Functional Analysis

Bengt Ove Turesson[1]

December 27, 2009

[1]Matematiska institutionen, Linköpings universitet, SE-581 83 Linköping, Sverige

Preface

Text

Bengt Ove Turesson

Linköping December 27, 2009

Contents

Chapter 1

Metric Spaces

1.1. Definition and First Examples

Definition 1.1.1. A **metric space** consists of a set X and a real-valued function d on $X \times X$ such that for any $x, y, z \in X$, the following properties hold:

(i) d is **non-negative**: $d(x, y) \geq 0$;

(ii) d is **definite**: $d(x, y) = 0$ if and only if $x = y$;

(iii) d is **symmetric**: $d(x, y) = d(y, x)$;

(iv) d satisfies the **triangle inequality**: $d(x, z) \leq d(x, y) + d(y, z)$.

The function d is called a **metric**.

We remark that the conditions (i)–(iv) are not independent; in fact, one can show that (iv) implies (i) and that (ii) and (iv) imply (iii).

We next give several examples of metric spaces that will reoccur throughout the text.

Example 1.1.2. From linear algebra, we know that

$$d(x, y) = |x - y| = \left(\sum_{j=1}^{d} |x^j - y^j|^2 \right)^{1/2},$$

where $x = (x^1, \dots, x^d)^t$ and $y = (y^1, \dots, y^d)^t$, is a metric on $\mathbf{R}^d$. This definition in fact also gives a metric on $\mathbf{C}^d$ (the space of d by 1 matrices with complex entries). We will use this as the standard metric on $\mathbf{R}^d$ and $\mathbf{C}^d$. In Example 1.2.4 and Example 1.4.2, we introduce other metrics on $\mathbf{R}^d$ and $\mathbf{C}^d$. □

Example 1.1.3. Let X be an arbitrary set. The **discrete metric** d on X is defined $d(x, y) = 0$ if $x = y$ and $d(x, y) = 1$ otherwise. It is easily verified that d satisfies the conditions (i)–(iv) in the definition. With this metric, we call X a **discrete metric space**. □

Using the following more or less obvious result, it is easy to construct new metric spaces from given spaces.

Proposition 1.1.4. *Let X be a metric space with metric d and Y a subset to X. Then Y is a metric space in itself with the restriction of d as the metric.*

We call Y with the metric $d|_Y$ a (**metric**) **subspace** to X and the metric $d|_Y$ the **induced metric**. When there is no risk of confusion, we shall denote the metric on Y by just d instead of $d|_Y$.

Example 1.1.5. Being a subset of $\mathbf{R}$, $\mathbf{Q}$ is a metric space with the metric d_2. □

1.2. Hölder's and Minkowski's inequalities

Let $1 < p < \infty$. We denote by p' the number defined by

$$\frac{1}{p} + \frac{1}{p'} = 1, \quad \text{that is} \quad p' = \frac{p}{p-1}.$$

Obviously, $1 < p' < \infty$. We also define $1' = \infty$ and $\infty' = 1$. Notice that this is consistent with the limits obtained by letting $p \to 1$ and $p \to \infty$ in the definition of p'. The number p' is called the **dual exponent** to p.

Lemma 1.2.1. *Let $x, y \geq 0$ and $1 < p < \infty$. Then*

$$xy \leq \frac{x^p}{p} + \frac{y^{p'}}{p'}. \tag{1}$$

Proof. Substituting $x^p = \exp(s)$, $y^{p'} = \exp(t)$ in (1), we see that the inequality is equivalent to

$$\exp\Big(\frac{s}{p} + \frac{t}{p'}\Big) \leq \frac{\exp(s)}{p} + \frac{\exp(t)}{p'},$$

which in turn follows from the fact that the exponential function is convex. ■

Theorem 1.2.2 (Hölder's inequality for $\mathbf{C}^d$). *Suppose that $x, y \in \mathbf{C}^d$. Then, for $1 < p < \infty$,*

$$\sum_{j=1}^{d} |x^j y^j| \leq \Big(\sum_{j=1}^{d} |x^j|^p\Big)^{1/p} \Big(\sum_{j=1}^{d} |y^j|^{p'}\Big)^{1/p'}. \tag{2}$$

In the case $p = 2$, Hölder's inequality is often called the **Cauchy–Schwarz inequality**.

Proof. Notice that if (2) is true for some x and y, then it is true for λx and μy, where $\lambda, \mu \in \mathbf{C}$. We may thus assume that $\sum_{j=1}^{d} |x^j|^p = \sum_{j=1}^{d} |y^j|^{p'} = 1$. It then follows from Lemma 1.2.1 that

$$\sum_{j=1}^{d} |x^j y^j| \leq \frac{1}{p}\sum_{j=1}^{d} |x^j|^p + \frac{1}{p'}\sum_{j=1}^{d} |y^j|^{p'} = 1 = \Big(\sum_{j=1}^{d} |x^j|^p\Big)^{1/p} \Big(\sum_{j=1}^{d} |y^j|^{p'}\Big)^{1/p'}. \quad ■$$

Theorem 1.2.3 (Minkowski's inequality $\mathbf{C}^d$). *Suppose that $x, y \in \mathbf{C}^d$. Then, for $1 \leq p < \infty$,*

$$\Big(\sum_{j=1}^{d} |x^j + y^j|^p\Big)^{1/p} \leq \Big(\sum_{j=1}^{d} |x^j|^p\Big)^{1/p} + \Big(\sum_{j=1}^{d} |y^j|^p\Big)^{1/p}. \tag{3}$$

Proof. The inequality (3) is obviously true if $p = 1$. We can therefore assume that $1 < p < \infty$. By Hölder's inequality,

$$\begin{aligned}\sum_{j=1}^{d} |x^j + y^j|^p &\leq \sum_{j=1}^{d} |x^j + y^j|^{p-1}|x^j| + \sum_{j=1}^{d} |x^j + y^j|^{p-1}|y^j| \\ &\leq \Big(\sum_{j=1}^{d} |x^j + y^j|^{(p-1)p'}\Big)^{1/p'} \Big(\Big(\sum_{j=1}^{d} |x^j|^p\Big)^{1/p} + \Big(\sum_{j=1}^{d} |y^j|^p\Big)^{1/p}\Big).\end{aligned}$$

All there remains is to use the fact that $(p-1)p' = p$ and divide the left and the right member by $(\sum_{j=1}^{d} |x^j + y^j|^p)^{1/p'}$. ■

Example 1.2.4. A generalization of the metric d_2 on $\mathbf{R}^d$ and $\mathbf{C}^d$ is

$$d_p(x, y) = \left(\sum_{j=1}^{d} |x^j - y^j|^p \right)^{1/p},$$

where $1 \le p < \infty$. It is clear that d_p is definite and symmetric, and the triangle inequality follows from Theorem 1.2.3. □

1.3. ℓ^p-spaces

Definition 1.3.1. For $1 \le p < \infty$, the set ℓ^p consists of all sequences $(x^j)_{j=1}^{\infty}$ of complex numbers such that $\sum_{j=1}^{\infty} |x^j|^p < \infty$.

Suppose that $x = (x^j)_{j=1}^{\infty} \in \ell^p$ and $y = (y^j)_{j=1}^{\infty} \in \ell^p$. Then, $\alpha x = (\alpha x^j)_{j=1}^{\infty} \in \ell^p$ for every number $\alpha \in \mathbf{C}$. Moreover, since

$$|x^j + y^j|^p \le 2^p(|x^j|^p + |y^j|^p) \quad \text{for every } j,$$

it follows that $x + y = (x^j + y^j)_{j=1}^{\infty} \in \ell^p$. This shows that ℓ^p is a vector space over $\mathbf{C}$. If we let d tend to ∞ in Theorem 1.2.2 and Theorem 1.2.3, we obtain Hölder's and Minkowski's inequalities for series.

Corollary 1.3.2 (Hölder's inequality for ℓ^p). *Suppose that $x \in \ell^p$ and $y \in \ell^{p'}$, where $1 < p < \infty$. Then*

$$\sum_{j=1}^{\infty} |x^j y^j| \le \left(\sum_{j=1}^{\infty} |x^j|^p \right)^{1/p} \left(\sum_{j=1}^{\infty} |y^j|^{p'} \right)^{1/p'}.$$

Corollary 1.3.3 (Minkowski's inequality for ℓ^p). *Suppose that $x, y \in \ell^p$, where $1 \le p < \infty$. Then*

$$\left(\sum_{j=1}^{\infty} |x^j + y^j|^p \right)^{1/p} \le \left(\sum_{j=1}^{\infty} |x^j|^p \right)^{1/p} + \left(\sum_{j=1}^{\infty} |y^j|^p \right)^{1/p}.$$

Example 1.3.4. It follows from Minkowski's inequality that

$$d_p(x, y) = \left(\sum_{j=1}^{\infty} |x^j - y^j|^p \right)^{1/p}, \quad x, y \in \ell^p,$$

is a metric on ℓ^p for $1 \le p < \infty$. □

1.4. ℓ^∞-spaces

Definition 1.4.1. We denote by $\ell^\infty(M)$ the set of all bounded, complex-valued functions on an arbitrary set M.

For $f, g \in \ell^\infty(M)$, we define αf, where $\alpha \in \mathbf{C}$, and $f+g$ pointwise:

$$(\alpha f)(x) = \alpha f(x) \quad \text{and} \quad (f+g)(x) = f(x) + g(x), \quad x \in M.$$

It is obvious that $\alpha f \in \ell^\infty(M)$ and $f + g \in \ell^\infty(M)$, so that ℓ^∞ is a vector space over $\mathbf{C}$. A metric on $\ell^\infty(M)$ is defined by

$$d_\infty(f, g) = \sup_{x \in M} |f(x) - g(x)|.$$

To prove the triangle inequality, suppose that $f, g, h \in \ell^\infty(M)$. Then

$$|f(x) - h(x)| \leq |f(x) - g(x)| + |g(x) - h(x)| \leq d_\infty(f, g) + d_\infty(g, h)$$

for every $x \in M$, which implies that $d_\infty(f, h) \leq d_\infty(f, g) + d_\infty(g, h)$.

Let $\ell^\infty(M, \mathbf{R})$ denote the set of all bounded, real-valued functions on M. It is clear that $\ell^\infty(M, \mathbf{R})$ is a subspace to $\ell^\infty(M, \mathbf{R})$ both as a metric space and as a vector space.

Example 1.4.2. If $M = \{1, 2, \dots, d\}$, then $\ell^\infty(M)$ can be identified with $\mathbf{C}^d$. Thus, another metric on $\mathbf{C}^d$ is

$$d_\infty(x, y) = \max_{1 \leq j \leq d} |x^j - y^j|.$$

Since $\mathbf{R}^d$ is a subspace to $\mathbf{C}^d$, this expression defines a metric on $\mathbf{R}^d$ too. □

Example 1.4.3. The space $\ell^\infty(\mathbf{N})$ consists of all bounded sequences $(x^j)_{j=1}^\infty$ of complex numbers; we will denote this space by ℓ^∞. The metric on ℓ^∞ is

$$d_\infty(x, y) = \sup_{1 \leq j < \infty} |x^j - y^j|.$$

The subspaces $\mathbf{c}$ and $\mathbf{c}_0$ to ℓ^∞ consist of all sequences $(x^j)_{j=1}^\infty \in \ell^\infty$ such that the limit $\lim_{j\to\infty} x^j$ exists and $\lim_{j\to\infty} x^j = 0$, respectively. □

1.5. Further Examples

Example 1.5.1. The space $C_b(M)$ consists of all bounded, continuous functions on a set $M \subset \mathbf{R}^d$. The metric on $C_b(M)$ is the same as in $\ell^\infty(M)$. If $K \subset \mathbf{R}^d$ is compact (i.e., closed and bounded), then all continuous functions on K are bounded. We will use the notation $C(K)$ for $C_b(K)$. In the case $K = [a, b]$, we will usually write $C[a, b]$ instead of $C([a, b])$. □

Example 1.5.2. Let E be a measurable subset of $\mathbf{R}^d$. For $1 \leq p \leq \infty$, $L^p(E)$ is a metric space with the metric

$$d_p(f, g) = \left(\int_E |f(x) - g(x)|^p \, dx \right)^{1/p} \quad \text{for } 1 \leq p < \infty$$

and

$$d_\infty(f, g) = \operatorname*{ess\,sup}_{x \in E} |f(x) - g(x)| \quad \text{for } p = \infty.$$

□

1.6. A Reverse Triangle Inequality

We close this chapter by proving a **reverse triangle inequality**:

Proposition 1.6.1. *If X is a metric space, then*

$$d(x,y) \geq |d(x,z) - d(z,y)| \quad \textit{for all } x,y,z \in X.$$

Proof. It follows from (iii) and (iv) in Definition 1.1.1 that

$$d(x,z) \leq d(x,y) + d(y,z) = d(x,y) + d(z,y), \quad \text{so that} \quad d(x,y) \geq d(x,z) - d(z,y),$$

and

$$d(z,y) \leq d(z,x) + d(x,y) = d(x,z) + d(x,y), \quad \text{so that} \quad d(x,y) \geq d(z,y) - d(x,z).$$

Notice finally that $|d(x,z) - d(z,y)|$ equals $d(x,z) - d(z,y)$ or $d(z,y) - d(x,z)$. ∎

Exercises

E1.1. In Definition 1.1.1, show that (iv) implies (i) and that (ii) and (iv) imply (iii).

Chapter 2

Topological Concepts in Metric Spaces

In this chapter, X will always denote a metric space with metric d.

2.1. Open Sets

Definition 2.1.1. If $x \in X$ and $r > 0$, the **open ball** with center x and radius r is the set

$$B_r(x) = \{y \in X : d(x, y) < r\}.$$

The corresponding **closed ball** is $\overline{B}_r(x) = \{y \in X : d(x, y) \leq r\}$.

