e. Observe that we need the statement only in the case when μ is a finite measure. A more general statement is proved in Folland. As $\phi(f) = \int fg$ for simple f and ϕ is bounded, we have that

$$M_q(g) = \sup\{|\int fg| : f \text{ is simple }, \|f\|_p = 1\}$$

is finite: $M_q(g) \leq \|\phi\| < \infty$. Moreover, since for any bounded measurable function f, $\|f\|_p = 1$, there is a sequence $\{f_n\}$ of simple functions such that $|f_n| \leq |f|$ and $f_n \to f$ pointwise, we have, by the dominated convergence theorem,

$$|\int fg| = \lim_{n\to\infty} |\int f_n g| \leq M_q(g) \tag{11.6}$$

Assume $q < \infty$. Let $\{g_n\}$ be a sequence of simple bounded functions such that $g_n \to g$ pointwise and $|g_n| \leq |g_{n+1}| \leq |g|$, it exists due to Theorem 2.10 in Folland. Let

$$f_n = \frac{|g_n|^{q-1}\overline{\mathrm{sgn} g}}{\|g_n\|_q^{q-1}}.$$

Then as in the proof of the previous theorem we have $\|f_n\|_p = 1$ and $\int |f_n g_n| = \|g_n\|_q$. By the monotone convergence theorem we have

$$\begin{aligned}\|g\|_q &= \lim_{n\to\infty} \|g_n\|_q = \lim_{n\to\infty} \int |f_n g_n| \leq \liminf \int |f_n||g| = (\text{as } f_n g \geq 0) \\ &= \liminf \int f_n g \leq (\text{by } 11.6) \leq M_q(g).\end{aligned}$$

Hence $g \in L^q$.

Assume now $q = \infty$ (and hence $p = 1$). In order to prove that $g \in L^\infty$ it is enough to see that $\|g\|_\infty \leq M_\infty(g)$. Assume contrary that for some $\epsilon > 0$ the set $A = \{x : |g(x)| > M_\infty(g) + \epsilon\}$ is of non-zero measure. Setting $f = \chi_A \overline{\mathrm{sgn} g}/\mu(A)$ we have $\|f\|_1 = 1$ and

$$\int fg = \int_A \frac{|g|}{\mu(A)} > M_\infty(g) + \epsilon$$

which is impossible by (11.6).

The proof is complete.

□

Let us summarise:

- $1 < p < \infty$: $(L^p)^* = L^q$ and $(L^p)^{**} = L^p$.

  We proved this for σ-finite measure, but the result holds for general μ.

- $p = 1$: $(L^1)^* = L^\infty$ if the measure μ is σ-finite. The inclusion $T : L^\infty \to (L^1)^*$ is isometric if μ is semi-finite.

  If μ is not semi-finite the injectivity of $T : L^\infty \to (L^1)^*, g \mapsto \phi_g$ fails (see explnation in Folland).

- $p = \infty$: $L_1 \subset (L^\infty)^*$, the mapping $T : L^1 \to (L^\infty)^*, g \mapsto \phi_g$ is an injective isometry but almost never surjective. We shall say more about this later.

12 The Hahn-Banach Theorem and consequences

Let X be a normed space over $\mathbb{R}$ or $\mathbb{C}$. Let X^* be the dual space of X. How can we be sure that there are enough elements in X^* so that the study of the dual space becomes interesting? Can we extend a linear bounded functional given on a subspace to a linear bounded functional on the whole space? The answers to these questions gives one of the most important results in functional analysis, the Hahn-Banach Theorem.

Assume first that X is a real vector space.

Definition 12.1. A sublinear functional on X is a map $p : X \to \mathbb{R}$ such that

$$p(x+y) \leq p(x) + p(y) \text{ and } p(\lambda x) = \lambda p(x) \quad \forall x, y \in X, \lambda \geq 0.$$

Example 12.2.

1. $p(x) = |f(x)|$, where f is a real linear functional;
2. $p(x) = \|x\|$, where $\|\cdot\|$ is a norm or seminorm.

Theorem 12.3. *(**The Hahn-Banach Theorem, real version**) Let X be a real vector space and p is a sublinear functional. Let M be a linear subspace of X and $f : M \to \mathbb{R}$ a linear functional such that*

$$f(x) \leq p(x), x \in M.$$

Then there exists a linear functional $F : X \to \mathbb{R}$ such that $F = f$ on M and $F(x) \leq p(x)$ for all $x \in X$.

Proof. 1. The result is trivial if $M = X$. Assume $M \neq X$ and let $x \in X \setminus M$. Consider the set $M' = M + \mathbb{R}x$, the subspace spanned by x and vectors in M. It is easy to see that each element in M' an be uniquely represented in the form $y + \lambda x$, $y \in M$, $\lambda \in \mathbb{R}$. If g is a linear extension of f to M' then

$$g(y + \lambda x) = f(y) + \lambda g(x), y \in M.$$

If $\alpha = g(x)$ then $g(y + \lambda x) = f(y) + \lambda\alpha$.

The aim is to find α such that $g(z) \leq p(z)$ for all $z \in M'$, i.e.

$$f(y) + \lambda\alpha \leq p(\lambda x + y). \tag{12.7}$$

For $\lambda > 0$, (12.7) is equivalent to $f\left(\frac{y}{\lambda}\right) + \alpha \leq p\left(x + \frac{y}{\lambda}\right)$ or

$$\alpha \leq p\left(x + \frac{y}{\lambda}\right) - f\left(\frac{y}{\lambda}\right) \tag{12.8}$$

and for $\lambda < 0$, to the condition $f\left(\frac{y}{\lambda}\right) + \alpha \geq -p\left(-x - \frac{y}{\lambda}\right)$ or

$$\alpha \geq -p\left(-x - \frac{y}{\lambda}\right) - f\left(\frac{y}{\lambda}\right). \tag{12.9}$$

Let us show that there exists α that satisfies both conditions (12.8) and (12.9).

Let y_1, $y_2 \in M$. Then since

$$f(y_2) - f(y_1) \leq p(y_2 - y_1) = p((y_2 + x) - (y_1 + x)) \leq p(y_2 + x) + p(-y_1 - x)$$

we have

$$-f(y_2) + p(y_2 + x) \geq -f(y_1) - p(-y_1 - x).$$

Since y_1, y_2 were arbitrary, we get

$$\inf_{y \in M}(-f(y) + p(y + x)) \geq \sup_{y \in M}(-f(y) - p(-y - x)).$$

Choosing α between these two quantities and letting

$$g(\lambda x + y) = \lambda\alpha + f(y)$$

we obtain the desired extension of f to the subspace $M + \mathbb{R}x$.

2. If in X one can find a countable set of elements $x_1, x_2, \dots$ that together with M generate the whole space X then the functional on X can be constructed inductively considering the following chain of subspaces:

$$M^{(1)} = \operatorname{span}\{M, x_1\}, M^{(2)} = \operatorname{span}\{M^{(1)}, x_2\}, \dots$$

Then any element in M belongs to some of the subspaces $M^{(k)}$ and the functional will be extended to the whole X.

3. In general case one uses Zorn's lemma. Let $\mathcal{F}$ be the family of linear extensions, F, of f to a subspace M_F with $F(x) \leq p(x)$. It is partially ordered by inclusion: $F_1 \leq F_2$ if $\{(x, F_1(x)) : x \in M_{F_1}\} \subset \{(x, F_2(x) : x \in M_{F_2}\}$. Since the union of increasing family of subspaces of X is a subspace, one has that the union of totally ordered subfamilies of $\mathcal{F}$ lies in $\mathcal{F}$. Hence by Zorn's lemma $\mathcal{F}$ has a maximal element. It is defined on the whole space since otherwise it would admit an extension and, hence, would not be the maximal element of $\mathcal{F}$. $\square$

Remark 12.4. 1. Step 2 is presented here with the main reason to demonstrate that, for separable spaces, the theorem can be proved without using the Zorn lemma.

2. If p is a seminorm/norm then

$$f(x) \leq p(x) \Leftrightarrow |f(x)| \leq p(x).$$

In fact, in this case, $|f(x)| = \pm f(x) = f(\pm x)$ and $f(x) \leq p(x) \Leftarrow -f(x) \leq p(-x) = p(x)$ giving the statement.

Theorem 12.5. *(***The complex Hahn-Banach Theorem***) Let X be a complex vector space, p a seminorm on X, M a subspace of X, and f a complex linear functional on M such that $|f(x)| \leq p(x)$ for $x \in M$. Then there exists a complex linear functional F on X such that $|F(x)| \leq p(x)$ for all $x \in X$ and $F(x) = f(x)$ for all $x \in M$.*

Proof. Consider X and M as vector spaces over reals and denote them by $X_{\mathbb{R}}$ and $M_{\mathbb{R}}$ respectively. Note that $X_{\mathbb{R}}$ ($M_{\mathbb{R}}$) and X (M respectively) are different as linear spaces but coincides as sets. Clearly p is a sublinear functional on $X_{\mathbb{R}}$.

Let $u = \mathrm{Re} f$. Then u is a real linear functional on $M_{\mathbb{R}}$ satisfying the condition

$$|u(x)| \leq p(x)$$

and hence $u(x) \leq p(x)$.

By the real Hahn-Banach theorem there exists a real linear functional U on $X_{\mathbb{R}}$ such that $U(x) \leq p(x)$, for all $x \in X_{\mathbb{R}}$ and $U(x) = u(x)$ for all $x \in M_{\mathbb{R}}$.

Clearly $-U(x) = U(-x) \leq p(-x) = p(x)$ and hence

$$|U(x)| \leq p(x), x \in X_{\mathbb{R}}.$$

Define a functional F on X by letting

$$F(x) = U(x) - iU(ix).$$

F is a complex linear functional: F is real linear since so is U and

$$F(ix) = U(ix) - iU(-x) = U(ix) + iU(x) = i(U(x) - iU(ix)) = iF(x), x \in X.$$

F is an extension of f to X, i.e. $F(x) = f(x)$, $x \in M$: as

$$\mathrm{Im} f(x) = -\mathrm{Re}(if(x)) = -\mathrm{Re}(f(ix)) = -u(ix)$$

we have

$$F(x) = u(x) - iu(ix) = \mathrm{Re} f(x) + i\mathrm{Im} f(x) = f(x).$$

It remains to prove that $|F(x)| \leq p(x)$ for all $x \in X$. Assume contrary that for some $x_0 \in X$, $|F(x_0)| > p(x_0)$. Write $F(x_0) = |F(x_0)|e^{i\varphi}$. Let $y_0 = e^{-i\varphi}x_0$. Then

$$U(y_0) = \mathrm{Re} F(y_0) = \mathrm{Re}[e^{-i\varphi}F(x_0)] = |F(x_0)| > p(x_0) = p(y_0).$$

A contradiction that gives the statement. □

Corollary 12.6. *Let X be a normed space and let M be its subspace. Then for every linear continuous functional f defined on M there exists a functional $F \in X^*$ such that $F(x) = f(x)$ for all $x \in M$ and $\|F\| = \|f\|$.*

Proof. Let $p(x) = \|f\|\|x\|$, $x \in X$. We have $|f(x)| \leq p(x)$ for any $x \in M$. By the complex Hahn-Banach theorem there exists a linear functional F on X such that $|F(x)| \leq p(x) = \|f\|\|x\|$ for all $x \in X$ and $F|_M = f$. Hence, F is bounded and $\|F\| \leq \|f\|$. As $\|F\| = \sup\{|F(x)| : \|x\| \leq 1, x \in X\} \geq \sup\{|F(x)| : \|x\| \leq 1, x \in M\} = \|f\|$, we have $\|F\| = \|f\|$. □

12.1 Corollaries of the Hahn-Banach Theorem

Theorem 12.7. *Let X be a normed space.*

1. *If M is a closed subspace of X then for every vector $x \in X \setminus M$, there exists $f \in X^*$ such that $f(x) \neq 0$ and $f|_M = 0$. In fact f can be taken to satisfy $\|f\| = 1$ and $f(x) = \rho(x, M) := \inf_{y \in M} \|x - y\|$.*

2. *If $x \neq 0$ in X, then there exists $f \in X^*$ such that $\|f\| = 1$ and $f(x) = \|x\|$.*

3. *The bounded linear functionals on X separate points: for x, $y \in X$, $x \neq y$, there exists $f \in X^*$ such that $f(x) \neq f(y)$.*

4. *A subset $M \subset X$ is total in X that is the closed linear span is dense in X if and only if $f \in X^*$, $f(x) = 0$, $x \in M \Rightarrow f = 0$.*

5. *If $x \in X$ define $\hat{x} : X^* \to \mathbb{C}$ by $\hat{x}(f) = f(x)$. Then the map $x \mapsto \hat{x}$ is a linear isometry from X to $X^{**} = (X^*)^*$.*

Proof. 1. We define a functional f on the space $M + \mathbb{C}x$ by

$$f(y + \lambda x) = \lambda \rho(x, M).$$

Then f is linear, $f|_M = 0$ and $f(x) = \rho(x, M)$. Let us find $\|f\|$. We have

$$\begin{aligned} \|f\| &= \sup\{\frac{|f(y + \lambda x))|}{\|y + \lambda x\|} : y \in M, \lambda \in \mathbb{C}\} \\ &= \sup\{\frac{|\lambda|\rho(x, M)}{|\lambda|\|y + \lambda^{-1}x\|} : y \in M, \lambda \in \mathbb{C}, \lambda \neq 0\} \\ &= \frac{\rho(x, M)}{\inf\{\|x + \lambda^{-1}y\|} = \frac{\rho(x, M)}{\rho(x, M)} = 1. \end{aligned}$$

By virtue of the Hahn-Banach theorem, there exists an extension $F \in X^*$ of the functional f with the norm $\|F\| = \|f\| = 1$. Thus the functional F possesses all the required properties.