Example 2.1.2. If $x \in \mathbf{R}$ and $r > 0$, then

$$B_r(x) = (x - r, x + r) \quad \text{and} \quad \overline{B}_r(x) = [x - r, x + r]. \qquad \square$$

Example 2.1.3. Let X be a discrete metric space (see Example 1.1.3). If $r < 1$, then $B_r(x) = \overline{B}_r(x) = \{x\}$, if $r = 1$, then $B_r(x) = \{x\}$ and $\overline{B}_r(x) = X$, and if $r > 1$, then $B_r(x) = \overline{B}_r(x) = X$. $\square$

Definition 2.1.4. A subset G to X is **open** if for every $x \in G$ there exists an open ball $B_r(x)$ such that $B_r(x) \subset G$. A **neighbourhood** of $x \in X$ is an open set G such that $x \in G$.

Example 2.1.5. An open ball $B_r(x) \subset X$ is of course open. Indeed, suppose that $y \in B_r(x)$ and $y \neq x$. Then the ball $B_s(y)$, where $s = r - d(x, y) > 0$, is a subset of $B_r(x)$, since if $z \in B_s(y)$, then

$$d(x, z) \leq d(x, y) + d(y, z) < d(x, y) + s = r. \qquad \square$$

Example 2.1.6. By Example 2.1.5, every interval of the form $(a, b) \subset \mathbf{R}$, where $-\infty < a < b < \infty$, is open (take x as the midpoint of the interval and r as half its length). As an exercise, show that all intervals $(-\infty, b)$ and (a, ∞) are open. $\square$

Example 2.1.7. If X is a discrete metric space and E a subset to X, then for every $x \in E$, $B_{1/2}(x) = \{x\} \subset E$. This shows that every subset to X is open. $\square$

The following theorem shows that every metric space is also a *topological space*.

Theorem 2.1.8. *The collection τ of open subsets of X is a topology on X:*

(i) $X, \varnothing \in \tau$;

(ii) *if $G_\alpha \in \tau$ for every $\alpha \in A$, then $\bigcup_{\alpha \in A} G_\alpha \in \tau$;*

(iii) *if $G_1 \ldots, G_n \in \tau$, then $\bigcap_{j=1}^{n} G_j \in \tau$.*

We remark that the index set A in (ii) may be infinite and even uncountable.

Proof. The first statement in the theorem is obvious; we prove the second and leave the third as an exercise. If $x \in \bigcup_{\alpha \in A} G_\alpha$, then x belongs to some set G_{α_0}, and since G_{α_0} is open, $B_r(x) \subset G_{\alpha_0}$ for some $r > 0$. But from this it follows that $B_r(x) \subset \bigcup_{\alpha \in A} G_\alpha$, which shows that the union is open. ■

Remark 2.1.9. We will later show that the metrics d_2, d_p, and d_∞ all give rise to the same topology, i.e., the same open sets, on $\mathbf{R}^d$ and $\mathbf{C}^d$.

Proposition 2.1.10. *A subset G of X is open if and only if it is the union of a (possibly empty) collection of open balls.*

Proof. First, suppose that G is open. Then, for every $x \in G$, there exists a ball $B_{r_x}(x) \subset G$. The union of these balls, as x varies over G, equals G. The converse follows directly from Theorem 2.1.8. ■

2.2. The Interior

Definition 2.2.1. The **interior** E° of a subset E to X is the union of all open subsets to E. The elements of E° are called **interior points**.

Property (a) in the following corollary follows from (ii) in Theorem 2.1.8; property (b) follows from (a).

Corollary 2.2.2. *Let E be a subset of X. Then*

(a) *E° is the largest open subset of E;*

(b) *E is open if and only if $E = E^\circ$.*

2.3. Closed Sets

Definition 2.3.1. A subset F to X is **closed** if its complement F^c is open.

Example 2.3.2. An interval $[a, b] \subset \mathbf{R}$ is closed since $[a, b]^c = (-\infty, a) \cup (b, \infty)$ and both $(-\infty, a)$ and (b, ∞) are open according to Example 2.1.6. In the same way, one can show that all intervals $(-\infty, b]$ and $[a, \infty)$ are closed. □

Example 2.3.3. We will show that every ball $\overline{B}_r(x) \subset X$ is closed. Suppose that y belongs to $\overline{B}_r(x)^c = \{y \in X : d(x, y) > r\}$. If $s = d(x, y) - r$ and $d(y, z) < s$, it then follows from Proposition 1.6.1 that $z \in \overline{B}_r(x)^c$ since

$$d(x, z) \geq |d(x, y) - d(y, z)| = d(x, y) - d(y, z) > d(x, y) - s = r.$$

Thus, $B_s(y)$ is a subset to $\overline{B}_r(x)^c$. Since this holds for every $y \in \overline{B}_r(x)^c$, $\overline{B}_r(x)^c$ is open, and hence is $\overline{B}_r(x)$ closed. □

Example 2.3.4. If X is a discrete metric space, then by Example 2.1.7 and the definition, every subset of X is both open and closed. □

The next theorem is a counterpart to Theorem 2.1.8 for closed sets.

Theorem 2.3.5.

(a) *The sets* $X, \varnothing$ *are closed.*

(b) *If* F_α *is closed for every* $\alpha \in A$*, then* $\bigcap_{\alpha \in A} F_\alpha$ *is closed.*

(c) *If* $F_1 \dots, F_n$ *are closed, then* $\bigcup_1^n F_{j=1}$ *is closed.*

Proof. Property (a) follows from (i) in Theorem 2.1.8. We will prove that (b) holds and leave (c) as an exercise. By De Morgan's law,

$$\left(\bigcap_{\alpha \in A} F_\alpha \right)^c = \bigcup_{\alpha \in A} F_\alpha^c.$$

Now, F_α^c is open, so it follows from (ii) in Theorem 2.1.8 that $\cup_{\alpha \in A} F_\alpha^c$ is open. ■

Example 2.3.6. Exactly as in Example 2.3.3, one can show that

$$\{x\}^c = \{y \in X : d(x, y) > 0\}$$

is open for every $x \in X$, i.e., the singleton $\{x\}$ is closed. From (c) in Theorem 2.3.5 it now follows that every finite subset of X is closed. □

2.4. The Relative Topology

Proposition 2.4.1. *Let* Y *be a subspace to* X*.*

(a) *A subset* G *to* Y *is open in* Y *if and only if* $G = G_1 \cap Y$*, where* G_1 *is open in* X*.*

(b) *A subset* F *to* Y *is closed in* Y *if and only if* $F = F_1 \cap Y$*, where* F_1 *is closed in* X*.*

Proof.

(a) By Theorem 2.1.10, G is open in Y if and only if G is a union of open balls B in Y, i.e., balls of the form $B_1 \cap Y$, where B_1 is an open ball in X. Again by Theorem 2.1.10, this is equivalent to $G = G_1 \cap Y$, where G_1 is open in X.

(b) By the definition, F is closed in Y if and only $Y \setminus F$ is open in Y. But by (a), this holds if and only if $Y \setminus F = G_1 \cap Y$, where G_1 is open in X, which is the same as $F = F_1 \cap Y$, where $F_1 = G_1^c$ is closed in X. ■

Definition 2.4.2. The **relative topology** τ_Y on a subspace Y to X is the topology induced by $d|_Y$.

Proposition 2.4.1 shows that τ_Y consists of all sets of the form $G_1 \cap X$, where G_1 is open in X. The next example illustrates the fact that a set can be open (closed) in the relative topology without being open (closed) in the topology of X.

Example 2.4.3. Let $X = \mathbf{R}$ and $Y = (0, 1]$. Then the set $G = (\frac{1}{2}, 1]$ is open in Y since we have $G = (\frac{1}{2}, 2) \cap Y$, but G is not open in $\mathbf{R}$. Similarly, $F = (0, \frac{1}{2}]$ is closed in Y since $F = [0, \frac{1}{2}] \cap Y$, but F is not closed in $\mathbf{R}$. □

2.5. Closure

Definition 2.5.1. The **closure** $\overline{E}$ of a subset E to X is the intersection of all closed subsets of X that include E.

Example 2.5.2. For an interval $(a, b) \subset \mathbf{R}$, $-\infty < a < b < \infty$, $\overline{(a, b)} = [a, b]$ since $[a, b]$ is closed and all closed subsets of $\mathbf{R}$, that include (a, b), will include $[a, b]$. □

Example 2.5.3. It is not necessarily true that the closure of a ball $B_r(x)$ is the closed ball $\overline{B}_r(x)$. To see this, let X be a discrete metric space with moe than one element. Then $B_1(x) = \{x\}$ for every $x \in X$, so that $\overline{B_1(x)} = \{x\}$ since $\{x\}$ is closed, but $\overline{B}_1(x) = X$. □

Using (b) in Theorem 2.3.5, we obtain the following result:

Corollary 2.5.4. *Let $E \subset X$. Then*

(a) *$\overline{E}$ is the smallest closed subset of X that includes E;*

(b) *E is closed if and only if $E = \overline{E}$.*

2.6. Accumulation Points

Definition 2.6.1. An element $x \in X$ is an **accumulation point** of a subset E to X if every neighbourhood of x contains at least one element $y \in E$ with $y \neq x$. The set of accumulation points of E is denoted E' and called the **derived set**.

Example 2.6.2. For an interval $(a, b) \subset \mathbf{R}$, $-\infty < a < b < \infty$, $(a, b)' = [a, b]$. □

The next proposition shows that the closure of a set E is what one obtains when one to E adds the accumulation points of E.

Proposition 2.6.3. *If $E \subset X$, then $\overline{E} = E \cup E'$.*

Proof. We first prove that $\overline{E} \subset E \cup E'$. Let $x \in \overline{E} \setminus E$. If $x \notin E'$, there exists a neghbourhood G of x such that $G \cap E = \varnothing$. But then $E \subset G^c$, from which it follows that $\overline{E} \subset G^c$ since G^c is closed. This is a contradiction since $x \in \overline{E} \cap G$. Thus, $x \in E'$.

We then prove that $E \cup E' \subset \overline{E}$. Let $x \in E' \setminus E$. If $x \in \overline{E}^c$, then, since $\overline{E}^c$ is open and $\overline{E}^c \subset E^c$, there exists a neighbourhood G of x such that $G \subset E^c$. But this is impossible since $x \in E'$. Thus $x \in \overline{E}$. ■

Corollary 2.6.4. *A subset F to X is closed if and only if $F' \subset F$.*

Proof. If F is closed, then by (a) in Proposition 2.5.4, $F = \overline{F}$. But according to Proposition 2.6.3, $\overline{F} = F \cup F'$, so $F' \subset F$. Conversely, if we assume that $F' \subset F$, then $\overline{F} = F \cup F' = F$, so F is closed by the same corollary. ■

2.7. Dense Subsets

Definition 2.7.1. Let Y be a subset to X. Then a subset E to Y is **dense** in Y if $\overline{E} \supset Y$.

In particular, $E \subset X$ is dense in X if $\overline{E} = X$. Since $\overline{E} = E \cup E'$, this equivalent to the fact that every element in X, that does not belong to E, is an accumulation point of E.

Example 2.7.2. It is well-known that $\mathbf{Q}$ is dense in $\mathbf{R}$: every neighbourhood of a real number contains a rational number not equal to the given number. This fact implies that the set of **rational complex numbers**, i.e., complex numbers with rational real and imaginary parts, is dense in $\mathbf{C}$. From this it follows that the set of vectors with rational entries is dense in $\mathbf{R}^d$ and that the set of vectors with rational complex entries is dense in $\mathbf{C}^d$. □

Example 2.7.3. For every subset E to X, E is obviously dense in $\overline{E}$. □

2.8. Separability

Definition 2.8.1. A subset Y to X is **separable** if it contains a countable, dense subset.

Example 2.8.2. By Example 2.7.2, $\mathbf{R}^d$ and $\mathbf{C}^d$ are separable. □

Proposition 2.8.3. *For $1 \leq p < \infty$, ℓ^p is separable.*

Proof. We claim that the set E of sequences $(y^j)_{j=1}^{\infty}$, where every y^j is a rational complex number and all but a finite number of y^j are zero, is dense in ℓ^p. It suffices to show that if $x \in \ell^p$, then x is an accumulation point of E, i.e., every neighbourhood of x contains an element from E. For an arbitrary $\varepsilon > 0$, choose N so large that $\sum_{j=N+1}^{\infty} |x^j|^p < \varepsilon^p$. Next, for $j = 1, \ldots, N$, choose $y^j \neq x^j$ such that $|x^j - y^j| < \varepsilon/N^{1/p}$, and set $y^j = 0$ for $j = N+1, N+2, \ldots$. If $y = (y^j)$, then

$$d_p(x,y)^p = \sum_{j=1}^{N} |x^j - y^j|^p + \sum_{j=N+1}^{\infty} |x^j|^p = 2\varepsilon^p,$$

and hence, $d_p(x,y) < 2\varepsilon$. This shows that x is an accumulation point of E. ■

Example 2.8.4. We will show that ℓ^∞ is not separable. Let Y be the subspace to ℓ^∞ that consists of all sequences, where all elements in the sequence are 0 or 1. Then Y is uncountable since every element in Y corresponds to a number in $[0,1]$ written in the basis 2. The distance between two non-equal elements of Y is 1, which implies that the balls $B_{1/2}(x)$, where $x \in Y$, are disjoint. Now, suppose that $\{x_1, x_2, \ldots\}$ were a countable, dense subset to ℓ^∞. Then every ball $B_{1/2}(x)$ would contain at least one x_n. But this is impossible since the set of such balls is uncountable. □

We leave it as an exercise to show that the subspaces $\mathbf{c}$ and $\mathbf{c}_0$ to ℓ^∞ are in fact separable.

Example 2.8.5. The argument that was employed in Example 2.8.4 shows that a discrete metric space X is separable if and only if X is countable. □

Example 2.8.6. It follows from the Weierstrass approximation theorem (Corollary 5.7.7) that the space $C(K)$, where K is a compact subset to $\mathbf{R}^d$ is separable: The set of polynomials with rational coefficients is dense in $C(K)$. □

Example 2.8.7. It shown in integration theory that the space $L^p(E)$ is separable for $1 \leq p < \infty$, but not for $p = \infty$. □

Exercises

E2.1. Show that all intervals $(-\infty, a)$ and (a, ∞), where $-\infty < a < \infty$, are open.