2. The statement is a special case of the previous one: set $M = \{0\}$.

3. If $x \neq y$ by (2) there exists $f \in X^*$ such that $f(x - y) = \|x - y\| \neq 0$ and hence $f(x) \neq f(y)$.

4. Assume first that a set M is total and $f(x) = 0$ for all $x \in M$. By linearity and continuity we have that $f(x) = 0$ for all x in the closed linear span of M, i.e. for all $x \in X$.

Assume that all functionals $f \in X^*$ vanishing on M are identically equal to zero. Suppose that M is not total and let G be the closed linear span of M. Then there exists $y \in X \setminus G$. By (1) there exists $f \in X^*$ such that $\|f\| = 1$ and $f(x) = 0$ for all $x \in G$, which contradicts the assumption.

5. For each $x \in X$, $\hat{x}$ is a linear functional on X^*:

$$\hat{x}(\lambda f_1 + \mu f_2) = (\lambda f_1 + \mu f_2)(x) = \lambda f_1(x) + \mu f_2(x) = \lambda \hat{x}(f_1) + \mu \hat{x}(f_2).$$

$\hat{x}$ is bounded:

$$|\hat{x}(f)| = |f(x)| \leq \|f\|\|x\|$$

and $\|\hat{x}\|_{X^{**}} \leq \|x\|_X$. On the other hand, since there exists $f \in X^*$, $\|f\| = 1$, $f(x) = \hat{x}(f) = \|x\|$, we have $\|f\| \geq \|x\|_X$. Therefore, $\|\hat{x}\|_{X^{**}} = \|x\|_X$ and $X \ni x \mapsto \hat{x} \in X^{**}$ is a linear isometry.

□

Remark 12.8. The third statement answers an important question about the structure of the dual space of a normed space: it is rich enough to separate the elements of the given space.

If X is a normed space, we let $\hat{X} = \{\hat{x} : x \in X\} \subset X^{**}$ and $i : X \to X^{**}$, $x \mapsto \hat{x}$ the isometric embedding. Then the closure $\overline{\hat{X}}$ is complete as a closed subspace of the complete space X^{**} and $i(X)$ is a dense set in $\overline{\hat{X}}$. Hence $\overline{\hat{X}}$ is the completion of X.

If $\hat{X} = X^{**}$ then X is called **reflexive**, i.e. if the canonical map $i : x \mapsto \hat{x}$ is an isometric isomorphism.

Exercise 12.9.

1. We have proved that $T : l^1 \to (l^\infty)^*$, $a \mapsto \Phi_a$, $\Phi_a(x) = \sum_i x_i a_i$, is a linear isometry. Use the Hahn-Banach Theorem to show that the embedding T is proper (that is $T(l^1) \neq (l^\infty)^*$) by proving that there exists a bounded linear functional $\Phi \in (l^\infty)^*$ such that $\Phi(x) = \lim_{n\to\infty} x_n$ for $x = (x_n) \in c = \{x = (x_n) \in l^\infty : \exists \lim_{n\to\infty} x_n\}$.

2. Show that there exists $\Phi \in L^\infty([0,1])^*$ such that $\Phi(f) = f(0)$ for $f \in C([0,1])$ and prove that the canonical embedding $T : L^1 \to (L^\infty)^*$ is proper.

Consider some geometrical concepts related to linear continuous functionals. Let X be a normed space and let $f \in X^*$, $f \neq 0$. Consider a set $\Gamma_0 = \operatorname{Ker} f = \{x \in X : f(x) = 0\}$. It is easy to see that this set is a closed subspace of X of codimension 1, i.e. $\operatorname{span}\{z, \operatorname{Ker} f\} = X$ (**Show**).

For $c \in \mathbb{C}$ the set $\Gamma_c = \{x \in X : f(x) = c\}$ is called a hyperplane. Then $\Gamma_c = z + \operatorname{Ker} f$ for some $z \in X$ (**Exercise**).

Assume X is a real normed space, $A \subset X$ a subset and x_0 is a boundary point of A.

Γ_c is called a **supporting hyperplane** of the set A containing x_0 if $x_0 \in \Gamma_c$ (i.e. $f(x_0) = c$) and either $f(x) - c \geq 0$ for all $x \in A$ or $f(x) - c \leq 0$ for all $x \in A$.

Consider a special case when $A = \overline{B(0,r)} = \{x \in X : \|x\| \leq r\}$. The boundary of A is the sphere $S(0,r) = \{x \in X : \|x\| = r\}$.

Theorem 12.10. *If $x_0 \in S(0,r)$ then there exists a supporting hyperplane of $\overline{B(0,r)}$ containing x_0.*

Proof. For $x_0 \in S(0,r)$ there exists $f \in X^*$ such that $\|f\| = 1$ and $f(x_0) = \|x_0\| = r$. Then $\Gamma_r = \{x : f(x) = r\}$ is the desired supporting hyperplane since $x_0 \in \Gamma_r$ and for all $x \in \overline{B(0,r)}$

$$f(x) \leq |f(x)| \leq \|x\| \leq r,$$

i.e. $f(x) - r \leq 0$. $\square$

13 The Baire Category Theorem

This a fundamental theorem in the theory of complete metric spaces. Recall some terminology. Let X be a metric space with metric denoted by ρ. For $a \in X$ and $r > 0$ we set

$$B(a, r) := \{x \in X : \rho(x, a) < r\},$$

the open ball with centrum in a and radius r in X. Let $M \subset X$. Recall that that $x \in M$ is called an **inner point** of M if there exists $r > 0$ such that $B(x, r) \subset M$. A subset M is called **open** if any $x \in M$ is an inner point, and closed if its complement is open. The set of all inner points of $M \subset X$ is the largest open set contained in M; it is called the **interior** of M and is denoted by M^o. The **closure**, $\overline{M}$, of M is the smallest closed set containing M which is the intersection of all closed sets $F \supset M$.

A set $E \subset X$ is called **dense** in X if $\overline{E} = X$. Note that E is dense iff $W \cap E \neq \emptyset$ for any non-empty open set W.

A set $E \subset X$ is **nowhere dense** if the closure of E has no inner point (i.e. has an empty interior). Note that $E \subset X$ is nowhere dense if and only if it is not dense in any open subset of X (Exercise!).

Theorem 13.1. *Let X be a complete metric space.*

1. *If $\{U_n\}_{n=1}^{\infty}$ is a sequence of open dense subsets of X then $\cap_1^{\infty} U_n$ is dense in X*

2. *X is not a countable union of nowhere dense sets.*

Proof. 1. It would be enough to show that if $W \subset X$ is a non-empty open set then $W \cap \cap_1^{\infty} U_n \neq \emptyset$. We note first that since U_1 is dense, $U_1 \cap W$ is non-empty. Since $W \cap U_1$ is also open, there exists a ball $B(x_1, r_1)$, $0 < r_1 < 1$, such that

$$\overline{B(x_1, r_1)} \subset W \cap U_1.$$

Similarly, we find $B(x_2, r_2)$, $0 < r_2 < 1/2$ such that

$$\overline{B(x_2, r_2)} \subset B(x_1, r_1) \cap U_2.$$

Repeating the arguments, we obtain balls $B(x_n, r_n)$, $0 < r_n < 2^{-n+1}$ such that

$$\overline{B(x_n, r_n)} \subset B(x_{n-1}, r_{n-1}) \cap U_n.$$

Then the sequence $\{x_n\}$ is Cauchy, as for any n, $m \geq N$ we have $x_n, x_m \in B(x_N, r_N)$ and hence $\rho(x_n, x_m) \leq \rho(x_n, x_N) + \rho(x_N, x_m) < 2r_N$ which goes to zero as N goes to infinity. As X is complete, the sequence converges to an element $x \in X$ and $\rho(x_N, x) \leq r_N$ for all N. Thus

$$x \in \overline{B(x_N, r_N)} \subset W \cap U_N$$

for all N. Hence $W \cap (\cap_1^\infty U_n) \neq \emptyset$.

2. Let E_n, $n \geq 1$, be nowhere dense sets in X. Then $(\overline{E_n})^c$ are open and dense in X (if $(\overline{E_n})^c$ were not dense in X we could find an open set W such that $W \cap (\overline{E_n})^c = \emptyset$, but then $W \subset \overline{E_n}$ contradicting the condition of E_n being nowhere dense). Hence by (1) $\cap_1^\infty (\overline{E_n})^c$ is dense in X (hence non-empty). Therefore,

$$(\cup_1^\infty E_n)^c = \cap_1^\infty E_n^c \supset \cap_1^\infty (\overline{E_n})^c \neq \emptyset$$

giving $\cup_1^\infty E_n \neq X$. $\square$

The theorem holds for any topological spaces that are homeomorphic to a complete metric space.

Application to Functional Analysis.

- **Open Mapping Theorem**

  Let X, Y be topological spaces. Recall that a map $f : X \to Y$ is called **open** if $f(U)$ is open whenever $U \subset X$ is open; it is called **continuous** if $f^{-1}(U)$ is open whenever $U \subset X$ is so. Hence if f is bijective then f is open if and only if the inverse f^{-1} is continuous.

  If X, Y are normed spaces and f is linear then this boils down to the condition $f(B(0,1)) \supset B(0,r)$ for some $r > 0$. In fact, given open U and $y \in f(U)$, there exists $x \in U$ and $\delta > 0$ such that $f(x) = y$ and $B(x, \delta) = x + B(0, \delta) \subset U$. Then using linearity of f and the condition $f(B(0,1)) \supset B(0,r)$ we can find $\tau > 0$ such that $y + B(0, \tau) \subset f(x) + f(B(0, \delta)) = f(B(x, \delta)) \subset f(U)$.

  Theorem 13.2. *Let X, Y be Banach spaces. If $T : X \to Y$ is a surjective linear bounded map then T is open.*

  Proof. By the remark before the theorem, we only need to show that $T(B(0,1)) \supset B(0,r)$ for some $r > 0$. Write B_r for $B(0,r)$ for the simplicity.

  Since T is surjective and $X = \cup_1^\infty B_n$,

  $$Y = \cup_1^\infty T(B_n).$$

  Hence by the Baire category theorem, one of the $T(B_n)$:s is not nowhere dense, i.e. $\overline{T(B_n)}$ has a non-empty interior. Since $\overline{T(B_n)} = n\overline{T(B_1)}$, $T(B_1)$

is not nowhere dense. Therefore there exist $y_0 \in \overline{T(B_1)}$, $r > 0$ such that $B(y_0, r) \subset \overline{T(B_1)}$.

Let us show first that we can find a ball B_r (with centrum in zero but different r) such that $B_r \subset \overline{T(B_1)}$. Pick $y_1 = Tx_1 \in T(B_1)$ such that $\|y_1 - y_0\| < r/2$. Then for any $y \in Y$, $\|y\| < r/2$, we have $y + y_1 \in B(y_0, r) \subset \overline{T(B_1)}$:

$$\|y + y_1 - y_0\| \leq \|y\| + \|y_1 - y_0\| < r.$$

Hence

$$y = -Tx_1 + (y + y_1) \in -Tx_1 + \overline{T(B_1)} \subset \overline{T(-x_1 + B_1)} \subset \overline{T(B_2)},$$

i.e. $B_{r/2} \subset \overline{T(B_2)}$ and hence $B_{r/4} \subset \overline{T(B_1)}$.

Now we show that for some $r > 0$, $B_r \subset T(B_1)$ (by shrinking the radius r again).

Since for some $r > 0$ and all $y \in Y$ with $\|y\| < r$, $y \in \overline{T(B_1)}$, we have $y \in \overline{T(B_{2^{-n}})}$ whenever $\|y\| < r2^{-n}$. Pick $y \in Y$ such that $\|y\| < r/2$. Then, since $y \in \overline{T(B_{1/2})}$, we can find $x_1 \in B_{1/2}$ such that $\|y - Tx_1\| < r/4$. This entails that $y - Tx_1 \in \overline{T(B_{1/4})}$. Therefore we can find $x_2 \in B_{1/4}$ with $\|y - Tx_1 - Tx_2\| < r/8$. Then $y - Tx_1 - Tx_2 \in \overline{T(B_{1/8})}$. Proceeding inductively, we find $x_n \in B_{2^{-n}}$ such that

$$\|y - \sum_{i=1}^{n} Tx_i\| < r2^{-n-1}. \tag{13.10}$$

Since X is complete and the series $\sum_n x_n$ is absolutely convergent, it converges in X to some $x \in X$. Moreover

$$\|x\| \leq \sum_{1}^{\infty} \|x_n\| \leq \sum_{1}^{\infty} 2^{-n} = 1,$$

i.e. $x \in B_1$. But then, as T is continuous, $\sum_{i=1}^{n} Tx_i \to Tx$. On the other hand, by (13.10), $\sum_{i=1}^{n} Tx_i \to y$, as $n \to \infty$. Hence $y = Tx \in T(B_1)$ and $T(B_1) \supset B_{r/2}$. $\square$

Corollary 13.3. *(Banach bounded inverse theorem) If X, Y are Banach spaces and $T \in L(X, Y)$ is bijective, then $T^{-1} \in L(Y, X)$ and there exists $C > 0$ such that*

$$C^{-1}\|x\| \leq \|Tx\| \leq C\|x\| \textit{ for all } x \in X.$$

Proof. Since T is bijective, T^{-1} exists. By the Open Mapping Theorem, T is open and hence T^{-1} is continuous and therefore bounded. Let $C = \max\{\|T\|, \|T^{-1}\|\}$ then for all $x \in X$,

$$\begin{aligned} C^{-1}\|x\| &\leq \|T^{-1}\|^{-1}\|x\| = \|T^{-1}\|^{-1}\|T^{-1}Tx\| \leq \|T^{-1}\|^{-1}\|T^{-1}\|\|Tx\| \\ &\leq \|Tx\| \leq \|T\|\|x\| \leq C\|x\|. \end{aligned}$$

□

- **The closed Graph Theorem.**

If X and Y are normed vector spaces and T is a linear map from X to Y we define a *graph* of T by

$$\Gamma(T) = \{(x,y) \in X \times Y : y = Tx\} = \{(x,Tx) : x \in X\}.$$

$X \times Y$ becomes a normed space when equipped with the *product norm*

$$\|(x,y)\| := \max\{\|x\|, \|y\|\}.$$

Here $\|x\|$ refers to the norm on X while $\|y\|$ refers to the norm on Y.

Definition 13.4. We say that T is **closed** if $\Gamma(T)$ is a closed subspace in $X \times Y$.

Clearly if T is continuous then $\Gamma(T)$ is closed, since if $x_n \to x$ then $Tx_n \to Tx$. Closedness means that if $x_n \to x$ and $Tx_n \to y$ then $y = Tx$. Next theorem says that if X and Y are complete then closedness of a linear operator gives the continuity.

Theorem 13.5. *If X and Y are Banach spaces and $T : X \to Y$ is a closed linear map then T is bounded.*

Proof. Let π_1 and π_2 be projections of $\Gamma(T)$ onto X and Y respectively:

$$\pi_1((x,Tx)) = x \quad \pi_2((x,Tx)) = Tx.$$

Then $\pi_1 \in L(\Gamma(T), X)$ and $\pi_2 \in L(\Gamma(T), Y)$ since

$$\|\pi_1((x,Tx))\| = \|x\| \leq \|(x,Tx)\| \text{ and } \|\pi_1((x,Tx))\| = \|Tx\| \leq \|(x,Tx)\|$$

($\|\pi_1\| \leq 1$, $\|\pi_2\| \leq 1$). Since X and Y are complete so is $X \times Y$ and $\Gamma(T)$ (as a closed subspace). The map π_1 is bijective as a map from $\Gamma(T)$ to X and by Corollary 13.3 π_1^{-1} is bounded. But then $T = \pi_2 \circ \pi_1^{-1}$ is also bounded. □

- **The Uniform Boundedness Principle.** The next theorem, known also as the Banach-Steinhaus theorem, allows one to deduce uniform estimates from pointwise estimates.

  Theorem 13.6. (**The Uniform Boundedness Principle.**) *Assume that X is a Banach space and Y is a normed space. Let $\{T_\alpha : \alpha \in \mathcal{A}\}$ be a family of operators in $L(X, Y)$. If $\sup_{\alpha\in\mathcal{A}} \|T_\alpha x\| \le c_x < \infty$ for all $x \in X$, then $\sup_{\alpha\in\mathcal{A}} \|T_\alpha\| < \infty$.*

  Proof. Let

  $$E_n = \{x : \sup_{\alpha\in\mathcal{A}} \|T_\alpha x\| \le n\} = \cap_{\alpha\in\mathcal{A}}\{x : \|T_\alpha x\| \le n\}.$$

  By assumption $X = \cup_n E_n$. Moreover, each E_n is closed as intersection of closed sets. By the Baire category theorem, some E_n must contain a ball, $\overline{B(x_0, r)}$. Since for any $x \in X$ such that $\|x\| \le r$ we have $x + x_0 \in \overline{B(x_0, r)} \subset E_n$, we obtain

  $$\|T_\alpha x\| \le \|(T_\alpha(x + x_0)\| + \|T_\alpha x_0\| \le 2n.$$

  Hence for x, $\|x\| \le 1$, this gives

  $$\|T_\alpha x\| = \|T_\alpha(rx)\|/r \le 2n/r,$$

  and hence $\|T_\alpha\| \le 2n/r$ for all $\alpha \in \mathcal{A}$. □

14 Weak and weak $*$ topologies

Weak convergence. Recall that for infinite-dimensional normed spaces closed and bounded sets are not necessarily (sequentially) compact in the topology induced by the norm: the unit sphere is compact if and only if the space is finite-dimensional (see Exercise sheet, week I). However we still hope for some type of convergence using a different topology.

Definition 14.1. Let X be a normed space, $\{x_n : n \geq 1\} \subset X$, $x \in X$. We say that $x_n \to x$ weakly if

$$f(x_n) \to f(x) \text{ for all } f \in X^*.$$

Remark 14.2.

1. If $x_n \to x$ weakly and $x_n \to y$ weakly then $x = y$ (the weak limit is unique). The reason is because then $f(x-y)$ for all $f \in X^*$ and by a consequence from the Hahn-Banach theorem, $x - y = 0$.

2. If $x_n \to x$ in norm, i.e. $\|x_n - x\| \to 0$ then $x_n \to x$ weakly. This follows from

$$|f(x_n) - f(x)\| \leq \|f\|\|x_n - x\| \to 0.$$

Even with this weaker notion of convergence, it is not always true that bounded sequences have convergent subsequences. We will have this for reflexive spaces. Let's look first at some characterisations of weak convergence.

Theorem 14.3. *Suppose $x_n \to x$ weakly. Then $\sup_n \|x_n\| < \infty$, i.e. the sequence is bounded.*

Proof. Apply the Uniform Boundedness Principle. Recall that the mapping $x \in X \mapsto \hat{x} \in X^{**}$, where $\hat{x}(f) = f(x)$, $f \in X^*$, is isometric so that

$$\|x\|_X = \|\hat{x}\|_{X^{**}}.$$

Then $\hat{x}_n \in X^{**}$ and given $f \in X^*$, $\sup_n |\hat{x}_n(f)| = \sup_n |f(x_n)| < \infty$ as $f(x_n)$ converges. Thus by the Uniform Boundedness Principle

$$\sup_n \|x_n\|_X = \sup_n \|\hat{x}_n\|_{X^{**}} < \infty.$$

□

Next statement is a characterisation of weak convergence.

Theorem 14.4. *Let $\{x_n : n \geq 1\} \subset X$, $x \in X$. Then $x_n \to x$ weakly if and only if*

1. $\sup_n \|x_n\| < \infty$*;*

2. $f(x_n) \to f(x)$ *for any* $f \in M$*, where* M *is a subset of* X^* *whose linear span (the set of all finite linear combinations of elements in* M*) is dense in* X^**.*

Proof. Assume first that $x_n \to x$ weakly. Then the condition (2) is obvious while the first one follows from Theorem 14.3.

To see the reverse statement, we take arbitrary $f \in X^*$ and $\{f_k : k \geq 1\} \subset \operatorname{span}\{M\}$ such that $f_k \to f$ in X^*. We must show that $f(x_n) \to f(x)$, as $n \to \infty$. Let $C > 0$ be such that $\|x_n\| < C$ and $\|x\| < C$. As $f_k \to f$, given $\epsilon > 0$, there exists $K > 0$ such that

$$\|f_k - f\| < \epsilon \quad \forall k \geq K.$$

By the second condition, $f_K(x_n) \to f_K(x)$ and hence there exists $N \geq 0$ such that $|f_K(x_n) - f_K(x)| < \epsilon$ for all $n \geq N$. Therefore, for all $n \geq N$ we have

$$\begin{aligned} |f(x_n) - f(x)| &\leq |f(x_n) - f_K(x_n)| + |f_K(x_n) - f_K(x)| + |f_K(x) - f(x)| \\ &\leq \|f - f_K\|\|x_n\| + |f_K(x_n) - f_K(x)| + \|f - f_K\|\|x\| \\ &\leq 2\epsilon C + \epsilon \end{aligned}$$

giving the statement. □

Example 14.5. 1. $X = \mathbb{R}^n$. The weak convergence coincides with the norm convergence (with respect to any norm on X). In fact, since all norms on a finite-dimensional vector space are equivalent, it is enough to see the statement for the Euclidean norm $\|(x^{(1)}, \dots x^{(n)})\| = \sqrt{((x^{(1)})^2 + \dots + (x^{(n)})^2}$. Let $x_k \to x$ weakly in $\mathbb{R}^n$. Clearly, $f_k(x) = x^{(k)}$, $x = (x^{(1)}, \dots x^{(n)})$, is a bounded linear functional. Hence $x_k^{(i)} \to x^{(i)}$ for all $i = 1, \dots, n$. But then

$$\|x_k - x\| = (\sum_{i=1}^{n} (x_k^{(i)} - x^{(i)})^2)^{1/2} \to 0.$$

The other implication follows from Remark 14.2.

2. $X = \ell^2$. Let $x_k = (x_k^{(1)}, x_k^{(2)}, \dots) \in \ell^2$, $x = (x^{(1)}, x^{(2)}, \dots) \in \ell^2$. Then $x_k \to x$ weakly iff $x_k^{(i)} \to x^{(i)}$ for all $i \geq 1$ and $\sup_k \|x_k\| < \infty$.

In fact, assume first that $x_k \to x$ weakly. As $f_i(x) = \langle x, e_i \rangle = x^{(i)}$ is a linear bounded functional, $x_k^{(i)} \to x^{(i)}$ for any i. The boundedness of the norm $\|x_k\|$ follows from Theorem 14.3.

To see the other inclusion let $x_k^{(i)} \to x^{(i)}$ for all i. Then $f(x_k) \to f(x)$ for all f: $f(x) = \langle x, a \rangle$, where a is a finite linear combination of vectors e_1, $e_2, \dots$, i.e. $a \in \ell_F^2$. By the Riesz-Fréchet theorem any bounded linear functional is given by $f(x) = \langle x, a \rangle$, where $a \in l^2$, and $\|f\| = \|a\|_2$. As ℓ_F^2 is dense in l^2, assuming also the boundedness of $\|x_k\|$, Theorem 14.4 will give the weak convergence.

Note that in ℓ^2 the weak convergence does not coincide with the norm convergence. In fact, the sequence $\{e_k\}$ converges weakly to 0 as $e_k^{(i)} \to 0$ for all i, as $k \to \infty$. But since $\|e_k\| = 1$, $e_k \not\to 0$ in norm.