E2.2. Prove (iii) in Theorem 2.1.8.

E2.3. Prove (c) in Theorem 2.3.5.

E2.4. Show that $\mathbf{c}$ and $\mathbf{c}_0$ are separable.

Chapter 3
Convergence in Metric Spaces

In this chapter, X will denote a metric space with metric d.

3.1. The Definition

Definition 3.1.1. A sequence $(x_n) \subset X$ is **convergent** with the **limit** $x \in X$ if $d(x, x_n) \to 0$. We denote this circumstance by writing $x_n \to x$ or $\lim_{n\to\infty} x_n = x$.

Proposition 3.1.2. *The limit of a convergent sequence is unique.*

Proof. Suppose that $x_n \to x$ and $x_n \to x'$. Then

$$d(x, x') \le d(x, x_n) + d(x_n, x') \to 0,$$

so $d(x, x') = 0$ and hence $x = x'$. ■

Example 3.1.3. In $\mathbf{R}^d$ and $\mathbf{C}^d$ with the standard metric d_2, a sequence converges if and only if every coordinate sequence converges:

$$x_n = (x_n^1, \dots, x_n^d)^t \to x = (x^1, \dots, x^d)^t \quad \text{if and only if} \quad x_n^j \to x^j,\ j = 1, \dots, d.$$

It is easy to see that the same holds with the metrics d_p, where $1 \le p \le \infty$. □

In the next example, we will show that in ℓ^∞ — unlike in $\mathbf{R}^d$ and $\mathbf{C}^d$ — coordinate-wise convergence does not imply convergence. The converse, however, is of course true: If a sequence converges in ℓ^∞, then every sequence of coordinates converges.

Example 3.1.4. Let $x_n = (x_n^j)_{j=1}^\infty \in \ell^\infty$, $n = 1, 2, \dots$, where $x_n^j = 1$ for $1 \le j \le n$ and $x_n^j = 0$ for $j > n$, and $x = (x^j)_{j=1}^\infty \in \ell^\infty$, where $x^j = 1$ for every j. Then, for every j, $x_n^j \to x^j$, but $x_n \nrightarrow x$ since $d(x, x_n) = \sup_{1 \le j < \infty} |x^j - x_n^j| = 1$. □

Example 3.1.5. In $\ell^\infty(M)$, $f_n \to f$ if $d(f, f_n) = \sup_{x\in M} |f(x) - f_n(x)| \to 0$. This type of convergence is known as **uniform convergence**. □

The next proposition shows that the accumulation points of a set E is the elements in X that can be approximated arbitrarily well by elements in E.

Proposition 3.1.6. *Let E be a subset to X. Then $x \in E'$ if and only if there exists a sequence $(x_n) \subset E$ such that $x_n \ne x$ for every n and $x_n \to x$.*

Proof. If $x \in E'$, then, for $n = 1, 2, \dots$, there exists an element $x_n \in E$ such that $x_n \ne x$ and $x_n \in B_{1/n}(x)$. But then $d(x, x_n) < \frac{1}{n}$, so $x_n \to x$. Conversely, suppose that $x_n \ne x$ for every n and $x_n \to x$. If G is an arbitrary neigbourhood of x, there exists a ball $B_r(x) \subset G$. Since $x_n \to x$, $x_n \in B_r(x)$ if n is large enough. This proves that $x \in E'$. ■

Corollary 3.1.7. *Let F be a subset to X. Then F is closed if and only if the limit of any convergent sequence of elements in F belongs to F.*

Proof. Suppose that F is closed and that $(x_n) \subset F$ such that $x_n \to x \in X$. If $x_n = x$ for some n, then $x \in F$. Otherwise, it follows from Theorem 3.1.6 that $x \in F'$. But according to Corollary 2.6.4, $F' \subset F$, so $x \in F$. To prove the converse, we let $x \in F'$ and choose a sequence $(x_n) \subset F$ such that $x_n \to x$. Then, by the assumption, $x \in F$. Thus, $F' \subset F$, so F is closed. ■

3.2. Cauchy Sequences

Definition 3.2.1. A sequence $(x_n) \subset X$ is a **Cauchy sequence** if

$$d(x_m, x_n) \to 0 \quad \text{as } m, n \to \infty.$$

Proposition 3.2.2. *Every convergent sequence in X is a Cauchy sequence.*

Proof. Suppose that $x_n \to x$. Then

$$d(x_m, x_n) \leq d(x_m, x) + d(x, x_n) \to 0 \quad \text{as } m, n \to \infty. \qquad \blacksquare$$

The converse to this proposition is false:

Example 3.2.3. The sequence $(\frac{1}{n})_{n=1}^{\infty} \subset (0, 1]$ is a Cauchy sequence since

$$\left|\frac{1}{m} - \frac{1}{n}\right| \leq \frac{1}{m} + \frac{1}{n} \to 0 \quad \text{as } m, n \to \infty.$$

However, it is not convergent in $(0, 1]$. □

3.3. Completeness

Definition 3.3.1. A subset Y to X is **complete** if every Cauchy sequence in Y is convergent.

Example 3.3.2. Accoording to Example 3.2.3, $(0, 1] \subset \mathbf{R}$ is not complete. □

Example 3.3.3. The space $\mathbf{Q}$ is not complete. Take for instance $x_n = (1 + \frac{1}{n})^n$ for $n = 1, 2, \ldots$. This is a Cauchy sequence in $\mathbf{Q}$ since it converges to e in $\mathbf{R}$. But since $e \notin \mathbf{Q}$, the sequence is not convergent in $\mathbf{Q}$. □

The next result, which we state without a proof, shows that $\mathbf{R}$ is complete.

Theorem 3.3.4 (The Cauchy Criterion). *Every Cauchy sequence in $\mathbf{R}$ is convergent.*

Corollary 3.3.5. *The spaces $\mathbf{R}^d$ and $\mathbf{C}^d$ are complete.*

Proof. As in Example 3.1.3, we see that if (x_n) is a Cauchy sequence in $\mathbf{R}^d$, then every coordinate sequence is a Cauchy sequence and thus convergent. Again using Example 3.1.3, this implies that (x_n) is convergent. The completeness of $\mathbf{C}^d$ follows by looking at the real and imaginary parts of a sequence. ■

Proposition 3.3.6. *For $1 \leq p < \infty$, ℓ^p is complete.*

Proof. Let (x_n) be a Cauchy sequence in ℓ^p, where $x_n = (x_n^j)_{j=1}^{\infty} \in \ell^p$. Given an arbitrary $\varepsilon > 0$, we choose N so large that

$$d_p(x_m, x_n)^p = \sum_{j=1}^{\infty} |x_n^j - x_m^j|^p < \varepsilon^p \quad \text{if } m, n \geq N. \tag{1}$$

This of course implies that, for every fixed $j \geq 1$, $|x_n^j - x_m^j| < \varepsilon$ if $m, n \geq N$, so that $(x_n^j)_{j=1}^\infty$ is a Cauchy sequence for every j. The completeness of $\mathbf{C}$ now implies that there exist numbers x^j such that $x_n^j \to x^j$ as $n \to \infty$. Let $x = (x^j)_{j=1}^\infty$. It follows from (1) that, for every $k \geq 1$,

$$\sum_{j=1}^{k} |x_n^j - x_m^j|^p < \varepsilon^p \quad \text{if } m, n \geq N.$$

If we first let $m \to \infty$ and then $k \to \infty$ in this inequality, we obtain

$$\sum_{j=1}^{\infty} |x_n^j - x^j|^p \leq \varepsilon^p \quad \text{if } n \geq N. \tag{2}$$

This shows that $x_n - x \in \ell^p$. But since we know that ℓ^p is a vector space, it follows that $x = x_n - (x_n - x) \in \ell^p$. The inequality (2) also shows that $d(x, x_n) \leq \varepsilon$ if $n \geq N$, and thus $x_n \to x$. ■

Example 3.3.7. We know from theory of integration that $L^p(E)$ is complete for $1 \leq p \leq \infty$. □

Proposition 3.3.8. *For an arbitrary set M, $\ell^\infty(M)$ is complete.*

It follows, for instance, that ℓ^∞ is complete.

Proof. Let $(f_n) \subset \ell^\infty(M)$ be a Cauchy sequence. Then

$$|f_m(x) - f_n(x)| \leq d(f_m, f_n) \to 0 \quad \text{for every } x \in M,$$

which shows that $(f_n(x))_{n=1}^\infty \subset \mathbf{C}$ is a Cauchy sequence for every $x \in M$. Since $\mathbf{C}$ is complete, this implies that $f(x) = \lim_{n\to\infty} f_n(x)$ exists for every $x \in M$. We next show that $f_n \to f$ in $\ell^\infty(M)$. Let $\varepsilon > 0$ be arbitrary and choose N so large that $d(f_m, f_n) < \varepsilon$ if $m, n \geq N$. Given $x \in M$, we then choose $m \geq N$ such that $|f(x) - f_m(x)| < \varepsilon$. It now follows that

$$|f(x) - f_n(x)| \leq |f(x) - f_m(x)| + |f_m(x) - f_n(x)| < \varepsilon + d(f_m, f_n) < 2\varepsilon$$

if $n \geq N$. This holds for every $x \in M$, so $d(f, f_n) \leq 2\varepsilon$ if $n \geq N$, i.e., $f_n \to f$. From the last inequality, we obtain that

$$|f(x)| \leq |f(x) - f_N(x)| + |f_N(x)| < 2\varepsilon + \sup_{y\in M} |f_N(y)|$$

for every $x \in M$, which shows that f is bounded on M. ■

The next result is often useful for proving that a metric space is complete.

Proposition 3.3.9.

(a) *If F is a complete subspace to X, then F is closed.*

(b) *If X is complete, then evey closed subspace F to X is complete.*

Thus, if X is complete, a subspace to X is complete if and only if it is closed.

Proof.

(a) Let $(x_n) \subset F$ be a sequence that converges to $x \in X$. Using Proposition 3.2.2 and the completeness of F, it follows directly that $x \in F$. According to Corollary 3.1.7, this implies that F is closed.

(b) Let $(x_n) \subset F$ be a Cauchy sequence. Since X is complete, there exists an element $x \in X$ such that $x_n \to x$. But F is closed, so $x \in F$. This shows that F is complete. ■

3.4. Completion

Example 3.4.1. It is known that the space $R^1[a, b]$ is not complete. □

The next theorem, which we will not use and hence state without a proof, shows that every incomplete may be embedded in a complete metric space.

Theorem 3.4.2. *There exists a metric space* $\widehat{X}$ *with metric* $\widehat{d}$ *such that*

(i) $\widehat{X}$ *is complete,*

(ii) *there exists an isometry* $\sigma : X \to \widehat{X}$ *such that* $\widehat{d}(\sigma(x), \sigma(y)) = d(x, y)$ *for all* $x, y \in X$*,*

(iii) $\sigma(X)$ *is dense in* $\widehat{X}$*.*

The space $\widehat{X}$ *is called the* **completion** *of* X*.*

Example 3.4.3. We give two examples of completions:

(a) $\mathbf{R}$ is the completion of $\mathbf{Q}$;

(b) $L^1(a, b)$ is the completion of $R^1[a, b]$. □

Exercises

E3.1. Show that if a Cauchy sequence in a metric space has a convergent subsequence, then the sequence is convergent.

Chapter 4

Continuity in Metric Spaces

Let X and Y denote metric spaces with metrics d_X and d_Y, respectively.

4.1. The Definition

The following definition mimics the one usually encountered in analysis courses.

Definition 4.1.1. A function f from X to Y is **continuous** at $x \in X$ if there for every $\varepsilon > 0$ exists a $\delta > 0$ such that

$$d_X(x, x') < \delta \quad \text{implies that} \quad d_Y(f(x), f(x')) < \varepsilon.$$

The function f is **continuous** if it is continuous at every point in X.

4.2. A Topological Characterisation of Continuity

Proposition 4.2.1. *A function $f : X \to Y$ is continuous if and only if $f^{-1}(V)$ is open in X for every open subset V to Y.*

Here, we use the notation $f^{-1}(V) = \{x \in X : f(x) \in V\}$, where V is a subset to Y.

Proof. Suppose that f is continuous and let V be an open subset to Y. We can obviously assume that $f^{-1}(V)$ is non-empty, so there exists a point $x \in f^{-1}(V)$. Since V is open, there exists an $\varepsilon > 0$ such that $B_\varepsilon(f(x)) \subset V$. If $\delta > 0$ is as in the definition of continuity, it follows that $f(x') \in B_\varepsilon(f(x))$ if $x' \in B_\delta(x)$, which means that $B_\delta(x) \subset f^{-1}(V)$, so $f^{-1}(V)$ is open.

Conversely, let $x \in X$ and $\varepsilon > 0$. If $V = B_\varepsilon(f(x))$, then since $f^{-1}(V)$ is open by assumption, there exists a number $\delta > 0$ such that $B_\delta(x) \subset f^{-1}(V)$, i.e., $d_X(x, x') < \delta$ implies that $d_Y(f(x), f(x')) < \varepsilon$, so f is continuous at x. ■

Example 4.2.2. Let X be a discrete metric space. According to Example 2.1.7, all subsets to X are open, so every function $f : X \to Y$ is trivially continuous. □

4.3. A Sequential Characterisation of Continuity at a Point

Proposition 4.3.1. *A function $f : X \to Y$ is continuous at $x \in X$ if and only if $f(x_n) \to f(x)$ for every sequence $(x_n) \subset X$ such that $x_n \to x$.*

Proof. Suppose that f is continuous at x and that $x_n \to x$. Given $\varepsilon > 0$, let $\delta > 0$ be as in the definition of continuity. Then $d_X(x, x_n) < \delta$ if n is sufficiently large, so it follows that $d_Y(f(x), f(x_n)) < \varepsilon$ for those n. Thus, $f(x_n) \to f(x)$.