Exercise 14.6. *1. Let $\{x_n\}$ be a sequence in $L^p([a,b])$, $x \in L^p([a,b])$, $1 < p < \infty$. Show that $x_n \to x$ weakly iff $\sup_n \|x_n\| < \infty$ and $\int_a^\tau x_n(t)dt \to \int_a^\tau x(t)dt$ for all $\tau \in [a,b]$.*

2. Let H be a Hilbert space, x, $x_n \in H$, $n \geq 1$. Show that $x_n \to x$ in H if $x_n \to x$ weakly and either (a) $\|x_n\| \to \|x\|$ or (b) $\limsup \|x_n\| \leq \|x\|$.

Even though weakly convergent sequence does not converge in norm in general, we have the following useful statement.

Theorem 14.7. *Let X be a normed space, $\{x_n : n \geq 1\} \subset X$, $x \in X$. If $x_n \to x$ weakly then there exists a sequence of finite linear combinations of elements x_1, $x_2, \dots$ that converges to x in norm.*

Proof. Let G be the closed linear span of x_n:s. It is enough to see that $x \in G$. Assume contrary that $x \notin G$. Then by a corollary from the Hahn-Banach theorem there exists $f \in X^*$ such that $f(x) = \rho(x, G)$, $\|f\| = 1$ and $f|_G = 0$. Then $f(x_n) = 0$ while $\lim f(x_n) = f(x) \neq 0$, a contradiction. $\square$

Weak Topology.

Let X be a normed space and X^* be its dual. If $f_1, f_2, \dots, f_n \in X^*$, $\epsilon > 0$, $x \in X$, let

$$U_{x,f_1,\dots,f_n,\epsilon_1,\dots\epsilon_n} = \{y \in X : |f_i(y - x)| < \epsilon_i, i = 1, \dots, n\}.$$

The set is open, contains x and hence is a neighborhood of x. The finite intersections of such sets is again a set of this type. Hence in X one can introduce a topology whose base is the family of neighborhoods of the above form (see Folland, Chapter 4). It is called the **weak topology** of X. Then the weak convergence can be formulated in the following way:

$x_n \to x$ weakly in X iff for any $U_{x,f_1,\dots,f_k,\epsilon_1,\dots,\epsilon_k}$ there exists $N > 0$ such that $x_n \in U_{x,f_1,\dots,f_k,\epsilon_1,\dots\epsilon_k}$ for all $n \geq N$ (**Show!**)

Weak* Topology

Consider now the dual space X^* of a normed space X. Besides the norm convergence on X^*, one can consider the following two topologies:

- Considering X^* as a Banach space we can equip X^* with the weak topology coming from the functionals in $(X^*)^* = X^{**}$:

$$f_n \to f \text{ weakly if } \psi(f_n) \to \psi(f) \quad \forall \psi \in X^{**}.$$

- The other topology, called the **weak* topology**, is induced by $\hat{X} = \{\hat{x} : x \in X\}$, here $\hat{x}(f) := f(x)$ for $f \in X^*$:

$$f_n \to f \text{ weakly}^* \text{ if} f_n(x) \to f(x) \quad \forall x \in X,$$

equivalently, $\hat{x}(f_n) \to \hat{x}(f)$ for all $x \in X$.

The neighborhoods

$$U_{x_1,\dots,x_n,\epsilon_1,\dots\epsilon_n} = \{f \in X^* : |f(x_i)| < \epsilon_i, i = 1, \dots, n\},$$

where $\epsilon > 0$, $x_1, \dots, x_n \in X$, is a neighborhood base for 0 this topology. We have:

$f_n \to 0$ weakly* iff for any $U_{x_1,\dots,x_k,\epsilon_1,\dots,\epsilon_k}$ there exists $N > 0$ such that $f_n \in U_{x_1,\dots,x_k,\epsilon_1,\dots,\epsilon_k}$ for any $n \geq N$ (**Show!**)

Remark 14.8. 1. When speaking of weak*-convergence on a normed space Y one must always be able to see Y as the dual of some normed space, i.e. $Y = X^*$. One says that X is a predual of Y.

2. If X is reflexive then clearly the two topologies, weak and weak* topologies, coincide.

Here is a characterisation of weak* convergence similar to one about weak convergence with a similar proof.

Theorem 14.9. *Let X be a Banach space. A sequence $\{f_n : n \geq 1\}$ in X^* converges weakly* to $f \in X^*$ if an only if*

1. *the sequence is bounded, i.e. $sup_n \|f_n\| < \infty$;*
2. *$f_n(x) \to f(x)$ for all x in a subset whose linear span is dense in X.*

Theorem 14.10. *Let X be a Banach space and let $\{f_n : n \geq 1\} \subset X^*$ such that the sequence $\{f_n(x) : n \geq 1\}$ is Cauchy for all $x \in X$. Then there exists $f \in X^*$ such that $f_n \to f$ weakly *.*

Proof. As $\{f_n(x)\}$ is Cauchy it is convergent. We let for each $x \in X$, $f(x) = \lim f_n(x)$. Show that $f \in X^*$. □

The final theorem of this section is the compactness theorem for bounded sets of the dual spaces.

Theorem 14.11. *(***The Banach-Alaoglu Theorem***). Let X be a separable normed space and $\{f_n : n \geq 1\}$ is a sequence in X^* with $\|f_n\| \leq 1$, $n = 1, 2, \ldots$. Then it has a subsequence convergent in the weak*-topology.*

Proof. Let $\{x_1, x_2, \ldots\}$ be a dense subset in X. If $\{f_n : n \geq 1\}$ is a bounded sequence of linear functionals on X then the sequence

$$f_1(x_1), f_2(x_1), \ldots, f_n(x_1), \ldots$$

is bounded ($|f_n(x_1)| \leq \|f_n\|\|x_1\|$). Therefore we can choose a subsequence

$$f_1^{(1)}, f_2^{(1)}, \ldots, f_n^{(1)}, \ldots$$

such that the subsequence $f_1^{(1)}(x_1), f_2^{(1)}(x_1), \ldots, f_n^{(1)}(x_1), \ldots$ converges. Then from $\{f_n^{(1)} : n \geq 1\}$ one can choose a subsequence

$$f_1^{(2)}, f_2^{(2)}, \ldots, f_n^{(2)}, \ldots$$

such that the subsequence $f_1^{(2)}(x_2), f_2^{(2)}(x_2), \ldots, f_n^{(2)}(x_2), \ldots$ converges. Proceeding inductively we obtain the following system of subsequences:

$$\begin{array}{l} f_1^{(1)}, f_2^{(1)}, \ldots, f_n^{(1)}, \ldots \\ f_1^{(2)}, f_2^{(2)}, \ldots, f_n^{(2)}, \ldots \\ \qquad\qquad\qquad \ldots \end{array}$$

each of them is contained in the previous one so that $\{f_n^{(k)}(x_i) : n \geq 1\}$ convergence for all $i = 1, 2, \ldots, k$. Taking now the "diagonal"

$$f_1^{(1)}, f_2^{(2)}, \ldots, f_n^{(n)}, \ldots$$

we obtain a subsequence of linear functionals such that $f_1^{(1)}(x_n), f_2^{(2)}(x_n), \ldots$ converges for all n. Then by Theorem 14.9, the sequence $f_1^{(1)}(x), f_2^{(2)}(x), \ldots$ converges for all $x \in X$. The statement now follows from Theorem 14.10 □

Remark 14.12. One can prove that for separable normed spaces the weak*-toplogy on the closed unit ball is metrizable and thus the compactness and sequential compactness are equivalent. Therefore the above theorem stays that the unit ball of the dual space of a separable space is compact.

Remark 14.13. If X is a reflexive space, then any bounded sequence of X has a weakly convergent subsequence. One proves this fact using the following steps:

Step 1: Let $\{x_n\}$ be a bounded sequence. Set Y to be the closed linear span of x_n:s. This is a separable subspace of X.

Step 2: As Y is closed one can prove that Y is reflexive as well.

Step 3: As $(Y^*)^* = Y$ is separable, one can show that Y^* is separable.

Step 4: exercise=complete the proof.

The "sequential" Banach-Alaoglu theorem is a special case of

Theorem 14.14. (**Alaoglu's Theorem**) *If X is a normed space, the closed unit ball*

$$B^* = \{f \in X^* : \|f\| \leq 1\}$$

in X^* is compact in the weak*-topology.

The proof of this general statement relies on Tychonoff's theorem about compactness of the product of any collection of compact topological spaces.

An interesting application of the Alaoglu Theorem is the following statement:

Proposition 14.15. *Any normed space X can be isometrically embedded into $C(\Omega)$, the Banach space of continuous functions on a compact Ω.*

Proof. Exercise. Hint: Take $\Omega = \{f \in X^* : \|f\| \leq 1\}$ and consider the mapping $X \ni x \mapsto \hat{x}|_\Omega \in C(\Omega)$. □

15 Hilbert spaces: Orthogonality. Riesz-Fréchet Theorem

Let H be a Hilbert space with inner product $\langle \cdot, \cdot \rangle$. We say that x, $y \in H$ are *orthogonal* if $\langle x, y \rangle = 0$. Write $x \perp y$.

If $E \subset H$ define

$$E^{\perp} = \{x \in H : \langle x, y \rangle = 0, y \in E\}$$

and call it the *orthogonal complement* to E.

Proposition 15.1. *If $E \subset H$ then*

- *$E^{\perp}$ is a closed subspace in H;*
- *$E \subseteq (E^{\perp})^{\perp}$;*
- *$(E^{\perp})^{\perp} = \overline{span\{E\}}$.*

Proof. Exercise. $\square$

Theorem 15.2. *(***The Pythagorean Theorem.***)If $x_i \perp x_j$ for all $1 \leq i < j \leq n$ then*

$$\| \sum_{1}^{n} x_j \|^2 = \sum_{1}^{n} \|x_j\|^2.$$

Proof. Exercise. $\square$

Proposition 15.3. *Let H be a Hilbert space, H_0 a closed subspace of H and $x \in H$. Then $x \perp H_0$ iff the distance from x to H_0 equals $\|x\|$.*

Proof. Let $x \perp H_0$ then for any $y \in H_0$ we have

$$\|x - y\|^2 = \langle x - y, x - y \rangle = \|x\|^2 + \|y\|^2 \geq \|x\|^2.$$

Hence $\inf\{\|x - y\| : y \in H_0\} = \|x - 0\| = \|x\|$.

Assume $\inf\{\|x - y\| : y \in H_0\} = \|x\|$. Then for any $y \in H_0$ and $\lambda \in \mathbb{C}$

$$\|x - \lambda y\| \geq \|x\|.$$

Hence $\langle x - \lambda y, x - \lambda y \rangle \geq \langle x, x \rangle$ and

$$\langle x, x \rangle - \lambda \langle y, x \rangle - \bar{\lambda} \langle x, y \rangle + \lambda \bar{\lambda} \langle y, y \rangle \geq \langle x, x \rangle.$$

Letting $\lambda = t \langle x, y \rangle$ we see that for all $t \geq 0$

$$-2t |\langle x, y \rangle|^2 + t^2 |\langle x, y \rangle|^2 \langle y, y \rangle \geq 0$$

and $\langle x, y \rangle|^2 t(t \langle y, y \rangle - 2) \geq 0$. Therefore $\langle x, y \rangle = 0$. $\square$

Proposition 15.4. *Let H be a Hilbert space, M a closed subspace, $x \in H$. Then there is a unique $z \in M$ such that*

$$\|z - x\| = \rho(x, M) := \inf\{\|y - x\| : y \in M\}.$$

Proof. Let $d = \rho(x, M)$ and let $\{y_n\} \subset M$ such that $\|x - y_n\| \to d$. By the parallellogram law (applied to $y_n - x$, $y_m - x$)

$$2(\|y_n - x\|^2 + \|y_m - x\|^2) = \|(y_n - x) - (y_m - x)\|^2 + \|(y_n - x) + (y_m - x)\|^2$$

i.e.

$$2(\|y_n - x\|^2 + \|y_m - x\|^2) = \|y_n - y_m\|^2 + \|y_n + y_m - 2x)\|^2 \qquad (15.11)$$

Since $\frac{1}{2}(y_n + y_m) \in M$, $\|\frac{y_n+y_m}{2} - x\| \geq d$ and we have from (15.11)

$$\begin{aligned} \|y_n - y_m\|^2 &= 2\|y_n - x\|^2 + 2\|y_m - x\|^2 - 4\|\frac{y_n + y_m}{2} - x\|^2 \\ &\leq 2\|y_n - x\|^2 + \|y_m - x\|^2 - 4d^2 \to 0 \text{ as } m, n \to \infty \end{aligned}$$

Hence $\{y_n\}$ is a Cauchy sequence. As H is complete, $y_n \to z$ for some $z \in H$. As M is closed, $z \in M$. As $\|x - y_n\| \to \|x - z\|$ we obtain $d = \|x - z\|$.

Uniqueness: Let z_1, $z_2 \in M$ be such that $\|z_1 - x\| = \rho(x, M) = \|z_2 - x\|$ and let $d = \rho(x, M)$. By the parallelogram law applied now to $(z_1 - x)$ and $(z_2 - x)$ we obtain

$$2(\|z_1 - x\|^2 + \|z_2 - x\|^2) = \|z_1 + z_2 - 2x\|^2 + \|z_1 - z_2|^2$$

giving $\|z_1 - z_2\|^2 = 4d^2 - 4\|x - (z_1 + z_2)/2\|^2 \leq 4d^2 - 4d^2 = 0$, i.e. $z_1 = z_2$. $\square$

Theorem 15.5. *(about orthogonal complement) If M is a closed subspace of a Hilbert space H then $H = M \oplus M^\perp$, i.e. each $x \in H$ can be expressed uniquely as $x = z + y$ where $z \in M$ and $y \in M^\perp$.*

Proof. Let $x \in H$ and let $z \in M$ be the closest to x vector in M which exists by Proposition 15.4. Let $y = x - z$. Then $\|y\| = \rho(x, M)$ and since $z \in M$

$$\begin{aligned} \rho(y, M) &= \inf\{\|y - \tilde{y}\| : \tilde{y} \in M\} = \inf\{\|x - z - \tilde{y}\| : \tilde{y} \in M\} \\ &= \inf\{\|x - \tilde{y}\| : \tilde{y} \in M\} = \rho(x, M) = \|y\|. \end{aligned}$$

By Proposition 15.3, $y \in M^\perp$. Hence $x = (x - z) + z = y + z$ is the desired decomposition.

To show uniqueness let $x = z_1 + y_1$, $z_1 \in M$, $y_1 \in M^\perp$ then $z_1 + y_1 = z + y$ and $M \ni y_1 - y = z - z_1 \in M^\perp$. As $M \cap M^\perp = \{0\}$, we obtain $y_1 = y$ and $z_1 = z$. $\square$

Theorem 15.6. *(Riesz-Fréchet) If $f \in H^*$ there is a unique $y \in H$ such that $f(x) = \langle x, y\rangle$ for all $x \in H$.*

Proof. Uniqueness: if $\langle x, y\rangle = \langle x, y'\rangle$ for all $x \in H$, taking $x = y - y'$ we obtain $\langle y - y', y - y'\rangle = 0$ and hence $y - y' = 0$.

Existence: If f is the zero functional them take $y = 0$. Assume $f \neq 0$ and let $M = \{x \in H : f(x) = 0\}$. Then M is a proper closed subspace and hence $M^\perp \neq \{0\}$. Take $y' \in M^\perp$ such that $\|y'\| = 1$ and let $y = \overline{f(y')}y'$. Then for any $x \in M$ we have $f(x) = \langle x, y\rangle = 0$ and for y' we have $f(y') = \langle y', y\rangle$. Thus $f = \Phi_y$ on $\operatorname{span}\{M, y'\}$, here $\Phi_y(x) = \langle x, y\rangle$.

As M is a subspace of codimension 1 (obs! $M = \operatorname{Ker} f$) we have $H = \operatorname{span}\{M, y'\}$ and hence $f = \Phi_y$ on the whole H. $\square$

We have established before that the mapping $H \ni a \mapsto \Phi_a$ is an anti-linear isometry. By the Riesz-Fréchet theorem we obtain that the mapping is surjective and hence $H \simeq H^*$ (H is self-dual).

15.1 Orthogonal systems

Definition 15.7. A subset $\{u_\alpha\}_{\alpha\in A}$ of a Hilbert space H is called *orthogonal* if $u_\alpha \perp u_\beta$ whenever $\alpha \neq \beta$. It is called *orthonormal* if in addition $\|u_\alpha\| = 1$.

Any orthogonal system $\{e_i\}_{i\in I}$ is linearly independent (i.e. if a finite linear combination $\sum_{i\in F} \lambda_i e_i = 0$, F is finite, then $\lambda_i = 0$).

Examples 15.8. 1. $H = \ell^2$, the standard basis $\{e_n\}_{n\geq 1}$ is an orthonormal system;

2. $H = L^2([-\pi, \pi])$ then $\{\frac{1}{\sqrt{2\pi}}e^{int}, n \in \mathbb{Z}\}$ is an orthonormal system;

3. $H = L^2(\mathbb{R})$. Consider $\{t^n e^{-t^2/2} : n = 0, 1, 2, \ldots\}$. It is a linearly independent system. Applying to it the Gram-Schmidt process we get a system $\{p_n(t)e^{-t^2/2}, n = 0, 1, 2, \ldots\}$ whose elements are called Hermite functions.

Definition 15.9. Let $\{e_n : n \in \mathbb{N}\}$ be an orthonormal system in a Hilbert space H, $x \in H$. The formal series $\sum_{n=1}^\infty \langle x, e_n\rangle e_n$ in H is called the *Fourier series* of x corresponding this system . The coefficients $\langle x, e_n\rangle$ are called the *Fourier coefficients* of x with respect to this system.

Examples 15.10. For $H = L^2([-\pi, \pi])$ and $\{e_n = \frac{1}{\sqrt{2\pi}}e^{int}, n \in \mathbb{Z}\}$ we obtain the classical Fourier series and Fourier coefficients.

Theorem 15.11. Bessel's Inequality *If $\{u_\alpha\}_{\alpha\in A}$ is an orthonormal set in H then for any $x \in H$,*

$$\sum_{\alpha\in A} |\langle x, u_\alpha\rangle|^2 \leq \|x\|^2.$$

In particular, $\{\alpha \in A : \langle x, u_\alpha\rangle \neq 0\}$ is countable.

Proof. For finite subset $F \subset A$ we have

$$\begin{aligned} 0 &\leq \|x - \sum_{\alpha \in F} \langle x, u_\alpha \rangle u_\alpha\|^2 = \langle x - \sum_{\alpha \in F} \langle x, u_\alpha \rangle u_\alpha, x - \sum_{\alpha \in F} \langle x, u_\alpha \rangle u_\alpha \rangle \\ &= \|x\|^2 + \sum_{\alpha \in F} |\langle x, u_\alpha \rangle|^2 - 2 \sum_{\alpha \in F} |\langle x, u_\alpha \rangle|^2 = \|x\|^2 - \sum_{\alpha \in F} |\langle x, u_\alpha \rangle|^2, \end{aligned}$$

and hence $\sum_{\alpha \in F} |\langle x, u_\alpha \rangle|^2 \leq \|x\|^2$. Therefore,

$$\sum_{\alpha \in A} |\langle x, u_\alpha \rangle = \sup\{\sum_{\alpha \in F} |\langle x, u_\alpha \rangle|^2 : F \subset A \text{ is finite}\} \leq \|x\|^2.$$

We have $B := \{\alpha \in A : \langle x, u_\alpha \rangle \neq 0\} = \cup_n A_n$, where $A_n = \{\alpha : |\langle x, u_\alpha \rangle| > 1/n\}$. If B were uncountable so would be A_n for at least one $n \geq 1$ and hence for any finite $F \subset A_n$ we would have $\sum_{\alpha \in F} |\langle x, u_\alpha \rangle|^2 > \text{card}(F)/n$ implying $\sum_{\alpha \in A} |\langle x, u_\alpha \rangle|^2 = \infty$. A contradiction. Thus each A_n and B are countable. $\square$

Definition 15.12. An orthonormal set is called an orthonormal basis for H if its linear span is dense in H.

Examples 15.13.

1. $\{e_n\}$ is an orthonormal basis in l^2.
2. $\{\frac{1}{\sqrt{2\pi}} e^{int}, n \in \mathbb{Z}\}$ is an orthonormal basis in $L^2([-\pi, \pi])$.
3. {The Hermite functions} is an orthonormal basis in $L^2(\mathbb{R})$.

Theorem 15.14. *Let $\{u_\alpha\}_{\alpha \in A}$ be an orthonormal set. The following are equivalent*

1. *$\{u_\alpha\}_{\alpha \in A}$ is an orthonormal basis for H;*
2. *(Completeness) If $\langle x, u_\alpha \rangle = 0$ for all α then $x = 0$.*
3. *(Parseval's Identity) $\|x\|^2 = \sum_\alpha |\langle x, u_\alpha|^2$ for all $x \in H$;*
4. *For each $x \in H$ $x = \sum_{\alpha \in A} \langle x, u_\alpha \rangle u_\alpha$, where the sum on the right hand side has only countable many non-zero terms and converges in the norm topology no matter how these terms are ordered.*

For the proof see Folland, 5.27.

Theorem 15.15. *Every Hilbert space has an orthonormal basis.*

Proof. A routine application of Zorn's lemma shows that the collection of orthogonal sets, ordered by inclusion, has a maximal element; and maximality is equivalent to the second property of Theorem 15.14. $\square$

Remark 15.16. Not all Banach spaces have a "topological basis".

Proposition 15.17. *A Hilbert space H is separable if and only if it has a countable orthonormal basis in which case every orthonormal basis for H is countable.*

Proof. If $\{x_n\}$ is a countable dense set in H, by discarding recursively any x_n that is in the linear span of $x_1, \dots, x_{n-1}$, we obtain a linearly independent sequence $\{y_n\}$ whose linear span is dense in H. Application of the Gram-Schmidt process to $\{y_n\}$ yields an orthonormal sequence $\{u_n\}$ whose linear span is dense in H and which is therefore a basis. Conversely, if $\{u_n\}$ is a countable orthonormal basis, the finite linear combinations of the u_n's with coefficients in a countable dense subsets of $\mathbb{C}$ from a countable dense set in H. Moreover, if $\{v_\alpha\}$ is another orthonormal basis, for each n the set $A_n = \{\alpha : \langle u_n, v_\alpha \rangle \neq 0\}$ is countable. By completeness of $\{u_n\}$, $A = \cup_1^\infty A_n$, so A is countable. $\square$

Most Hilbert spaces that arise in practice are separable. Next statement says that all such infinite-dimensional spaces are isomorphic to ℓ^2.

Definition 15.18. Let H_1 and H_2 are Hilbert spaces with inner products $\langle \cdot, \cdot \rangle_1$ and $\langle \cdot, \cdot \rangle_2$ respectively. A unitary map from H_1 to H_2 is an invertible linear map $U : H_1 \to H_2$ that preserves inner product, i.e.

$$\langle Ux, Uy \rangle_2 = \langle x, y \rangle_1 \text{ for any } x, y \in H_1.$$

Proposition 15.19. *Let $\{u_\alpha\}$ be an orthonormal basis for H. Then the correspondence $x \mapsto \hat{x}$ defined by $\hat{x}(\alpha) = \langle x, u_\alpha \rangle$ is a unitary map from H to $\ell^2(A)$.*

Proof. The map $U : x \mapsto \hat{x}$ is clearly linear and it is isometry from H to $\ell^2(A)$ by the Parseval identity $\|x\|^2 = \sum |\hat{x}(\alpha)|^2$. If $f \in \ell^2(A)$ then $\sum |f(\alpha)|^2 < \infty$ (in the series only a countable number of terms are non-zero). The Pythagorean theorem shows that the partial sums of the series $\sum f(\alpha) u_\alpha$ is Cauchy; hence $x = \sum f(\alpha) u_\alpha$ exists in H and $\hat{x} = f$, i.e. the map U is surjective and hence invertible. As U is an isometry,it follows from the polarization identity

$$\langle x, y \rangle = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2 + i\|x+iy\|^2 - i\|x-iy\|^2)$$

it follows that U preserves inner product. $\square$

16 The Riesz Representation Theorem. Dual of $C_0(X)$

We recall first some topological facts. A topological space X is said to be **locally compact** if for every $x \in X$ there is a compact set K such that $x \in K^0$ (i.e. x is an inner point of K).

One says that X is *Hausdorff* if for any distinct x, $y \in X$ there are disjoint open sets V and W such that for $x \in V$ and $y \in W$. When X is locally compact and Hausdorff we write X is LCH. Throughout this section, X will denote an LCH space.