To prove the converse, we suppose that f is not continuous at x. Then there exists an $\varepsilon > 0$ and a sequence $(x_n) \subset X$ such that $d_X(x, x_n) < \frac{1}{n}$ for every n, but $d_Y(f(x), f(x_n)) \geq \varepsilon$. This, of course, contradicts the assumption. ■

Proposition 4.3.1 in conjunction with the ordinary rules for limits of numerical sequences gives us the following result.

Proposition 4.3.2. *Suppose that $f, g : X \to \mathbf{C}$ are continuous at $x \in X$. Then the functions $\alpha f + \beta g$, where $\alpha, \beta \in \mathbf{C}$, and fg are continuous at x. If, in addition, we assume that $g(x) \neq 0$, then the function f/g is continuous at x.*

It follows that if f and g are continuous, then $\alpha f + \beta g$ and fg also are continuous, and that f/g is continuous if $g(x) \neq 0$ for every $x \in X$.

4.4. Spaces of Continuous Functions

Definition 4.4.1. We denote by $C(X)$ the set of continuous, complex-valued functions on X. The subset of bounded functions in $C(X)$ is denoted $C_b(X)$.

We denote by $C(X, \mathbf{R})$ and $C_b(X, \mathbf{R})$ the real-valued functions and bounded, real-valued functions in $C(X)$, respectively. It is clear that $C_b(X)$ and $C_b(X, \mathbf{R})$ are metric spaces with the metric

$$d_\infty(f, g) = \sup_{x \in X} |f(x) - g(x)|.$$

According to Proposition 4.3.2, all these spaces are vector spaces.

Proposition 4.4.2. *The space $C_b(X)$ is a closed subspace of $\ell^\infty(X)$.*

Proof. Suppose that $(f_n) \subset C_b(X)$ and that $f_n \to f \in \ell^\infty(X)$; we will show that f is continuous at every $x \in X$. Let $\varepsilon > 0$ be given and choose n so large that $d_\infty(f, f_n) < \varepsilon$. Next, using the fact that f_n is continuous at x, choose $\delta > 0$ such that $|f_n(x) - f_n(x')| < \varepsilon$ for every $x' \in X$ such that $d_X(x, x') < \delta$. It then follows that

$$\begin{aligned} |f(x) - f(x')| &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(x')| + |f_n(x') - f(x')| \\ &\leq 2d_\infty(f, f_n) + \varepsilon < 3\varepsilon \end{aligned}$$

if $d_X(x, x') < \delta$, which is exactly what we wanted to prove. ■

Corollary 4.4.3. *The space $C_b(X)$ is complete.*

Proof. Since $C_b(X)$ is a closed subspace of $\ell^\infty(X)$ and $\ell^\infty(X)$ is complete, it follows from Proposition 3.3.9 that $C_b(X)$ is complete. ■

Since $C_b(X, \mathbf{R})$ obviously is a closed subspace to $C_b(X)$, it follows that $C_b(X, \mathbf{R})$ is complete.

Chapter 5

Compactness in Metric Spaces

In this chapter, X will denote a metric space with metric d.

5.1. Definition and Examples

Definition 5.1.1. An **open covering** of X is a collection $\{G_\alpha\}_{\alpha\in A}$ of open subsets to X such that $X \subset \bigcup_{\alpha\in A} G_\alpha$. A **finite subcovering** to such a covering is an open covering $\{G_\alpha\}_{\alpha\in A'}$ of X, where A' is a finite subset to A.

Definition 5.1.2. The space X is **compact** if every open covering of X has a finite subcovering. A subspace K to X is compact if K is compact with respect to the relative topology.

It follows from Proposition 2.4.1 that an open covering of a subspace K to X has the form $\{G_\alpha \cap K\}_{\alpha\in A}$, where every G_α is an open subset to X.

Example 5.1.3.

(a) The interval $(0,1) \subset \mathbf{R}$ is not compact since the covering $(\frac{1}{n}, 1)$, $n = 1, 2, \dots$, of $(0,1)$ obviously has no finite subcovering.

(b) The space $\mathbf{R}$ is not compact since the covering $(-n, n)$, $n = 1, 2, \dots$, of $\mathbf{R}$ lacks a finite subcovering.

(c) Every finite subset to $\mathbf{R}^d$ is obviously compact. □

Proposition 5.1.4. *If X is compact and $F \subset X$ is closed, then F is compact.*

Proof. Suppose that $F \subset \bigcup_{\alpha\in A} G_\alpha \cap F$, where every G_α is open in X. It then follows that $X \subset \bigcup_{\alpha\in A} G_\alpha \cup F^c$. But since X is compact, there exist a finite number of indices $\alpha_1, \dots, \alpha_n$ such that $X \subset \bigcup_{j=1}^n G_{\alpha_j} \cup F^c$. But this implies that $F \subset \bigcup_{j=1}^n G_{\alpha_j} \cap F$, i.e., F is compact. ■

Definition 5.1.5. A subset E to X is bounded if $\sup_{x,y\in E} d(x,y) < \infty$.

Proposition 5.1.6. *Every compact subset K to X is closed and bounded.*

According to the Heine–Borel theorem below (Theorem 5.2.5), the converse to this proposition is true in $\mathbf{R}^d$ and $\mathbf{C}^d$. In general, however, the converse is false; see Example 5.2.6.

Proof. We will first show that K is closed. Let $y \in K^c$. For every $x \in K$, there exist two disjoint, open balls B_x och B_x^y such that $x \in B_x$ and $y \in B_x^y$. Since $K \subset \bigcup_{x\in K} B_x \cap K$ and K is compact, there are elements $x_1, \dots, x_n \in K$ such that $K \subset \bigcup_{j=1}^n B_{x_j} \cap K$. The set $G = \bigcap_{j=1}^n B_{x_j}^y$ is then a neighbourhood of y such that $G \cap K = \varnothing$. This shows that y is not an accumulation point of K. Since this is true for every $y \in K^c$, $K' \subset K$, so K is closed.

To show that K is bounded, we first extract a finite subcovering $\{B_1(x_j)\cap K\}_{j=1}^n$ from the covering $\{B_1(x)\cap K\}_{x\in K}$ of K. Suppose that $x, y \in K$ and that $x \in B_1(x_r)$ and $y \in B_1(x_s)$. Then

$$d(x,y) \le d(x,x_r) + d(x_r,x_s) + d(x_s,y) < 2 + \max_{1\le j,k\le N} d(x_j,x_k),$$

which shows that K is bounded. ■

5.2. Totally Bounded Sets, Sequential Compactness

Definition 5.2.1. The space X is **totally bounded** if there for every $\varepsilon > 0$ exist elements $x_1, \dots, x_n \in X$ such that $X \subset \bigcup_{j=1}^n B_\varepsilon(x_j)$.

As in the proof of proposition 5.1.6, one can show that a compact space is totally bounded and that a totally bounded space is bounded. A bounded space, however, does not have to be totally bounded; see Example 5.2.6.

Definition 5.2.2. The space X is called **sequentially compact** is every sequence in X has a convergent subsequence.

The following theorem summarises the relations between the concepts introduced so far. Notice that a consequence of the theorem is that every compact space is complete.

Theorem 5.2.3. *The following conditions are equivalent*:

(i) X *is compact*;

(ii) X *sequentially compact*;

(iii) X *complete and totally bounded.*

Proof. We will first show that (i) implies (ii). Let $(x_n)_{n=1}^\infty$ be a sequence of elements from X. The set $F_n = \overline{\{x_n, x_{n+1}, \dots\}}$, $n = 1, 2, \dots$, is by definition closed, so its complement $G_n = X \setminus F_n$ is open. Suppose that $\bigcap_{n=1}^\infty F_n = \varnothing$. It then follows that

$$\bigcup_{n=1}^\infty G_n = \bigcup_{n=1}^\infty (X \setminus F_n) = X \setminus \bigcap_{n=1}^\infty F_n = X,$$

so the sets G_n form an open covering of X. Since X is assumed to be compact, X is covered by a finite number of the sets G_n: $X \subset \bigcup_{n=1}^N G_n$. But this implies that $\bigcap_1^N F_{n=1} = \varnothing$, which is impossible since $x_j \in \bigcap_{n=1}^N F_n$ for $j \ge N$. The assumption was thus incorrect, so we have $\bigcap_{n=1}^\infty F_n \ne \varnothing$. A moment's reflection shows that this implies that $(x_n)_1^\infty$ has a convergent subsequence.

We next show that (ii) implies (iii). If $(x_n) \subset X$ is a Cauchy sequence, then by the assumption, the sequence has a convergent subsequence, which implies that the whole sequence is convergent (see Exercise E3.1). Thus, X is complete. Now suppose that X is not totally bounded. Then there exists a number $\varepsilon > 0$ such that $X \not\subset \bigcup_{n=1}^N B_\varepsilon(x_n)$ no matter how N and $x_1, \dots, x_N$ are chosen. First choose $x_1 \in X$ arbitrary, then $x_2 \in B_\varepsilon(x_1)^c$, then $x_3 \in (B_\varepsilon(x_1) \cup B_\varepsilon(x_2))^c$ and so on. But since $d(x_m, x_n) \ge \varepsilon$ for $m, n = 1, 2, \dots$, the sequence (x_n) cannot have a convergent subsequence (it is not even a Cauchy sequence). This contradiction shows that X is totally bounded.

Let us finally show that (iii) implies (i). Suppose that $\{G_\alpha\}_{\alpha\in A}$ is a covering of X with no finite subcovering. Since X is assumed to be totally bounded, there are elements $x_1^{(1)}, \dots, x_{N_1}^{(1)} \in X$ such that $X \subset \bigcup_{j=1}^{N_1} B_1(x_j^{(1)})$. At least one of the balls in the covering — say B_1 — cannot be covered with a finite number of sets G_α (otherwise X could also be covered with a finite number of sets G_α). Now suppose that $X \subset \bigcup_{j=1}^{N_2} B_{2^{-1}}(x_j^{(2)})$. Of the balls $B_{2^{-1}}(x_j^{(2)})$, that intersect B_1, at least one — say B_2 – has no finite subcovering. Continuing in this way, we obtain a sequence of open balls B_n with centers x_n and radii 2^{-n}, such that none of the balls can be covered with a finite number of sets G_α. Suppose that $y \in B_n \cap B_{n+1}$. Then

$$d(x_n, x_{n+1}) \le d(x_n, y) + d(y, x_{n+1}) < 2^{-n} + 2^{-(n+1)} < 2^{-(n-1)},$$

which implies that for $m > n \ge N$,

$$\begin{aligned} d(x_n, x_m) &\le d(x_n, , x_{n+1}) + \dots + d(x_{m-1}, x_m) < 2^{-(n-1)} + \dots + 2^{-(m-2)} \\ &< 2^{-(N-2)}. \end{aligned}$$

Thus, (x_n) is a Cauchy sequence and therefore the sequence converges to some element $x \in X$. Suppose that $x \in G_\alpha$ and choose r so small that $B_r(x) \subset G_\alpha$ and n so large that $d(x, x_n) < r/2$ and $2^{-n} < r/2$. If $y \in B_n$, then $d(x, y) < r$, so $B_n \subset B_r(x) \subset G_\alpha$, which is a contradiction. We conclude that X is compact. ■

Corollary 5.2.4. *If X is complete, then $K \subset X$ is compact if and only if K is closed and totally bounded.*

Using this corollary, we get a short proof of the classical Heine–Borel theorem.

Theorem 5.2.5 (Heine–Borel). *If K is closed and bounded set in $\mathbf{R}^d$ or $\mathbf{C}^d$, then K is compact.*

Proof. We prove the the theorem for subsets to $\mathbf{R}^d$ and leave the rest of the proof as an exercise. Since K is bounded, $K \subset [-a, a]^d$ if a is sufficiently large. By Corollary 5.2.4, the set $[-a, a]^d$ is compact since it is closed and totally bounded. From Proposition 5.1.4, it now follows that K is compact. ■

Example 5.2.6. The set $\overline{B_1(0)} \subset \ell^2$ is obviously both closed and bounded; it is, however, not compact. In fact, let $(\delta_n)_{n=1}^\infty \subset \overline{B_1(0)}$ be defined by $\delta_n^j = 1$ if $j = n$ and $\delta_n^j = 0$ otherwise. Then $d_2(\delta_m, \delta_n) = \sqrt{2}$ if $m \ne n$, which shows that the sequence does not contain a convergent subsequence. Hence, $\overline{B_1(0)}$ is not sequentially compact and not compact. This example also shows that a bounded set does not have to be totally bounded. □

5.3. Relatively Compact Sets

Definition 5.3.1. A subspace Y to X is **relatively compact** if $\overline{Y}$ is compact.

Example 5.3.2. Every bounded subset E to $\mathbf{R}^d$ is relatively compact. Indeed, $\overline{E}$ is closed and bounded, and thus compact by the Heine–Borel theorem. □

The following necessary and sufficient condition for a subspace to be relatively compact follows from Theorem 5.2.3.

Proposition 5.3.3. *A subspace Y to X is relatively compact if and only if every sequence in Y has a subsequence that converges to an element of $\overline{Y}$.*

Proof. If Y is relatively compact, then every sequence in $\overline{Y}$ and, in particular, every sequence in Y has a subsequence with limit in $\overline{Y}$. For the converse, let (x_n) be a sequence in $\overline{Y}$. Then, according to Proposition 3.1.6, there exist a sequence of elements $y_n \in Y$ such that $d(x_n, y_n) < \frac{1}{n}$ for every n. The assumption now shows that some subsequence (y_{n_k}) converges to $y \in Y$. But then (by the triangle inequality), x_{n_k} also converges to y. ■

Corollary 5.3.4 (The Bolzano–Weierstrass Theorem). *In $\mathbf{R}^d$ and $\mathbf{C}^d$, every bounded sequence has a convergent subsequence.*

Proof. We prove the the corollary for real sequences and leave the rest of the proof as an exercise. Suppose that (x_n) is a bounded sequence in $\mathbf{R}^d$. Then the set $Y = \{x_n : n = 1, 2, ...\}$ is bounded and thus relatively compact as a subspace to $\mathbf{R}^d$. ■

We leave the proof of the following proposition as an exercise.