We write $C(X)$ for the set of all continuous functions $f : X \to \mathbb{C}$,

$$C_0(X) = \{f \in C(X) : \{x : |f(x)| \geq \varepsilon\} \text{ is compact for any } \varepsilon > 0\}.$$

$C_0(X)$ is a Banach space with respect to the uniform norm $\|f\|_\infty = \sup_{x \in X} |f(x)|$.

Radon measures. Let μ be a Borel measure on X and $E \subset X$ be a Borel subset of X. The measure μ is called **outer regular** on E if

$$\mu(E) = \inf\{\mu(U) : U \supset E, U \text{ is open}\}$$

and **inner regular** on E if

$$\mu(E) = \sup\{\mu(U) : K \subset E, K \text{ is compact}\}.$$

If μ is outer and inner regular on all Borel sets μ is called **regular**.

Example 16.1. The Lebesgue measure on $\mathbb{R}^n$ is regular.

Definition 16.2. A **Radon measure** on X is a Borel measure that is finite on all compacts, outer regular on all Borel subsets and inner regular on all open sets.

The following are proved in Folland, 7.2:

- Every σ-finite Radon measure is regular.
- If X is σ-compact, i.e. there exist a sequence $\{K_n\}$ of compact subsets of X such that $K_n \subset K_{n+1}$, $n \geq 1$, and $X = \cup_n K_n$, then every Radon measure is regular.
- If X is σ-compact then every Borel measure on X that is finite on compact sets is regular and hence Radon.

Definition 16.3. A **signed Radon measure** is a signed Borel measure whose positive and negative variations are Radon (Recall: if ω is a signed measure then there exist positive measures ω_+, ω_- such that $\omega(E) = \omega_+(E) - \omega_-(E)$ for any measurable E), a complex measure is Radon if its real and imaginary parts are signed Radon measure.

Note that if X is σ-compact then every complex Borel measure is Radon. This follows from the fact that complex measures are bounded and the remark given above.

We denote the space of complex Radon measures on X by $M(X)$ and for $\mu \in M(X)$ we define

$$\|\mu\| = |\mu|(X),$$

where $|\mu|$ is the total variation of μ.

Theorem 16.4. The Riesz Representation Theorem. *Let X be an LCH space, and for $\mu \in M(X)$ and $f \in C_0(X)$ let $I_\mu(f) = \int f d\mu$. Then the map $\mu \mapsto I_\mu$ is an isometric isomorphism from $M(X)$ to $C_0(X)^*$.*

We prove a partial case of the theorem and find the dual $(C_{\mathbb{R}}([a,b]))^*$ for the space of real valued continuous functions on the interval $[a,b]$.

Signed measures on $\mathbb{R}$ are related to functions of bounded variation. We recall the definition.

Definition 16.5. If $f : [a,b] \to \mathbb{R}$ and $\pi : a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b$ a partition of $[a,b]$ we let $V_\pi(f;[a,b]) = \sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)|$. If $\{V_\pi(f;[a,b])\}_\pi$ is bounded then f is called a **function of bounded variation**, and we denote the space of all such f by BV. The number

$$V(f;[a,b]) = \sup_\pi \sum_{k=0}^{n-1} |f(x_{k+1}) - f(x_k)|$$

is called the **variation** of f.

$V(f;[a,b])$ is denoted by T_f in Folland.

If f is a function of bounded variation one can write it as a difference of two increasing functions: $f(x) = \varphi(x) - \psi(x)$ (take e.g. $\varphi(x) = V(f;[a,x])$) To a function of bounded variation which is right continuous one can associate a signed measure: Let $f = \varphi - \psi$ with increasing functions φ and ψ. If f is right continuous, φ, ψ can be chosen to be right continuous as well. Set $\omega_f(E) = \mu_\varphi(E) - \mu_\psi(E)$ where μ_φ, μ_ψ are the Lebesgue-Stieltjes measures given on the half open intervals $(c,d]$ by $\mu_\varphi((c,d]) = \varphi(d) - \varphi(c)$, $\mu_\varphi((c,d]) = \psi(d) - \psi(c)$. We have $|\omega_f|([a,b]) = V(f;[a,b])$.

Conversely, if μ is a bounded signed Borel measure on $\mathbb{R}$ and $F(x) = \mu(-\infty, x])$ then F is a right continuous function of bounded variation.

Theorem 16.6. *For any $l \in (C_{\mathbb{R}}([a,b]))^*$ there exists a function of bounded variation g such that*

$$l(x) = \int_a^b x(t)dg(t) \tag{16.12}$$

(meaning the Riemann-Stiltjes integral over the interval $[a,b]$) and $V(g;[a,b]) = \|l\|$.

Proof. In what follows we write $C([a,b])$ for $C_{\mathbb{R}}([a,b])$. $C([a,b])$ can be considered as a subspace of the space $B([a,b])$ of all bounded functions on $[a,b]$ with the same norm. Let l be a bounded linear functional on $C([a,b])$. Then by Hahn-Banach theorem it can be extended to a bounded linear functional F on $B([a,b])$ with the same norm. In particular, this extension will be defined on functions $h_\tau(x) = \chi_{[a,\tau]}(x)$, $\tau \in (a,b]$, $h_a(x) = 0$. Let $g(\tau) = F(h_\tau)$. Then g is a function of bounded variation. In fact, let $\pi : a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b$ be a partition of $[a,b]$ and let $\varepsilon_k = \operatorname{sgn}(g(x_k) - g(x_{k-1}))$. Then

$$\begin{aligned}\sum_{k=1}^n |g(x_k) - g(x_{k-1})| &= \sum_{k=1}^n \varepsilon_k(g(x_k) - g(x_{k-1})) = \sum_{k=1}^n \varepsilon_k F(h_{x_k} - h_{x_{k-1}}) \\ &= F(\sum_{k=1}^n \varepsilon_k(h_{x_k} - h_{x_{k-1}})) \leq \|l\| \|\sum_{k=1}^n \varepsilon_k(h_{x_k} - h_{x_{k-1}})\|\end{aligned}$$

But the function $\sum_{k=1}^n \varepsilon_k(h_{x_k} - h_{x_{k-1}})$ takes values 1, -1 and 0 and hence its norm is less or equal 1. Thus

$$\sum_{k=1}^n |g(x_k) - g(x_{k-1})| \leq \|l\|$$

and $V(g;[a,b]) \leq \|l\|$.

We show next the representation (16.12). Let f be a continuous function on $[a,b]$. Given $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(x_1) - f(x_2)| < \varepsilon$ for any x_1, x_2, $|x_1 - x_2| < \delta$. Choose a partition of $[a,b]$ such that the diameter is less than δ. Consider the following step function

$$f_\varepsilon(x) = \begin{cases} f(x_k), & x_{k-1} < x \leq x_k, k > 1 \\ f(x_1), & a \leq x \leq x_1 \end{cases}$$

It can be written in the form

$$f_\varepsilon(x) = \sum_{k=1}^n f(x_k)(h_{x_k}(x) - h_{x_{k-1}}(x)).$$

Clearly, $|f(x) - f_\varepsilon(x)| < \varepsilon$, $x \in [a,b]$, i.e. $\|f - f_\varepsilon\| \leq \varepsilon$.

Consider

$$\begin{aligned} F(f_\varepsilon) &= \sum_{k=1}^{n} f(x_k)(F(h_{x_k}) - F(h_{x_{k-1}})) \\ &= \sum_{k=1}^{n} f(x_k)(g(x_k) - g(x_{k-1})). \end{aligned}$$

Hence making the partition small enough we get

$$|F(f_\varepsilon) - \int_a^b f(x)dg(x)| < \varepsilon.$$

On the other hand

$$|F(f) - F(f_\varepsilon)| \le \|l\| \|f - f_\varepsilon\|_\infty \le \|l\|\varepsilon.$$

Therefore

$$|l(f) - \int_a^b f(x)dg(x)| = |F(f) - F(f_\varepsilon) + F(f_\varepsilon) - \int_a^b f(x)dg(x)| \le \varepsilon(\|l\| + 1)$$

and $l(f) = \int_a^b f(x)dg(x)$.

We have proved already that $V(g; [a, b]) \le \|l\|$. On the other hand,

$$|\int_a^b f(x)dg(x)| \le \|f\|_\infty V(g; [a, b]),$$

i.e. $|l(f)| \le \|f\|_\infty V(g; [a, b])$ and hence $\|l\| \le V(g; [a, b])$ and we get $\|l\| = V(g; [a, b])$.

□

Remark 16.7. Clearly $l_g(f) = \int_a^b f(x)dg(x)$ is a bounded linear functional for any function $g : [a, b] \to \mathbb{R}$ of bounded variation.

It is known that if $g_1 = g_2$ everywhere except for finite or countable number of points in (a, b) then $l_{g_1} = l_{g_2}$.

If $g_1 = g_2 + K$, K is a constant, $l_{g_1} = l_{g_2}$. Conversely if g_1, g_2 define the same functional one can prove that $g_1 - g_2 = \text{const}$ everywhere except for some finite or countable number of points.

Let for $g \in BV([a, b])$, we denote by $[g]$ the set of functions $g_1 \in BV([a, b])$ such that the difference $g_1 - g$ is constant except for finite or countable number of points in (a, b). In each of such subsets one can choose a unique function g_1 such that $g_1(a) = 0$ and right continuous. Therefore we will get a one-to one correspondence between $(C([a, b])^*$ and the set of right continuous functions of bounded variation vanishing in a and therefore with measures w_g defined by such a function g.

The weak* topology on $M(X) = C_0(X)^*$, in which $\mu_n \to \mu$ iff $\int f d\mu_n \to \int f d\mu$ for all $f \in C_0(X)$, is of considerable importance in applications.

Here is a useful criterion for weak* convergence for signed measures in $M([a,b])$. Let NBV denote the space of all $g \in BV$ that are right continuous on $(a,b]$ and such that $g(a) = 0$ (N for "normalized").

Proposition 16.8. *Let $\{g_n\} \subset NBV$, $g \in NBV$. Assume that $\sup_n V(g_n; [a,b]) < \infty$ and $g_n(x) \to g(x)$ for every $x \in [a,b]$. Then $\int_a^b f dg_n \to \int_a^b f dg$ for any $f \in C([a,b])$.*

Proof. Let l_n and l be the linear functionals on $C[a,b])$ corresponding to g_n and g respectively. We must prove that $l_n \to l$ weakly$*$. For this we will use Theorem 14.9 (a characterisation of weak*-convergence). By the Hahn-Banach theorem we can extend l_n and l to bounded linear functionals on

$$F = \operatorname{span}\{C[a,b] \cup \{\chi_{(\alpha,\beta]}, a \leq \alpha < \beta \leq b\}\}$$

equipped with the sup-norm. One can show that

$$l_n(\chi_{(\alpha,\beta]}) = \int \chi_{(\alpha,\beta]}(t) dg_n(t) = g_n(\beta) - g_n(\alpha) \to g(\beta) - g(\alpha).$$

As $\operatorname{span}\{\chi_{(\alpha,\beta]} : a \leq \alpha < \beta \leq b\}$ is dense in F and $\|l_n\| = V(g_n; [a,b]) \leq C < \infty$ for some $C > 0$ and all n, we have $l_n \to l$ weak* in F^* and hence in $C([a,b])^*$. $\square$

Weak topology in $C_0(X)$:

Theorem 16.9. *A sequence $\{f_n\} \in C_0(X)$ converges to $f \in C_0(X)$ weakly if and only if $\sup_{x \in X} \|f_n(x)\| < \infty$ and $f_n(x) \to f(x)$ for any $x \in X$.*

Proof. Assume $f_n \to f$ weakly. Then by Theorem 14.4 $\sup_{x \in X} \|f_n(x)\| < \infty$ and as the point evaluation is a bounded linear functional on $C_0(X)$ we obtain the pointwise convergence of f_n to f. To see the other direction, note that for any $\mu \in M(X)$,

$$|\int_X (f_n(x) - f(x)) d\mu(x)| \leq \int_X |f_n(x) - f(x)| d|\mu|(x) \to 0$$

as $\sup_{x \in X} |f_n(x) - f(x)| < \infty$ and $|\mu|$ is a finite measure. The statement now follows from the Riesz Representation theorem. $\square$

17 Adjoint Operators

Definition 17.1. Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be normed spaces and let $T : X \to Y$ be a bounded linear operator. The *adjoint operator of* T is the linear operator $T^* : Y^* \to X^*$ defined by

$$[T^*(f)](x) = f(Tx)$$

for all $f \in Y^*$ and $x \in X$.