Proposition 5.3.5. *Every relatively compact subspace to X is totally bounded, and if X is complete, every totally bounded subspace to X is relatively compact.*

5.4. Continuous Functions on Compact Sets

We next prove the classical results about continuous functions on compact sets in a general setting. Below, Y will denote a metric space with metric d_Y.

Theorem 5.4.1. *If X is compact and $f : X \to Y$ is a continuous function, then the range $f(X)$ of f is also compact.*

Proof. Suppose that $f(X) = \bigcup_{\alpha \in A} V_\alpha$, where every V_α is open in $f(X)$. Then, according to Proposition 4.2.1, the set $U_\alpha = f^{-1}(V_\alpha)$ is open for every α. Moreover, there holds $X = \bigcup_{\alpha \in A} U_\alpha$. Now, since X is assumed to be compact, there are indices $\alpha_1, \dots, \alpha_n \in A$ such that $X = \bigcup_{k=1}^{n} U_{\alpha_k}$. This implies that $f(X) = \bigcup_{k=1}^{n} V_{\alpha_k}$, so $f(X)$ is compact. ■

Theorem 5.4.2. *If X is compact and $f : X \to \mathbf{R}$ is continuous, then f attains a maximal and a minimal value on X.*

Proof. We will show that f has a maximal value on X; the existence of a minimal value follows by considering the function $-f$. By Theorem 5.4.1, $f(X)$ is compact, and hence bounded and closed by the Heine–Borel theorem. The fact that $f(X)$ is bounded implies that $\sup f(X) < \infty$, and the fact that $f(X)$ is closed that $\sup f(X) \in f(X)$. Thus, there exists a point $x \in X$ such that $f(x) = \sup f(X)$. ■

Definition 5.4.3. A function $f : X \to Y$ is **uniformly continuous** if there for every $\varepsilon > 0$ exists a $\delta > 0$ such that

$$d(x, x') < \delta \quad \text{implies that} \quad d_Y(f(x), f(x')) < \varepsilon.$$

Theorem 5.4.4. *If X is compact, then every continuous function $f : X \to Y$ is uniformly continuous.*

Proof. Let $\varepsilon > 0$ be given. Then, for every $x \in X$, there exists a number $\delta(x) > 0$ such that $d_Y(f(x), f(x')) < \varepsilon$ if $d(x, x') < \delta(x)$. The balls $B_{\delta(x)/2}(x)$, where x varies over X, of course cover X, so by compactness, we can find a finite number of points $x_1, \dots, x_n \in X$ such that $X = \bigcup_{j=1}^{n} B_j$, where $B_j = B_{\delta(x_j)/2}(x_j)$. Next put $\delta = \frac{1}{2}\min(x_1, \dots, x_n)$ and suppose that $d(x, x') < \delta$. If $x \in B_j$ for some j, then

$$d(x_j, x') \le d(x_j, x) + d(x, x') < \tfrac{1}{2}\delta(x_j) + \delta \le \delta(x_j),$$

so $x' \in B_j$. Using the fact that f is continuous at x_j, it now follows that

$$d_Y(f(x), f(x')) \le d_Y(f(x), f(x_j)) + d_Y(f(x_j), f(x')) < 2\varepsilon. \qquad \blacksquare$$

5.5. Compactness Criteria — The Arzelà–Ascoli theorem

Let X denote a compact metric space with metric d.

Definition 5.5.1. A subset E to $C(X)$ is said to be **equicontinuous** if for every $\varepsilon > 0$, there exists a number $\delta > 0$ such that, for every $f \in E$,

$$d(x, x') < \delta \quad \text{implies that} \quad |f(x) - f(x')| < \varepsilon.$$

Example 5.5.2. The subset

$$E = \{f \in C[a, b] : |f(s) - f(t)| \le C|s - t|^{\alpha},\ s, t \in [a, b]\}$$

to $C[a, b]$, where $\alpha > 0$ is fixed, is equicontinuous. □

Theorem 5.5.3 (The Arzelà–Ascoli theorem). *A subspace E to $C(X)$ is relatively compact if and only if E is bounded and equicontinuous.*

The sufficiency part of the theorem was proved by G. Ascoli in 1882-83 and the necessity part by C. Arzelà in 1889.

Proof. We can obviously assume that both X and E are non-empty. First, suppose that the subspace E is relatively compact. Then, according to Proposition 5.3.5, E is totally bounded, so for a given $\varepsilon > 0$, there are functions $f_1, \dots, f_n \in E$ such that $E \subset \bigcup_{j=1}^{n} B_\varepsilon(f_j)$. It also follows from Theorem 5.4.4 that each of these functions is uniformly continuous, that is, we can find a number $\delta > 0$ such that $|f_j(x) - f_j(x')| < \varepsilon$ for every j if $d(x, x') < \delta$. Now, suppose that $f \in E \cap B_\varepsilon(f_j)$. If $d(x, x') < \delta$, then

$$|f(x) - f(x')| \le |f(x) - f_j(x)| + |f_j(x) - f_j(x')| + |f_j(x') - f(x')| < 3\varepsilon,$$

thus proving that E is equicontinuous. The set E is finally bounded since it is totally bounded.

Suppose conversely that E is bounded and equicontinuous. Let $\varepsilon > 0$ be arbitrary and $\delta > 0$ as in the definition of equicontinuity. Since X is compact and hence totally bounded, there are points $x_1, \dots, x_m \in X$ such that $X = \bigcup_{j=1}^m B_\delta(x_j)$. From the boundedness of E, it follows that the set $\{(f(x_1), \dots, f(x_n))^t : f \in E\} \subset \mathbf{C}^n$ is bounded and therefore totally bounded, so there are functions $f_1, \dots, f_n \in E$ such that if $f \in E$, then

$$\left(\sum_{j=1}^n |f(x_j) - f_k(x_j)|^2 \right)^{1/2} < \varepsilon$$

for some k. If $x \in B_\delta(x_j)$, then

$$|f(x) - f_k(x)| \le |f(x) - f(x_j)| + |f(x_j) - f_k(x_j)| + |f_k(x_j) - f_k(x)| < 3\varepsilon.$$

Taking the supremum over all $x \in X$, it follows that $d_\infty(f, f_k) \le 3\varepsilon$. Thus, E is totally bounded and hence relatively compact by Corollary 4.4.3 and Proposition 5.3.5. ■

5.6. Peano's Existence Theorem

As an application of the Arzelà–Ascoli theorem, we prove the Peano existence theorem. Consider the following system of differential equations with initial conditions:

$$\begin{cases} y'(t) = f(t, y(t)), \ t \ge 0 \\ y(0) = y_0 \end{cases} . \tag{1}$$

Here we assume that the function f in the right-hand side is continuous on the compact set $A = \{(t, y) \in \mathbf{R} \times \mathbf{R}^d : 0 \le t \le T, \ |y - y_0| \le K\}$ with values in $\mathbf{R}^d$. Suppose also that $|f| \le M$ on this set.

Theorem 5.6.1 (G. Peano 1890). *The problem* (1) *has at least one solution* y *defined for* $0 \le t \le T_1$*, where* $T_1 = \min(T, K/M)$*.*

In the proof we will use the space $C[0, T_1; \mathbf{R}^d]$ of continuous functions on $[0, T_1]$ with values in $\mathbf{R}^d$. The metric in this space is $d_\infty(x, y) = \sup_{0 \le t \le T_1} |x(t) - y(t)|$. It is not so hard to show that $C[0, T_1; \mathbf{R}^d]$ is complete.

Proof. For $n = 1, 2, \dots$, put $y_n(t) = y_0$ for $0 \le t \le T_1/n$. Next, define $y_n(t)$ for $T_1/n \le t \le 2T_1/n$ by

$$y_n(t) = y_0 + \int_0^{t - T_1/n} f(\tau, y_n(\tau)) \, d\tau. \tag{2}$$

Notice that y_n is well-defined through this formula and that $|y_n(t) - y_0| \le MT_1/n$. Using the fact that $MT_1 \le K$, we now see that equation (2) also defines $y_n(t)$ for $2T_1/n \le t \le 3T_1/n$ and that $|y_n(t) - y_0| \le 2MT_1/n \le K$ for t belonging to this interval. We continue in this way and obtain that (2) defines y_n on $[0, T_1]$ and that $|y_n| \le |y_0| + K$. One can also show that

$$|y_n(t) - y_n(s)| \le M|t - s|$$

for $0 \le s \le t \le T_1$. If, for instance, $T_1/n \le s \le t \le T_1$, then

$$|y_n(t) - y_n(s)| = \left| \int_{s-T_1/n}^{t-T_1/n} f(\tau, y_n(\tau))\, d\tau \right| \le M|t-s|.$$

From this, we obtain that the subspace $E = \{y_n : n = 1, 2, ...\}$ to $C[0, T_1; \mathbf{R}^d]$ is equicontinuous. Hence, according to the Arzelà–Ascoli theorem, E is relatively compact. It now follows from Proposition 5.3.3 and Proposition 4.4.2 that there exists a subsequence to $(y_n)_{n=1}^\infty$, which we after renumbering still can denote $(y_n)_{n=1}^\infty$, that converges to a function $y \in C[0, T_1; \mathbf{R}^d]$. Letting $n \to \infty$ in (2) and using the fact that f is uniformly continuous, we obtain that

$$y(t) = y_0 + \int_0^t f(\tau, y(\tau))\, d\tau, \quad 0 \le t \le T_1,$$

which implies that y is continuously differentiable and satisfies (1). ■

Example 5.6.2. The solution to (1), whose existence is guaranteed by Theorem 5.6.1, does not necessarily have to be unique. For instance, the problem

$$\begin{cases} y'(t) = 2\sqrt{y(t)},\ t \ge 0 \\ y(0) = 0 \end{cases}$$

has both the solution $y(t) = 0,\ t \ge 0$, and the solution $y(t) = t^2,\ t \ge 0$. □

5.7. The Stone–Weierstrass Theorem

In this section, X denotes a *compact* metric space. Thus, all continuous real- or complex-valued functions are bounded.

Definition 5.7.1. A subspace $\mathcal{A}$ to $C(X)$ is a **subalgebra** to $C(X)$ if $\alpha f + \beta g \in \mathcal{A}$ and $fg \in \mathcal{A}$ whenever $f, g \in \mathcal{A}$ and $\alpha, \beta \in \mathbf{C}$.

Notice that every subalgebra is a vector subspace to $C(X)$ and that $C(X)$ is a subalgebra to itself. The notion of a subalgebra to $C(X, \mathbf{R})$ is defined similarly.

Lemma 5.7.2. *For every number $\varepsilon > 0$, there exists a real polynomial p such that*

$$\max_{-1 \le t \le 1} \big| |t| - p(t) \big| < \varepsilon.$$

Proof. Given $\varepsilon > 0$, put $f_\varepsilon(t) = (t + \varepsilon^2/4)^{1/2},\ 0 \le t \le 1$. It is not so hard to see that the Taylor series at $t = \frac{1}{2}$ for f_ε converges uniformly to f_ε on $[0, 1]$. Thus, there exists a polynomial q such that $|f_\varepsilon(t) - q(t)| < \frac{\varepsilon}{2}$ for $0 \le t \le 1$. Now, if $-1 \le t \le 1$, then

$$\big| |t| - q(t^2) \big| \le \big| |t| - f_\varepsilon(t^2) \big| + \big| f_\varepsilon(t^2) - q(t^2) \big| < \varepsilon.$$

Thus, the assertion follows if we take $p(t) = q(t^2)$. ■

Proposition 5.7.3. *Suppose that $\mathcal{A}$ is a closed subalgebra to $C(X)$ or $C(X, \mathbf{R})$. Then, $|f| \in \mathcal{A}$ if $f \in \mathcal{A}$. Moreover, $\mathcal{A}$ is a lattice, i.e., if $f, g \in \mathcal{A}$, then* $\max(f, g)$ *and* $\min(f, g)$ *both belong to $\mathcal{A}$.*

Proof. We first prove that if $f \in \mathcal{A}$, then $|f| \in \mathcal{A}$. First suppose that $|f| \leq 1$. Let $\varepsilon > 0$ and let p denote the polynomial in Lemma 5.7.2. Then $p(f) \in \mathcal{A}$ and $\big||f(x)| - p(f(x))\big| < \varepsilon$ for every $x \in X$. It follows that $|f| \in \overline{\mathcal{A}}$, i.e., $|f| \in \mathcal{A}$ since $\mathcal{A}$ is closed. For a general function f, let $M = \max|f|$. Then $|f|/M \in \mathcal{A}$, so $|f| = M(|f|/M) \in \mathcal{A}$. The second assertion follows from the identities

$$\max(f,g) = \tfrac{1}{2}(|f+g| + |f-g|), \quad \min(f,g) = \tfrac{1}{2}(|f+g| - |f-g|). \qquad \blacksquare$$

Definition 5.7.4. A subalgebra $\mathcal{A}$ to $C(X)$ or $C(X,\mathbf{R})$ **separates points** if, for all $y, z \in X$ such that $y \neq z$, there exists a function $g \in \mathcal{A}$ such that $g(y) \neq g(z)$.