Remark 17.2. One can see that T^*f is the composition of bounded linear operators T and f, so $T^*f \in X^*$.

Theorem 17.3. *Let $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ be normed spaces and let $T : X \to Y$ be a bounded linear operator. Then T^* is a bounded linear operator and $\|T^*\| \leq \|T\|$.*

Proof. (i) For all $f, g \in Y^*$, $\lambda, \mu \in \mathbb{C}$ and $x \in X$,

$$\begin{aligned}
[T^*(\lambda f + \mu g)](x) &= (\lambda f + \mu g)(Tx) \\
&= \lambda f(Tx) + \mu g(Tx) \\
&= \lambda [T^*(f)](x) + \mu [T^*(g)](x) \\
&= [\lambda T^*(f) + \mu T^*(g)](x).
\end{aligned}$$

Thus T^* is linear.

(ii) For all $f \in Y^*$,

$$\begin{aligned}
\|T^*f\|_{X^*} &= \sup_{\|x\|_X \leq 1} |[T^*(f)](x)| \\
&= \sup_{\|x\|_X \leq 1} |f(Tx)| \\
&\leq \sup_{\|x\|_X \leq 1} \|f\|_{Y^*} \|T(x)\|_Y \\
&\leq \sup_{\|x\|_X \leq 1} \|f\|_{Y^*} \|T\| \|x\|_X \\
&= \|f\|_{Y^*} \|T\|.
\end{aligned}$$

Thus T^* is bounded and

$$\begin{aligned}
\|T^*\| &= \sup_{\|f\|_{Y^*} \leq 1} \|T^*f\|_{X^*} \\
&\leq \sup_{\|f\|_{Y^*} \leq 1} \|T\| \|f\|_{Y^*} \\
&= \|T\|.
\end{aligned}$$

$\square$

Remark 17.4. Let H_1, H_2 be Hilbert spaces and let $T : H_1 \to H_2$ be a bounded linear operator. By the Riesz-Fréchet theorem, the *adjoint operator of* T is the unique bounded linear operator $T^* : H_2 \to H_1$ such that

$$\langle x, T^*y\rangle = \langle Tx, y\rangle$$

for all $x \in H_1$ and $y \in H_2$.

Example 17.5. *The adjoint of the shift operator.* Recall that

$$S : \ell^2 \to \ell^2 : (x_1, x_2, \ldots) \mapsto (0, x_1, x_2, \ldots).$$

Find S^*, SS^* and S^*S.

Solution. (i) Since ℓ^2 is a Hilbert space, by the Riesz-Fréchet theorem, S^* is the linear operator $S^* : \ell^2 \to \ell^2$ such that

$$\langle x, S^*y\rangle = \langle Sx, y\rangle$$

for all $x = (x_i)_{i=1}^{\infty}, y = (y_i)_{i=1}^{\infty} \in \ell^2$. Note

$$\begin{aligned}\langle Sx, y\rangle &= \sum_{i=1}^{\infty}(Sx)_i\overline{y}_i \\ &= 0 \cdot \overline{y}_1 + x_1 \cdot \overline{y}_2 + \ldots \\ &= \sum_{i=1}^{\infty} x_i\overline{y}_{i+1} \\ &= \langle x, z\rangle,\end{aligned}$$

where $z = (y_2, y_3, \ldots)$. Therefore, since $\langle x, S^*y\rangle = \langle Sx, y\rangle$, we have

$$\langle x, z\rangle = \langle x, S^*y\rangle$$

for all $x \in \ell^2$. This implies

$$S^*y = z = (y_2, y_3, \ldots).$$

Thus S^* is the *backward shift operator.*

(ii) For all $x = (x_i)_{i=1}^{\infty} \in \ell^2$,

$$\begin{aligned}SS^*(x_1, x_2, \ldots) &= S(x_2, x_3, \ldots) \\ &= (0, x_2, x_3, \ldots).\end{aligned}$$

(iii) For all $x = (x_i)_{i=1}^{\infty} \in \ell^2$,

$$\begin{aligned}S^*S(x_1, x_2, \ldots) &= S^*(0, x_1, x_2, \ldots) \\ &= (x_1, x_2, \ldots).\end{aligned}$$

Thus $S^*S = I_{\ell^2}$, the identity operator on ℓ^2. □

Remark 17.6. For a matrix $A = (a_{ij})_{i,j=1}^{m,n}$ the adjoint of A is the usual transpose conjugate:

$$A^* = (b_{ij})_{i,j=1}^{n,m},$$

where

$$b_{ij} = \overline{a}_{ji}$$

for $i = 1, \ldots n, j = 1, \ldots, m$.

18 Spectrum

If A is an $n \times n$ complex matrix then A has at least one *eigenvalue*: that is, there exists $\lambda \in \mathbb{C}$ such that

$$Ax = \lambda x$$

for some *non-zero* vector x. It is an important tool, and in the infinite dimensional case, the *spectrum* of a linear operator is the analogue of the set of eigenvalues.

Remark 18.1. The shift operator S on ℓ^2:

$$S(x_1, x_2, \ldots) = (0, x_1, x_2, \ldots)$$

has *no eigenvalues.*

Solution. Let there be $\lambda \in \mathbb{C}$ such that

$$Sx = \lambda x$$

for some $x \in \ell^2$. Hence

$$(0, x_1, x_2, \ldots) = (\lambda x_1, \lambda x_2, \ldots),$$

so $\lambda x_1 = 0, \lambda x_2 = x_1, \ldots, \lambda x_n = x_{n-1}, \ldots$. Therefore $\lambda = 0$ or $x_1 = 0$. If $\lambda = 0$, then $x_n = \lambda x_{n+1} = 0$ for all $n \geq 1$, and so $x = 0$. If $x_1 = 0$ and $\lambda \neq 0$, then

$$\begin{aligned} x_2 &= \frac{x_1}{\lambda} = 0, \\ x_3 &= \frac{x_2}{\lambda} = 0, \\ &\ldots \\ x_n &= \frac{x_{n-1}}{\lambda} = 0, \\ &\ldots. \end{aligned}$$

Thus $x = 0$. Hence $Sx = \lambda x$ implies $x = 0$, and so S has no eigenvalues. $\square$

Definition 18.2. Let $(E, \|\cdot\|_E)$, $(F, \|\cdot\|_F)$ be normed spaces. A bounded linear operator $T : E \to F$ is called *invertible* if there exists a bounded linear operator $S : F \to E$ such that $TS = I_F$ and $ST = I_E$ where I_F, I_E are the identity operators on F and E respectively. The operator S is called the *inverse* of T, denoted by T^{-1}.

Theorem 18.3. *Let $T : E \to F$ be invertible. Then T is injective and surjective.*

Proof. (i) Suppose $x_1, x_2 \in E$ are such that $x_1 \neq x_2$ and $Tx_1 = Tx_2$. Then $x_1 = T^{-1}Tx_1 = T^{-1}Tx_2 = x_2$. Thus $x_1 \neq x_2$ implies $Tx_1 \neq Tx_2$, so T is injective.

(ii) For any $y \in F$ there exists $x = T^{-1}y$ such that

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

Thus T is surjective. $\square$

Example 18.4. The *backward shift* operator $T : \ell^2 \to \ell^2$ given by

$$T(x_1, x_2, x_3, \ldots) = (x_2, x_3, \ldots)$$

is *not injective*, because $(1, 0, \ldots) \neq (0, 0, \ldots)$ but $T(1, 0, \ldots) = (0, 0, \ldots) = T(0, 0, \ldots)$. Thus T is *not invertible*.

Definition 18.5. Let $T : X \to X$ be a bounded linear operator on a Banach space X. The *spectrum* $\operatorname{Sp} T$ of T is the set

$$\{\lambda \in \mathbb{C} : \lambda I_X - T \text{ is not invertible }\}.$$

Example 18.6. Consider the Banach space $(C[0,1], \|\cdot\|_\infty)$ and the bounded linear operator

$$T : C[0,1] \to C[0,1]$$

defined by the formula

$$(Tf)(t) = g(t)f(t)$$

where $g \in C[0,1]$. Find the spectrum $\operatorname{Sp} T$ of T.

Solution. By the definition

$$\operatorname{Sp} T = \{\lambda \in \mathbb{C} : \lambda I_X - T \text{ is not invertible }\}.$$

Note that, for all $\lambda \in \mathbb{C}$,

$$T_\lambda f \stackrel{\text{def}}{=} (\lambda I_{C[0,1]} - T)(f) = \lambda f - Tf,$$

so

$$\begin{aligned} T_\lambda f(t) &= [(\lambda I_{C[0,1]} - T)(f)](t) \\ &= \lambda f(t) - g(t)f(t) \\ &= (\lambda - g(t))f(t) \end{aligned}$$

for all $f \in C[0,1]$ and $t \in [0,1]$.

1. For all $\lambda \in \mathbb{C}$ such that

$$\lambda \notin \operatorname{Im} g = \{z \in \mathbb{C} : z = g(t) \text{ for some } t \in [0,1]\}$$

the function $\lambda - g(t) \neq 0$ as $t \in [0,1]$. Thus, for all $\lambda \notin \operatorname{Im} g$, we can define a mapping

$$S_\lambda : C[0,1] \to C[0,1]$$

by the formula

$$(S_\lambda f)(t) = \frac{1}{\lambda - g(t)} f(t).$$

We proved before that S_λ is a bounded linear operator and

$$\|S_\lambda\| = \left\| \frac{1}{\lambda - g(t)} \right\|_\infty .$$

Let us show that S_λ is the *inverse* of the operator T_λ. For all $f \in C[0,1]$ and $t \in [0,1]$,

$$\begin{aligned}[S_\lambda T_\lambda f](t) &= [S_\lambda(T_\lambda f)](t) \\ &= \frac{1}{\lambda - g(t)}(T_\lambda f)(t) \\ &= \frac{1}{\lambda - g(t)}(\lambda - g(t)) f(t) = f(t)\end{aligned}$$

and

$$\begin{aligned}[T_\lambda S_\lambda f](t) &= [T_\lambda(S_\lambda f)](t) \\ &= (\lambda - g(t))(S_\lambda f)(t) \\ &= (\lambda - g(t))\frac{1}{\lambda - g(t)} f(t) = f(t).\end{aligned}$$

Thus $S_\lambda T_\lambda = I_{C[0,1]}$ and $T_\lambda S_\lambda = I_{C[0,1]}$, so T_λ *is invertible for all* $\lambda \notin \operatorname{Im} g$.

2. For *every* $\lambda \in \operatorname{Im} g$, the operator T_λ is *not surjective*, because, for any $f \in C[0,1]$, the function

$$(T_\lambda f)(t) = (\lambda - g(t)) f(t)$$

is equal to zero at a point t_λ such that $\lambda = g(t_\lambda)$. Therefore T_λ is *not invertible* for all $\lambda \in \operatorname{Im} g$.

3. The spectrum $\operatorname{Sp} T = \operatorname{Im} g$. □

Lemma 18.7. *Let $T : X \to X$ be a bounded linear operator on a Banach space X. Then every eigenvalue λ of T belongs to $\operatorname{Sp} T$.*

Proof. By the assumption, there exists a non-zero vector $x \in X$ such that

$$Tx = \lambda x.$$

Hence $(\lambda I_X - T)x = \lambda x - Tx = \underline{0}$ and $(\lambda I_X - T)\underline{0} = \underline{0}$. Therefore $\lambda I_X - T$ is not injective, so $\lambda I_X - T$ is not invertible. Thus $\lambda \in \operatorname{Sp} T$. $\square$

Remark 18.8. The spectrum is *larger* than the *set of eigenvalues*.

Example 18.9. Let $T : C[0,1] \to C[0,1]$ be the bounded linear operator defined by

$$(Tf)(t) = tf(t), \quad t \in [0,1],$$

on the Banach space $(C[0,1], \|\cdot\|_\infty)$.

1. We have proved that $\operatorname{Sp} T = [0,1]$.

2. *T has no eigenvalues.*

Proof. If λ is an eigenvalue of T with eigenvector $f \in C[0,1]$, then $(Tf)(t) = \lambda f(t)$ for all $t \in [0,1]$, so $tf(t) = \lambda f(t)$, and therefore

$$(t - \lambda)f(t) = 0$$

for all $t \in [0,1]$. Thus $f(t) = 0$ for all t, and hence T has no eigenvalues. $\square$

Notation 18.10. Let X be a normed space. We shall abbreviate $B(X,X)$ to $B(X)$. For $T \in B(X)$ we denote $TT, TTT, \ldots$ by $T^2, T^3, \ldots$ etc.