Theorem 5.7.5 (The Real Stone–Weierstrass Theorem). *If $\mathcal{A}$ is a closed subalgebra to $C(X,\mathbf{R})$ such that $\mathcal{A}$ separates points and $1 \in \mathcal{A}$, then $\mathcal{A} = C(X,\mathbf{R})$.*

Proof. Let $f \in C(X,\mathbf{R})$. If $y \neq z$, then, since $\mathcal{A}$ separates points, there exists a function $g \in \mathcal{A}$ such that $g(y) \neq g(z)$. The function $f_{y,z}$, defined by

$$f_{y,z}(x) = f(z) + \big(f(y) - f(z)\big)\frac{g(x) - g(z)}{g(y) - g(z)}, \quad x \in X,$$

belongs to $\mathcal{A}$ and satisfies $f_{y,z}(z) = f(z)$. Since f and $f_{y,z}$ are continuous, it follows that for $\varepsilon > 0$, there exists an open ball B_z with center z such that $f_{y,z}(x) < f(x) + \varepsilon$ for every $x \in B_z$. Using the fact that X is compact, we can cover X with a finite number of balls $B_{z_1}, \dots, B_{z_m}$. If $f_y = \min(f_{y,z_1}, \dots, f_{y,z_m})$, then $f_y \in \mathcal{A}$ and $f_y(x) < f(x) + \varepsilon$ for every $x \in X$. Moreover, since $f_y(y) = f(y)$, there is an open ball B_y with center y such that $f_y(x) > f(x) - \varepsilon$ for every $x \in B_y$. By compactness, X can be covered by a finite number of balls $B_{y_1}, \dots, B_{y_n}$. If we let $g = \max(f_{y_1}, \dots, f_{y_n})$, then $g \in \mathcal{A}$ and

$$f(x) - \varepsilon < g(x) < f(x) + \varepsilon,$$

that is, $|f(x) - g(x)| < \varepsilon$ for every $x \in X$. Since ε was arbitrary and $\mathcal{A}$ is closed, this shows that $f \in \mathcal{A}$. $\blacksquare$

Theorem 5.7.6 (The Stone–Weierstrass Theorem). *If $\mathcal{A}$ is a closed subalgebra to $C(X)$ such that $\mathcal{A}$ separates points, $1 \in \mathcal{A}$, and $\overline{f} \in \mathcal{A}$ whenever $f \in \mathcal{A}$, then $\mathcal{A} = C(X)$.*

Without the assumption that $\mathcal{A}$ is closed under complex conjugation, the conclusion in the Stone–Weierstrass Theorem may be false.

Proof. Notice that if $f \in \mathcal{A}$, then $\operatorname{Re} f = \frac{1}{2}(f + \overline{f}) \in \mathcal{A}$ and $\operatorname{Im} f = \frac{1}{2i}(f - \overline{f}) \in \mathcal{A}$. Moreover, the set $\operatorname{Re}(\mathcal{A}) = \{\operatorname{Re} f : f \in \mathcal{A}\}$ is a closed subalgebra to $C(X,\mathbf{R})$ that satisfies the assumptions of Theorem 5.7.5. Thus, $\operatorname{Re}(\mathcal{A}) = C(X,\mathbf{R})$. It follows that $\mathcal{A} = C(X)$ since $\mathcal{A} = \operatorname{Re}(\mathcal{A}) + i\operatorname{Re}(\mathcal{A})$ and $C(X) = C(X,\mathbf{R}) + iC(X,\mathbf{R})$. $\blacksquare$

Corollary 5.7.7 (Weierstrass' Approximation Theorem). *Suppose that K is a non-empty, compact subset to $\mathbf{R}^d$. Then every continuous function on K can be approximated uniformly by polynomials.*

Proof. Let $\mathcal{A}$ be the closure in $C(K)$ of the set of complex polynomials in d variables. Then $\mathcal{A}$ satisfies the assumptions of Theorem 5.7.6, so $\mathcal{A} = C(K)$. ■

Basically the same proof gives the following corollary. Here, a **trigonometric polynomial** is a function of the form $\sum_{n=-N}^{N} c_n e^{int}$, $t \in \mathbf{R}$, where each c_n is a complex number.

Corollary 5.7.8. *Every continuous function on* $\mathbf{R}$ *with period* 2π *can be approximated uniformly with trigonometric polynomials.*

Exercises

E5.1. Finish the proof of Theorem 5.2.5.

E5.2. Show that the set $[-a, a]^d \subset \mathbf{R}^d$ is closed.

E5.3. Prove Proposition 5.3.5.

E5.4. Show that $C[0, T_1; \mathbf{R}^d]$ is complete.

E5.5. Let $f(t) = (t + \varepsilon^2/4)^{1/2}$, $0 \le t \le 1$. Show that the Taylor series of f at $t = \frac{1}{2}$ converges uniformly to f on $[0, 1]$.

E5.6. Show that the Stone–Weierstrass theorem may be false without the assumption that the algebra is closed under complex conjugation.

E5.7. Prove Corollary 5.7.8.

Chapter 6

Cantor's Nested Set Theorem, Baire's Category Theorem

Below, X will denote a metric space with metric d. If E is a subset to X, then the **diameter** of E is the number $\operatorname{diam}(E) = \sup_{x,y\in E} d(x,y)$.

6.1. Cantor's Nested Set Theorem

Theorem 6.1.1. *Suppose that X is complete. If $F_1 \supset F_2 \supset \ldots$ is a decreasing sequence of non-empty, closed subsets to X such that $\operatorname{diam}(F_n) \to 0$, then the intersection $\bigcap_{n=1}^{\infty} F_n$ contains exactly one element.*

Proof. First choose one element x_n in every set F_n, $n = 1, 2, \ldots$. If $m > n$, then $d(x_m, x_n) \le \operatorname{diam}(F_n)$. Since $\operatorname{diam}(F_n) \to 0$, this shows that (x_n) is a Cauchy sequence and hence that $x_n \to x$ for some $x \in X$ because X is complete. Using the facts that $x_m \in F_n$ for $m \ge n$ and every F_n is closed, it follows that $x \in F_n$ for every n, i.e., $x \in \bigcap_{n=1}^{\infty} F_n$. Finally, if $x, y \in \bigcap_{n=1}^{\infty} F_n$, then $d(x,y) \le \operatorname{diam}(F_n)$ for every n, which implies that $d(x,y) = 0$, that is $x = y$. ■

6.2. Baire's Category Theorem

Definition 6.2.1. A subset E to X is **nowhere dense** in X if $\overline{E}^{\circ} = \varnothing$.

Definition 6.2.2. The space X is of the **first category** if $X = \bigcup_{n=1}^{\infty} X_n$, where every set X_n is nowhere dense in X and otherwise of the **second category**.

Example 6.2.3. As a metric space, $\mathbf{Q}$ is of the first category. Indeed, if $(r_n)_{n=1}^{\infty}$ is any enumeration of $\mathbf{Q}$, then $\mathbf{Q} = \bigcup_{n=1}^{\infty}\{r_n\}$ och $\{r_n\}$ is nowhere dense in $\mathbf{Q}$. □

For the case $X = \mathbf{R}^d$, the following theorem was proved by René-Louis Baire in 1894.

Theorem 6.2.4. *Every complete metric space X is of the second category.*

Proof. Suppose on the contrary that $X = \bigcup_{n=1}^{\infty} X_n$, where every set X_n is nowhere dense in X. Take $x_0 \in X$ and put $r_0 = 2$. Since x_0 is not an interior point to $\overline{X_1}$, there exists a point $x_1 \in B_{r_0}(x_0)$ such that $x_1 \notin \overline{X_1}$. Now, using the fact that $\overline{X_1}^{c}$ is open, choose $r_1 \le 1$ such that $\overline{B_{r_1}(x_1)} \cap \overline{X_1} = \varnothing$. We can also assume that $\overline{B_{r_1}(x_1)} \subset \overline{B_{r_0}(x_0)}$. There also exists a point $x_2 \in B_{r_1}(x_1)$ such that $x_2 \notin \overline{X_2}$ and a radius $r_2 \le \frac{1}{2}$ such that $\overline{B_{r_2}(x_2)} \cap \overline{X_2} = \varnothing$ and $\overline{B_{r_2}(x_2)} \subset \overline{B_{r_1}(x_1)}$. Continuing in the same manner, we obtain a decreasing sequence of balls $\overline{B_{r_n}(x_n)}$, where $r_n \le \frac{1}{n}$ and $\overline{B_{r_n}(x_n)} \cap \overline{X_n} = \varnothing$. It now follows from Cantor's theorem that $\bigcap_{n=1}^{\infty} \overline{B_{r_n}(x_n)}$ contains a unique point $x \in X$. This is a contradiction since x does not belong to any set X_n. ■

6.3. Continuous, Nowhere Differentiable Functions

As an application of Baire's theorem, we prove that there exist continuous, nowhere differentiable functions.

Proposition 6.3.1. *There exists a continuous function f on $[0,1]$ that is not differentiable at any point.*

Proof. For $n = 1, 2, \dots$, let E_n denote the set of continuous functions f on $[0,1]$ for which there exists a point $x \in [0, 1 - \frac{1}{n}]$ such that

$$\left| \frac{f(x+h) - f(x)}{h} \right| \le n \quad \text{for } 0 < h < \frac{1}{n}. \tag{1}$$

Notice that if f has a right-hand derivative at $x \in [0,1)$, then $f \in E_n$ for some n. If we can show that every E_n is nowhere dense in $C[0,1]$, it follows from Baire's theorem that $\bigcup_{n=1}^{\infty} E_n$ is not the whole of $C[0,1]$, so there exists a continuous function which does not belong to any set E_n and hence is not differentiable at any point in $[0,1)$. If this function happens to be differentiable at 1, we can easily make it nondifferentiable by adding a continuous function which is differentiable everywhere except at 1.

To show that E_n is nowhere dense in $C[0,1]$, it suffices to show that E_n is closed and that E_n^c is dense in $C[0,1]$. Suppose that $(f_k) \subset E_n$ and that $f_k \to f \in C[0,1]$. For every k, there exists a point $x_k \in [0, 1 - \frac{1}{n}]$ such that

$$\left| \frac{f_k(x_k+h) - f_k(x_k)}{h} \right| \le n \quad \text{for } 0 < h < \frac{1}{n}. \tag{2}$$

From the Bolzano–Weierstrass theorem (Corollary 5.3.4), it follows that (x_k) has a subsequence, which we after renumbering still may denote by (x_k), that converges to $x \in [0, 1 - \frac{1}{n}]$. For a fixed h such that $0 < h < \frac{1}{n}$, let

$$g(t) = \frac{f(t+h) - f(t)}{h} \quad \text{and} \quad g_k(t) = \frac{f_k(t+h) - f_k(t)}{h} \quad \text{for } 0 \le t \le \frac{1}{n}.$$

Then

$$|g(x) - g_k(x_k)| \le |g(x) - g(x_k)| + |g(x_k) - g_k(x_k)| \le |g(x) - g(x_k)| + d_\infty(g, g_k),$$

and since $d_\infty(g, g_k) \to 0$ and g is continuous at x, it follows that $g_k(x_k) \to g(x)$. The inequality (1) thus follows from (2). This shows that f belongs to E_n, so E_n is closed. We finally show that E_n^c is dense in $C[0,1]$. Let $f \in C[0,1]$ and let $\varepsilon > 0$ be arbitrary. From Weierstrass' approxiamation theorem (Corollary 5.7.7), it follows that there exists a polynomial p such that $d_\infty(f, p) < \varepsilon$. Let M be the maximal value of $|p'|$ on $[0,1]$ and let s be a "saw-tooth function" such that $|s| \le \varepsilon$ and the right-hand derivative of s is greater than or equal to $n + M$ everywhere. Then $d_\infty(f, p+s) < 2\varepsilon$ and the absolute value of the right-hand derivative of $p+s$ is greater than or equal to n everywhere, that is, $p+s$ does not belong to E_n. ■

Exercises

E6.1. Formulate and prove a converse to Theorem 6.1.1.

Chapter 7
Banach's Fixed Point Theorem

Let X be a complete metric space with metric d and let F be a mapping from X into X.

7.1. Fixed Points

Definition 7.1.1. An element $x \in X$ is a **fixed point** of F if $F(x) = x$.

Example 7.1.2. If T is a linear mapping on $\mathbf{R}^d$ and x is an eigenvector of T with eigenvalue 1, then x is a fixed point of T since, by definition, $Tx = x$. □

7.2. Lipschitz Mappings and Contractions

Definition 7.2.1. The mapping F is a **Lipschitz mapping** if there exists a constant $L \geq 0$ such that

$$d(F(x), F(y)) \leq Ld(x, y) \quad \text{for all } x, y \in X. \tag{1}$$

The infimum over all such constants L is called the **Lipschitz constant** of F. If the Lipschitz constant is less than 1, F is said to be a **contraction**.

Remark 7.2.2.

(a) Notice that every Lipschitz mapping is uniformly continuous. Instead of saying that F is a Lipschitz mapping, we will sometimes say that F is **Lipschitz continuous**.

(b) In the case $X = \mathbf{R}^d$, the Lipschitz condition (1) takes the form

$$|F(x) - F(y)| \leq L|x - y| \quad \text{for all } x, y \in \mathbf{R}^d.$$

It follows that all derivatives of F (if they exist) are less than or equal L.

Example 7.2.3. Suppose that $F : \mathbf{R} \to \mathbf{R}$ is continuously differentiable satisfying $|F'(x)| \leq L$ for every $x \in \mathbf{R}$. Then, according to the mean value theorem,

$$|F(x) - F(y)| = |F'(\xi)||x - y| \leq L|x - y|$$

for all $x, y \in \mathbf{R}$, which shows that F is a Lipschitz mapping. In particular, if $|F'(x)| \leq \alpha < 1$ for every $x \in \mathbf{R}$, then F is a contraction. □

7.3. The Banach Fixed Point Theorem

The following fixed point theorem not only shows that a contraction on a complete metric space has a unique fixed point, it also gives a practical method for finding the fixed point numerically.

Theorem 7.3.1 (S. Banach 1922). *Every contraction $F : X \to X$ has a uniquely determined fixed point.*

Proof. Let $\alpha < 1$ denote the Lipschitz constant of F. Define the sequence $(x_n)_{n=0}^{\infty}$ by $x_{n+1} = F(x_n),\ n = 0, 1, \ldots$, where $x_0 \in X$ is arbitrary. Then

$$d(x_{k+1}, x_k) = d(F(x_k), F(x_{k-1})) \le \alpha d(F(x_{k-1}), F(x_{k-2})) \le \ldots \le \alpha^k d(x_1, x_0)$$

for $k = 0, 1, \ldots$. If $m > n \ge 0$, this implies that

$$d(x_m, x_n) \le \sum_{k=n}^{m-1} d(x_{k+1}, x_k) \le d(x_1, x_0) \sum_{k=n}^{m-1} \alpha^k < \frac{\alpha^n}{1-\alpha} d(x_1, x_0). \qquad (2)$$

This shows that $(x_n)_{n=0}^{\infty}$ is a Cauchy sequence and hence convergent since X is complete; let $x \in X$ denote the limit of the sequence. Since F is continuous, we have

$$F(x) = F(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} x_{n+1} = x,$$

so x is a fixed point of F. It remains to show that x is unique. If $F(x) = x$ and $F(y) = y$, then

$$d(x, y) = d(F(x), F(y)) \le \alpha d(x, y),$$

which is impossible unless $x = y$. ■

Remark 7.3.2.

(a) The iterations $x_{n+1} = F(x_n),\ n = 0, 1, \ldots$, are called **fixed point iterations**. Notice that these iterations converge however the initial approximation x_0 is chosen.

(b) Letting $m \to \infty$ in (2), we obtain the following error estimate:

$$d(x, x_n) < \frac{\alpha^n}{1-\alpha} d(x_1, x_0).$$

(c) It suffices to assume that $d(F(x), F(y)) < d(x, y)$ for all $x, y \in X$ for the uniqueness part of the theorem.

The next example shows that it is not sufficient that $d(F(x), F(y)) < d(x, y)$ for all $x, y \in X$ for F to have a fixed point.

Example 7.3.3. Let $F(x) = x + \frac{\pi}{2} - \arctan x,\ x \in \mathbf{R}$. Then, by the mean value theorem,

$$|F(x) - F(y)| = \left|1 - \frac{1}{1+\xi^2}\right| |x - y| = \frac{\xi^2}{1+\xi^2}|x - y| < |x - y|,$$

where ξ is between x and y. Since $\xi^2/(1+\xi^2)$ tends to 1 as x and y tend to ∞, this is also the best possible estimate. It is easy to see that F indeed does not have a fixed point. □

7.4. Picard's Existence Theorem

Consider the initial value problem:

$$\begin{cases} y'(t) = f(t, y(t)),\ 0 \le t \le a \\ y(0) = y_0 \end{cases}. \tag{3}$$

Here, f is defined on the set $\{(t, x) \in \mathbf{R} \times \mathbf{R}^d : 0 \le t \le a,\ x \in \mathbf{R}^d\}$ with values in $\mathbf{R}^d$. Moreover, it is assumed that f is continuous and satisfies a Lipschitz condition with respect to the second variable (uniformly with respect to t): There exists a constant $L \ge 0$ such that

$$|f(t, y) - f(t, z)| \le L|y - z| \quad \text{for } 0 \le t \le a \text{ and } y, z \in \mathbf{R}^d.$$

Theorem 7.4.1 (E. Picard 1890). *The problem* (3) *has a unique solution.*

Proof. The initial value problem (3) is equivalent to the following integral equation:

$$y(t) = y_0 + \int_0^t f(\tau, y(\tau))\, d\tau, \quad 0 \le t \le a. \tag{4}$$

In fact, integrating (3) gives (4), and differentiating (4) (which is allowed since y and f are continuous) gives (3). If $F(y)(t)$ denotes the right-hand side of this equation, we are looking for a fixed point of F. Let X denote the space of continuous function on $[0, a]$, with values $\mathbf{R}^d$, equipped with the metric

$$d(x, y) = \max_{0 \le t \le a} |x(t) - y(t)| e^{-2Lt}.$$

We leave it as an exercise to show that this space is complete. It is easy to see that F maps X into X. Moreover, if $y, z \in X$ and $0 \le t \le a$, then

$$\begin{aligned} |F(y)(t) - F(z)(t)| &\le \int_0^t |f(\tau, y(\tau)) - f(\tau, z(\tau))|\, d\tau \le L \int_0^t |y(\tau) - z(\tau)|\, d\tau \\ &\le L \int_0^t |y(\tau) - z(\tau)| e^{-2L\tau} e^{2L\tau}\, d\tau \le L d(y, z) \int_0^t e^{2L\tau}\, d\tau \\ &\le \frac{1}{2} e^{2Lt} d(x, y). \end{aligned}$$

It follows that $d(F(y), F(z)) \le \frac{1}{2} d(x, y)$, so F is a contraction. Banach's fixed point theorem now shows that F has a unique fixed point y. ∎

Example 7.4.2. We will solve the equation $y' = y,\ t \ge 0$, with the initial condition $y(0) = 1$, using fixed point iterations. The right-hand side of the equation satisfies a Lipschitz condition with $L = 1$ on every interval $[0, a]$, so it follows from Theorem 7.4.1, the problem has a unique solution on $[0, \infty)$. The corresponding integral equation is

$$y(t) = 1 + \int_0^t y(\tau)\, d\tau, \quad 0 \le t < \infty.$$

If we take $y_0 = 1$, then

$$\begin{aligned} y_1(t) &= 1 + \int_0^t y_0(\tau)\, d\tau = 1 + t, \\ y_2(t) &= 1 + \int_0^t y_1(\tau)\, d\tau = 1 + t + \frac{t^2}{2}, \\ y_3(t) &= 1 + \int_0^t y_2(\tau)\, d\tau = 1 + t + \frac{t^2}{2} + \frac{t^3}{6}. \end{aligned}$$

Using induction, we see that $y_n(t) = \sum_{k=0}^{n} \frac{t^k}{k!}$, $n = 0, 1, \dots$. As n tends to ∞, $y_n(t)$ tends to the anticipated solution $y(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!} = e^t$. □

Example 7.4.3. The initial value problem in Example 5.6.2 does not have a unique solution. The reason for this is that the right-hand side $f(y) = 2\sqrt{y}$, $y \geq 0$, is not a Lipschitz mapping, for instance since its derivative is unbounded on $(0, \infty)$ (cf. Remark 7.2.2). □

7.5. A Fredholm Equation

Example 7.5.1. We will consider the following integral equation of **Fredholm** type:[1]

$$x(t) - \mu \int_a^b K(t,\tau)x(\tau)\, d\tau = f(t), \quad a \leq t \leq b.$$

Here, $f \in C[a,b]$ and $K \in C([a,b]^2)$ are given functions and μ a parameter; the solution x should belong to $C[a,b]$. If we define

$$F(x)(t) = f(t) + \mu \int_a^b K(t,\tau)x(\tau)\, d\tau, \quad a \leq t \leq b,$$

for $x \in C[a,b]$, then F maps $C[a,b]$ into $C[a,b]$ and the equation is equivalent to the fixed point equation $F(x) = x$. As in the proof of Theorem 7.4.1, one has

$$\|F(x) - F(y)\|_\infty \leq |\mu| \|K\|_\infty \|x - y\|_\infty$$

for all $x, y \in C[a,b]$. Thus, if we assume that $|\mu| \|K\|_\infty < 1$, then F is a contraction, so the equation has a unique solution. □

Exercises

E7.1. Show that the space X of continuous function on an interval $[0,a]$ with values in $\mathbf{R}^d$, equipped with the metric $d(x,y) = \max_{0 \leq t \leq a} |x(t) - y(t)| e^{-2Lt}$, $x, y \in X$, is complete.

[1] Ivar Fredholm (1866–1927), Swedish mathematician

Chapter 8

Normed Spaces, Banach Spaces

In this chapter, X will denote a vector space over a field $\mathbf{K}$ which is either $\mathbf{R}$ or $\mathbf{C}$.

8.1. Normed Spaces

Definition 8.1.1. A **norm** on X is a function $\|\cdot\| : X \to \mathbf{R}$ such that for all $x, y \in X$ the following properties hold:

(i) $\|\cdot\|$ is **positive**: $\|x\| \geq 0$;

(ii) $\|\cdot\|$ is **definite**: if $\|x\| = 0$, then $x = 0$;

(iii) $\|\cdot\|$ is **homogeneous**: $\|\alpha x\| = |\alpha|\|x\|$ for every $\alpha \in \mathbf{K}$;

(iv) $\|\cdot\|$ satisfies the **triangle inequality**: $\|x + y\| \leq \|x\| + \|y\|$.

Equipped with a norm, X is called a **normed space**.

Remark 8.1.2.

(a) Notice that if $x = 0$, then $\|x\| = \|0x\| = 0\|x\| = 0$, so (ii) is in fact an equivalence.

(b) It is easy to see that $d(x, y) = \|x - y\|$, $x, y \in X$, defines a metric on X.

Example 8.1.3. In Chapter 1, the following spaces were introduced as metric spaces. They are, in fact, also normed spaces:

(a) $\mathbf{R}^d$ och $\mathbf{C}^d$ with the norm $\|x\|_2 = \big(\sum_{j=1}^{d} |x^j|^2\big)^{1/2}$, $x = (x^1, \dots, x^d)^t \in \mathbf{R}^d$; other norms on these spaces are $\|x\|_p = \big(\sum_{j=1}^{d} |x^j|^p\big)^{1/p}$, $1 \leq p < \infty$, and $\|x\|_\infty = \max_{1 \leq j \leq d} |x^j|$;

(b) ℓ^p, $1 \leq p < \infty$, with the norm $\|x\|_p = \big(\sum_{j=1}^{\infty} |x^j|^p\big)^{1/p}$;

(c) $\ell^\infty(M)$ with the so called **supremum norm** $\|x\|_\infty = \sup_{t \in M} |x(t)|$;

(d) ℓ^∞, $\mathbf{c}$ and $\mathbf{c}_0$ with the norm $\|x\|_\infty = \sup_{j \geq 1} |x^j|$;

(e) $C_b(X)$, where X is a metric space, and $C(X)$, where X is a compact metric space, with the norm $\|f\|_\infty = \sup_{x \in X} |f(x)|$;

(f) $L^p(E)$, $1 \leq p < \infty$, where E is a measurable subset to $\mathbf{R}^d$, with the norm $\|f\|_p = \big(\int_E |f(x)|^p \, dx\big)^{1/p}$. □

In the rest of this chapter, X denotes a normed space. The following reverse triangle inequality follows directly from Proposition 1.6.1.

Proposition 8.1.4. *For all $x, y \in X$, there holds* $\big|\|x\| - \|y\|\big| \leq \|x - y\|$.

It follows from this inequality that the function $X \ni x \mapsto \|x\| \in \mathbf{R}_+$ is uniformly continuous.

Exercises

E8.1. Show that if X is a normed space, then $d(x,y) = \|x-y\|,\ x,y \in X$, defines a metric on X.

8.2. Banach Spaces

Definition 8.2.1. A **Banach space** is a complete normed space (with the metric given by the norm).

Example 8.2.2. All spaces in Example 8.1.3 are Banach spaces. □

Example 8.2.3. Let $H^\infty(D)$ denote the space of bounded analytic functions on the open unit disk $D = \{z \in \mathbf{C} : |z| < 1\}$ in the complex plane equipped with the norm $\|f\|_\infty = \sup_{z\in D} |f(z)|$. We will prove that $H^\infty(D)$ is a Banach space by showing that $H^\infty(D)$ is a closed subset of $C_b(D)$. Suppose that $f_n \in H^\infty(D)$ and that $f_n \to f \in C_b(D)$. If γ is a simple, closed curve in D, then, by Cauchy's integral theorem,

$$\int_\gamma f(z)\,dz = \lim_{n\to\infty} \int_\gamma f_n(z)\,dz = 0.$$

Morera's theorem now shows that f is analytic. □

Example 8.2.4. According to Example 3.4.1, the space $C[a,b]$ of Riemann integrable functions on an interval $[a,b]$ with the norm $\|f\| = \int_a^b |f(x)|\,dx$ is not a Banach space. □

8.3. Series in Banach Spaces

Definition 8.3.1. A series $\sum_{n=1}^\infty x_n$, where the terms x_n belong to X, is **convergent** with the sum $x \in X$ if $\sum_{n=1}^N x_n \to x$ as $N \to \infty$. The series $\sum_{n=1}^\infty x_n$ is **absolutely convergent** if $\sum_{n=1}^\infty \|x_n\| < \infty$.

Theorem 8.3.2. *A normed space X is a Banach space if and only if every absolutely convergent series in X is convergent.*

Proof. First suppose that X is a Banach space and that $\sum_{n=1}^\infty x_n$ is absolutely convergent. Put $S_N = \sum_{n=1}^N x_n,\ N = 0,1,2,\dots$. If $M > N$, then

$$\|S_M - S_N\| = \Big\| \sum_{n=N+1}^M x_n \Big\| \le \sum_{n=N+1}^M \|x_n\|.$$

Letting $M,N \to \infty$, it follows that $(S_N)_{n=1}^\infty$ is a Cauchy sequence and therefore convergent.