Exercise 18.11. Show that, for any $n \in \mathbb{N}$,

$$\|T^n\| \leq \|T\|^n.$$

Theorem 18.12. *Let $T : X \to X$ be a bounded linear operator on a Banach space X. If $\|T\| < 1$ then the operator $I_X - T$ is invertible and*

$$(I_X - T)^{-1} = \sum_{n=0}^{\infty} T^n = \lim_{n\to\infty} (I_X + T + T^2 + \ldots T^n)$$

in the Banach space $B(X)$.

Proof. Let $S_n = I_X + T + \ldots + T^n$. We have

$$\|S_n - S_m\| = \| \sum_{k=n+1}^{m} T^k \| \leq \sum_{k=n+1}^{m} \|T\|^k.$$

As $\|T\| < 1$, the series $\sum_{k=0}^{\infty} \|T\|^k$ is convergent and hence $\{S_n\}_n$ is a Cauchy sequence. Since $B(X)$ is a Banach space, $\{S_n\}$ has a limit $S \in B(X)$. It is easy to see that

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

Letting $n \to \infty$, we obtain $(I_X - T)S = I_X = S(I_X - T)$. Thus $I_X - T$ is invertible with the inverse $S = \lim_{n\to\infty}(I_X + T + \ldots + T^n)$. □

Corollary 18.13. *Let $T : X \to X$ be a bounded linear operator on a Banach space X. Then*

$$\operatorname{Sp} T \subset \{\lambda \in \mathbb{C} : |\lambda| \leq \|T\|\}.$$

Proof. Note that, for λ such that $|\lambda| > \|T\|$,

$$\left\| \frac{1}{\lambda} T \right\| = \frac{1}{|\lambda|} \|T\| < 1.$$

Thus the operator $(I_X - \frac{1}{\lambda}T)$ is invertible, so there exists a bounded linear operator $S : X \to X$ such that

$$S(I_X - \frac{1}{\lambda}T) = I_X = (I_X - \frac{1}{\lambda}T)S.$$

Then

$$\left(\frac{1}{\lambda}S\right)(\lambda I_X - T) = I_X = (\lambda I_X - T)\left(\frac{1}{\lambda}S\right).$$

Thus $\frac{1}{\lambda}S$ is the inverse of $(\lambda I_X - T)$, so $(\lambda I_X - T)$ is invertible for all λ such that $|\lambda| > \|T\|$. Therefore

$$\operatorname{Sp} T \subset \{\lambda \in \mathbb{C} : |\lambda| \leq \|T\|\}.$$

□

Theorem 18.14. *Let $T : X \to X$ be a bounded linear operator on a Banach space X. Then* $\operatorname{Sp} T$ *is a* closed bounded non-empty subset *of* $\mathbb{C}$ *and is contained in the closed disc* $\{\lambda \in \mathbb{C} : |\lambda| \leq \|T\|\}$.

Let's first look at some examples.

Example 18.15. Let T be the left shift operator on ℓ^2.

(i) Show that vectors $(1, \lambda, \lambda^2, \lambda^3, \ldots)$, where $\lambda \in \mathbb{C}$ and $|\lambda| < 1$, are eigenvectors of T.

(ii) Find the spectrum $\operatorname{Sp} T$ of T.

Solution. (i) For $\lambda \in \mathbb{C}$ such that $|\lambda| < 1$, the series

$$\sum_{n=0}^{\infty} |\lambda^n|^2 = \sum_{n=0}^{\infty} |\lambda|^{2n} < \infty.$$

Thus $x_\lambda = (1, \lambda, \lambda^2, \lambda^3, \ldots) \in \ell^2$. One can see that

$$\begin{aligned} Tx_\lambda &= T(1, \lambda, \lambda^2, \ldots) \\ &= (\lambda, \lambda^2, \lambda^3, \ldots) \\ &= \lambda(1, \lambda, \lambda^2, \ldots) \\ &= \lambda x_\lambda. \end{aligned}$$

Therefore, for $\lambda \in \mathbb{C}$ such that $|\lambda| < 1$, x_λ is an eigenvector with eigenvalue λ.
(ii) We know that every eigenvalue of T belongs to $\operatorname{Sp} T$. Hence

$$\{\lambda \in \mathbb{C} : |\lambda| < 1\} \subset \operatorname{Sp} T.$$

We proved above that

$$\operatorname{Sp} T \subset \{\lambda \in \mathbb{C} : |\lambda| \leq \|T\|\}.$$

Recall that $\|T\| = 1$ and $\operatorname{Sp} T$ is a closed subset of $\mathbb{C}$. Therefore

$$\operatorname{Sp} T = \{\lambda \in \mathbb{C} : |\lambda| \leq \|T\|\}.$$

$\square$

Example 18.16. Volterra Integral Operator. Let $V : C([0,1]) \to C([0,1])$ be a linear operator given by

$$(Vf)(s) = \int_0^s f(r)dr.$$

Then $\operatorname{Sp} V = \{0\}$.

Proof. Using induction and integration by parts one can prove that

$$(V^n f)(s) = \frac{1}{(n-1)!}\int_0^s (s-r)^{n-1} f(r)dr.$$

Hence for any $s \in [0,1]$ we have

$$|(V^n f)(s)| \leq \frac{1}{(n-1)!}\int_0^s (s-r)^n \|f\|_\infty dr \leq \frac{\|f\|_\infty}{n!}.$$

Thus $\|V^n f\|_\infty \leq \|x\|_\infty/n!$ and $\|V^n\| \leq 1/n!$ that gives $\lim \|V^n\|^{1/n} = 0$. If $\lambda \in \operatorname{Sp} V$ then $\lambda^n \in \operatorname{Sp} V^n$. Hence $|\lambda|^n \leq \|V^n\|$, $|\lambda| \leq \|V^n\|^{1/n}$ and therefore $\lambda = 0$. $\square$

To prove Theorem 18.14 we need the following lemmas.

Lemma 18.17. *Let $T : X \to X$ be a bounded linear operator on a Banach space X. Then $\operatorname{Sp} T$ is closed.*

Proof. Let $\rho(T) = \{\lambda \in \mathbb{C} : \lambda I - T \text{ is invertible}\}$. As $\rho(T) = \operatorname{Sp} T^c$ we must see that $\rho(T)$ is closed. Let $\lambda \in \rho(T)$ and $\mu \in \mathbb{C}$ such that $|\lambda - \mu| < \|(\lambda - T)^{-1}\|^{-1}$. Then

$$\mu I - T = (\mu - \lambda)I + (\lambda I - T) = (\lambda I - T)(I + (\mu - \lambda)(\lambda I - T)^{-1}).$$

As $\|(\mu - \lambda)(\lambda I - T)^{-1}\| = |\mu - \lambda|\|(\lambda I - T)^{-1}\| < 1$, by Theorem 18.12, $I + (\mu - \lambda)(\lambda I - T)^{-1}$ is invertible. As product of two invertible operators is invertible we obtain that $\mu I - T$ is invertible and hence

$$B(\lambda, \|(\lambda I - T)^{-1}\|^{-1}) = \{\mu \in \mathbb{C} : |\mu - \lambda| < \|(\lambda I - T)^{-1}\|^{-1} \subset \rho(T).$$

The proof is done. $\square$

Lemma 18.18. *The map $\lambda \mapsto (\lambda I - T)^{-1}$ is differentiable on $\rho(T)$, i.e. for any $\lambda \in \rho(T)$ there exists a bounded linear operator $U_{\lambda,T}$ such that*

$$\lim_{h \to 0} \frac{\|((\lambda + h)I - T)^{-1}\| - (\lambda - T)^{-1} - U_{\lambda,T}h}{|h|} = 0.$$

Proof. Let $\lambda \in \rho(T)$. By the previous lemma, $\rho(T)$ is open and $B(\lambda, r) \subset \rho(T)$ for some $r > 0$. Take $U_{\lambda,T} = -(\lambda I - T)^{-2}$ and $h \in \mathbb{C}$ such that $|h| < \max((\lambda I - T)^{-1}\|^{-1}, r)$. Then

$$\begin{aligned}
&((\lambda + h)I - T)^{-1} - (\lambda I - T)^{-1} + (\lambda I - T)^{-2}h \\
&= (\lambda I - T)^{-1}(I + h(\lambda I - T)^{-1})^{-1} - (\lambda I - T)^{-1} + (\lambda I - T)^{-2}h
\end{aligned}$$

As $\|h(\lambda I - T)^{-1}\| < 1$, by Theorem 18.12, $I + h(\lambda I - T)^{-1}$ is invertible with the inverse $(I + h(\lambda I - T)^{-1})^{-1} = \sum_{k=0}^{\infty} h^k(\lambda I - T)^{-k}$. Hence

$$
\begin{aligned}
&((\lambda+h)I - T)^{-1} - (\lambda I - T)^{-1} + (\lambda I - T)^{-2}h \\
&= \sum_{k=0}^{\infty} h^k(\lambda I - T)^{-k-1} - (\lambda I - T)^{-1} + (\lambda I - T)^{-2}h \\
&= \sum_{k=2}^{\infty} h^k(\lambda I - T)^{-k-1} = h^2 \sum_{k=0}^{\infty} h^k(\lambda I - T)^{-k-3}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
&\lim_{h\to 0} \frac{(\|(\lambda+h)I - T)^{-1} - (\lambda I - T)^{-1} + (\lambda I - T)^{-2}h\|}{|h|} \\
&= \lim_{h\to 0} \frac{|h^2|\|\sum_{k=0}^{\infty} h^k(\lambda I - T)^{-k-3}\|}{|h|} = 0.
\end{aligned}
$$

□

Proof of Theorem 18.14. From Lemma 18.17 and Corollary 18.13, we have that the spectrum is a closed subset of $\{\lambda \in \mathbb{C} : |\lambda| \le \|T\|\}$. We must only see that $\operatorname{Sp} T$ is non-empty. Assume contrary that $\operatorname{Sp} T = \emptyset$. By Lemma 18.18, $\lambda \to (\lambda I - T)^{-1}$ is differentiable on $\mathbb{C}$ and hence continuous. As

$$
|\|(\lambda I - T)^{-1}\| - \|(\mu I - T)^{-1}\|| \le \|(\lambda I - T)^{-1} - (\mu I - T)^{-1}\|,
$$

one can easily see that the scalar function $\lambda \mapsto \|(\lambda I - T)^{-1}\|$ is continuous on $\mathbb{C}$. Thus the latter is bounded on the compact subset $\{\lambda \in \mathbb{C} : |\lambda| \le 2\|T\|\}$. For λ, $|\lambda| > 2\|T\|$, we have

$$
\begin{aligned}
\|(\lambda I - T)^{-1}\| &= |\lambda|^{-1}\|\|(1 - \lambda^{-1}T)^{-1}\| \\
&\le (\text{ as } \|\lambda^{-1}T\| < 1/2) = |\lambda|^{-1}\| \sum_{k=0}^{\infty} \lambda^{-k}T^k\| \\
&\le |\lambda|^{-1} \sum_{k=0}^{\infty} |\lambda|^{-k}\|T\|^k = |\lambda|^{-1} \frac{1}{1 - |\lambda|^{-1}\|T\|} \\
&< \frac{|\lambda|^{-1}}{2} \le \|T\|^{-1}
\end{aligned}
$$

Thus the function $\lambda \mapsto \|(\lambda I - T)^{-1}\|$ is bounded on the whole complex plane.

Take now $f \in X^*$, $\xi \in X$. Then the function $\psi : \mathbb{C} \to \mathbb{C}$, $\lambda \to f(\lambda I - T)^{-1}\xi)$ is bounded. In fact, by the previous argument, we have

$$
|f(\lambda I - T)^{-1}\xi)| \le \|f\|\|(\lambda I - T)^{-1}\xi\| \le \|f\|\|(\lambda I - T)^{-1}\|\|\xi\| < C
$$

for some $C > 0$. Moreover, as $\lambda \to (\lambda I - T)^{-1}$ is differentiable on $\mathbb{C}$ one can easily see that so is $\psi : \mathbb{C} \to \mathbb{C}$. By Liouville's theorem from complex analysis, we have that ψ is a constant function. In particular, $\psi(0) = \psi(1)$, i.e. $f(-T^{-1}\xi) = f((I-T)^{-1}\xi)$. As this holds for any $f \in X^*$ and $\xi \in X$ we obtain $-T^{-1} = (I-T)^{-1}$ and hence $-T = I - T$ which is impossible. Thus $\operatorname{Sp} T$ is non-empty.