For the converse, let $(x_n)_{n=1}^\infty \subset X$ be a Cauchy sequence. For $k = 1,2,\dots$, choose n_k such that $\|x_m - x_n\| < 2^{-k}$ if $m,n \ge n_k$. The series $\sum_{k=1}^\infty (x_{n_{k+1}} - x_{n_k})$ is then absolutely convergent and hence convergent by assumption. This implies that the limit

$$x = \lim_{p\to\infty} x_{n_{p+1}} = x_{n_1} + \lim_{p\to\infty} \sum_{k=1}^p (x_{n_{k+1}} - x_{n_k}) = x_{n_1} + \sum_{k=1}^\infty (x_{n_{k+1}} - x_{n_k})$$

exists. Finally, since $(x_n)_{n=1}^{\infty}$ is a Cauchy sequence and a subsequence converges to x, the whole sequence converges to x (see Exercise E3.1). ■

Example 8.3.3. Using Theorem 8.3.2, we get a new proof of the completeness of $\mathbf{R}$. Indeed, suppose that $\sum_{n=1}^{\infty} x_n$ is absolutely convergent, where $(x_n)_{n=1}^{\infty} \subset \mathbf{R}$. Put $a_n = |x_n| - x_n$ for $n = 1, 2, \dots$. Then the positive series $\sum_{n=1}^{\infty} a_n$ converges according to one of the comparison tests since $a_n \leq |x_n|$ for every n. But this implies that $\sum_{n=1}^{\infty} x_n$ converges since $x_n = |x_n| - a_n$. □

8.4. Schauder Bases

Definition 8.4.1. A sequence $(e_n)_{n=1}^{\infty} \subset X$ is a **Schauder basis** for X if there for every $x \in X$ exist unique numbers α_n, $n = 1, 2, \dots$, such that

$$x = \sum_{n=1}^{\infty} \alpha_n e_n.$$

Example 8.4.2. The sequence $(\delta_n)_{n=1}^{\infty}$ in Example 5.2.6, defined by $\delta_n^j = 1$ for $j = n$ and $\delta_n^j = 0$ otherwise, is a Schauder basis for ℓ^p, $1 \leq p < \infty$. Indeed, if $x \in \ell^p$, then

$$\left\| x - \sum_{n=1}^{N} x^j \delta_n \right\|^p = \sum_{n=N+1}^{\infty} |x^j|^p \to 0 \quad \text{as } N \to \infty.$$

A Schauder basis is not the same as a **basis** or a **Hamel basis** for a vector space. A basis for a vector space is by a collection of linearly independent vectors such that every vector in the space is a *finite* linear combination of the vectors in the basis. Using Zorn's lemma, one can show that every vector space has a basis. For finite-dimensional spaces, these two concepts of course coincide.

Proposition 8.4.3. *If X is an infinite-dimensional Banach space, then every Hamel basis for X, is uncountable.*

Proof. To produce a contradiction, suppose that $(e_n)_{n=1}^{\infty}$ is a countable Hamel basis for X and put $X_k = \operatorname{span}\{e_1, \dots, e_k\}$, $k = 1, 2, \dots$. Then each X_k is closed and $X = \bigcup_{k=1}^{\infty} X_k$ since $(e_n)_{n=1}^{\infty}$ is a basis for X. But X_k, being finite-dimensional, cannot contain interior points, and is thus nowhere dense in X, which contradicts Baire's theorem (Theorem 6.2.4). ■

Proposition 8.4.4. *If X has a Schauder basis, then X is separable.*

Proof. Finite linear combinations with rational coefficients of the vectors in the basis are dense in X. ■

The proposition implies, for instance, that ℓ^∞ does not have a Schauder basis. In 1973, the Swedish mathematician Per Enflo gave an example of a separable Banach space without a Schauder basis, thus refuting the conjecture by Stefan Banach from 1930 stating that every separable Banach space has a Schauder basis.

8.5. Equivalent Norms and Finite-dimensional Spaces

Definition 8.5.1. Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on X are **equivalent** if there exist two positive constans C and D such that

$$C\|x\|_1 \leq \|x\|_2 \leq D\|x\|_1 \quad \text{for every } x \in X.$$

Example 8.5.2. The norms $\|\cdot\|_p$, $1 \leq p < \infty$, and $\|\cdot\|_\infty$ on $\mathbf{K}^d$ are equivalent. It is in fact easy to see that, for every $x \in \mathbf{K}^d$,

$$\|x\|_\infty \leq \|x\|_p \leq d^{1/p}\|x\|_\infty. \qquad \square$$

More generally, we have the following theorem.

Theorem 8.5.3. *All norms on a finite-dimensional, normed space X are equivalent.*

The converse to the statement in this theorem is in fact also true: If all norms on X are equivalent, then X has to be finite-dimensional. This gives us our first characterization of finite-dimensional, normed spaces: These are exactly those normed spaces for which all norms are equivalent. Another characterization is given in Corollary 8.6.2 below.

The theorem is a direct consequence of the following lemma.

Lemma 8.5.4. *Suppose that* $\dim(X) = d < \infty$ *and that* $e_1, \dots, e_d$ *is a basis for* X. *Put*

$$\|x\|_2 = \left(\sum_{j=1}^{d} |x^j|^2\right)^{1/2} \quad \text{for } x = x^1 e_1 + \dots + x^d e_d \in X.$$

Then $\|\cdot\|$ *and* $\|\cdot\|_2$ *are equivalent.*

Notice that the norm $\|x\|_2$ equals the Euclidian norm of the coordinate vector of $x \in X$.

Proof. To estimate $\|x\|$ with $\|x\|_2$ from above is straightforward:

$$\|x\| = \left\|\sum_{j=1}^{d} x^j e_j\right\| \leq \sum_{1}^{d} |x^j|\|e_j\| \leq \left(\sum_{j=1}^{d} |x^j|^2\right)^{1/2} \left(\sum_{j=1}^{d} \|e_j\|^2\right)^{1/2} = D\|x\|_2,$$

where $D = \left(\sum_{j=1}^{d} \|e_j\|^2\right)^{1/2}$. To prove the reverse inequality, we notice that the set $S = \{x \in X : \|x\|_2 = 1\}$ is compact since the unit sphere $\{(x^1, \dots, x^d)^t \in \mathbf{K}^d : \sum_{j=1}^{d} |x^j|^2 = 1\}$ in $\mathbf{K}^d$ is compact by the Heine–Borel theorem (Theorem 5.2.5). Moreover, the function $S \ni x \mapsto \|x\| \in \mathbf{R}$ is continuous and positive on S, so it follows from Theorem 5.4.2 that there exists a constant $C > 0$ such that $\|x\| \geq C$ for every $x \in S$. Finally, if $x \neq 0$, then $x/\|x\|_2 \in S$, so

$$\left\|\frac{x}{\|x\|_2}\right\| \geq C \quad \text{and hence} \quad \|x\| \geq C\|x\|_2. \qquad \blacksquare$$

Proposition 8.5.5. *Suppose that X is finite-dimensional with* $\dim(X) = d$. *Then the following properties hold*:

(a) *if* $e_1, \dots, e_d$ *is a basis for X, then*

$$x_n = x_n^1 e_1 + \dots + x_n^d e_d \longrightarrow x = x^1 e_1 + \dots + x^1 e_d$$

if and only if $x_n^j \to x^j$ *for* $j = 1, \dots, n$;

(b) *X is complete*;

(c) *a subset K to X is compact if and only if K is bounded and closed.*

Proof.

(a) By Lemma 8.5.4, convergence with respect to $\| \cdot \|$ is the same as convergence with respect to $\| \cdot \|_2$ which in turn is the same as convergence in every coordinate.

(b) If (x_n) is a Cauchy sequence in X, then, again according to Lemma 8.5.4, every sequence of coordinates (x_n^j), $j = 1, \dots, d$, is a Cauchy sequence in $\mathbf{K}$ and hence convergent. By (a), this implies that (x_n) is convergent.

(c) If K is compact, then, by Proposition 5.1.6, K is bounded and closed. Suppose conversely that K is bounded and closed, and let $(x_n) \subset K$. Then (x_n) is bounded with respect to the 2-norm since $\|x_n\|_2 \le C\|x_n\|$ for every n. The Bolzano–Weierstrass Theorem (Corollary 5.3.4) now shows that (x_n) has a subsequence in which every sequence of coordinates converges. Again, (a) implies that (x_n) is convergent. Hence, by Theorem 5.2.3, K is compact. ■

Corollary 8.5.6. *If Y is a finite-dimensional subspace to X, then Y is complete and in particular closed.*

Proof. By (b) in Proposition 8.5.5, Y is complete. Moreover, all complete metric spaces are closed. ■

8.6. Riesz' Lemma

We denote by

$$S_X = \{x \in X : \|x\| = 1\} \quad \text{and} \quad B_X = \{x \in X : \|x\| \le 1\}$$

the unit sphere in X and the closed unit ball in X, respectively.

Lemma 8.6.1 (F. Riesz 1918). *Suppose that $Y \neq X$ is a closed subspace to X and let the number τ satisfy $0 < \tau < 1$. Then there exists an element $\widehat{x} \in S_X$ such that*

$$\|\widehat{x} - y\| \ge \tau \quad \text{for every } y \in Y.$$

Proof. For $x \in X \setminus Y$, put $d = \inf_{y \in Y} \|x - y\|$. Then $d > 0$, because otherwise there would exist a sequence $(y_n) \subset Y$ such that $\|x - y_n\| \to 0$, which would imply that $x \in Y$ since Y is closed. Then choose $\widehat{y} \in Y$ such that $d \leq \|x - \widehat{y}\| \leq d/\tau$ and define $\widehat{x} = (x - \widehat{y})/\|x - \widehat{y}\|$. It now follows that

$$\|\widehat{x} - y\| = \frac{\big\|x - (\widehat{y} + \|x - \widehat{y}\| y)\big\|}{\|x - \widehat{y}\|} \geq \frac{d}{d/\tau} = \tau. \qquad \blacksquare$$

Corollary 8.6.2. *The closed unit ball B_X is compact if and only if* $\dim(X) < \infty$.

Proof. The sufficiency part follows from (c) in Proposition 8.5.5. Now, if B_X is compact, then, by Theorem 5.2.3, B_X is totally bounded, so there exist a finite number of points $x_1, \ldots, x_N \in B_X$ such that $B_X \subset \bigcup_{n=1}^{N} B_{1/2}(x_n)$. Then the set $Y = \operatorname{span}\{x_1, \ldots, x_N\}$ is a closed subspace to X. If $\dim(X) = \infty$, then $Y \neq X$, so according to Riesz' lemma, there exists a point $\widehat{x}$ on the unit sphere S_X such that $\|\widehat{x} - x_n\| \geq \frac{1}{2}$ for every n, which obviously is a contradiction. $\blacksquare$

Exercises

E8.2. Show by an example that the statement Riesz' lemma may be false for $\tau = 1$.

Chapter 9

Hilbert Spaces

Throughout this chapter, X will denote a vector space over a field $\mathbf{K}$ which is either $\mathbf{R}$ or $\mathbf{C}$.

9.1. Inner Product Spaces, Hilbert Spaces

Inner Product

Definition 9.1.1. A function $(\,\cdot\,,\,\cdot\,) : X \times X \to \mathbf{K}$ is called an **inner product** if

(i) the function $(\,\cdot\,, z) : X \to \mathbf{K}$ is linear for every $z \in X$, that is,

$$(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z) \quad \text{for all } x, y \in X,\ \alpha, \beta \in \mathbf{K};$$

(ii) $(x, y) = \overline{(y, x)}$ for all $x, y \in X$;

(iii) $(x, x) \geq 0$ for every $x \in X$;

(iv) $(x, x) = 0$ implies that $x = 0$.

Equipped with an inner product, X is called an **inner product space**.

For the rest of this chapter, X will always denote an inner product space.

Remark 9.1.2.

(a) Notice that if $x = 0$, then by (i), $(x, x) = 0$, so (iv) is really an equivalence.

(b) It follows from (i) and (ii) that, for $x, y, z \in X$ and $\lambda \in \mathbf{K}$,

$$(x, y + z) = (x, y) + (x, z) \quad \text{and} \quad (x, \lambda y) = \overline{\lambda}(x, y).$$

In the case $\mathbf{K} = \mathbf{R}$, this means that $(\,\cdot\,,\,\cdot\,)$ is **bilinear** (linear in both arguments), and in the case $\mathbf{K} = \mathbf{C}$, that $(\,\cdot\,,\,\cdot\,)$ is **sesquilinear** (linear in the first argument but only additive in the second).[1]

Example 9.1.3. Let us give a few examples of inner product spaces:

(a) $\mathbf{K}^d$ with $(x, y) = \sum_{j=1}^d x^j \overline{y^j},\ x, y \in \mathbf{K}^d$;

(b) ℓ^2 with $(x, y) = \sum_{j=1}^\infty x^j \overline{y^j},\ x, y \in \ell^2$; the series is absolutely convergent since $2|x_j \overline{y_j}| \leq |x_j|^2 + |y_j|^2$ for all j;

(c) $L^2(E)$, where $E \subset \mathbf{R}^d$ is measurable, with $(f, g) = \int_E f\overline{g}\, dx,\ f, g \in L^2(E)$; this definition makes sense since $f\overline{g}$ is measurable and belongs to $L^1(E)$ because $2|f\overline{g}| \leq |f|^2 + |g|^2$, where $|f|^2 + |g|^2 \in L^1(E)$. □

---

[1] In Latin, *sesqui* means one *one and a half.*

The Cauchy–Schwarz Inequality

Theorem 9.1.4 (The Cauchy–Schwarz Inequality). *For $x, y \in X$,*

$$|(x,y)|^2 \le (x,x)(y,y).$$

Equality holds if and only if x and y are linearly dependent.

Proof. The inequality obviously holds true if $y = 0$. If $y \neq 0$, put $e = ty$, where $t^{-1} = \sqrt{(y,y)}$. Then $(e,e) = 1$, and

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

from which the Cauchy–Schwarz inequality follows directly. Equality holds if and only if $x - (x,e)e = x - t^2(x,y)y = 0$, which means that x and y are linearly dependent. ∎

Example 9.1.5. We formulate the Cauchy–Schwarz inequality for the spaces in Example 9.1.3:

(a) For $\mathbf{K}^d$: $\left|\sum_{j=1}^{d} x^j \overline{y^j}\right| \le \left(\sum_{j=1}^{d} |x^j|^2\right)^{1/2} \left(\sum_{j=1}^{d} |y^j|^2\right)^{1/2}$.

(b) For ℓ^2: $\left|\sum_{j=1}^{\infty} x^j \overline{y^j}\right| \le \left(\sum_{j=1}^{\infty} |x^j|^2\right)^{1/2} \left(\sum_{j=1}^{\infty} |y^j|^2\right)^{1/2}$.

(c) For $L^2(E)$: $\left|\int_E f\overline{g}\,dx\right| \le \left(\int_E |f|^2\,dx\right)^{1/2} \left(\int_E |g|^2\,dx\right)^{1/2}$. □

The Norm on an Inner Product Space

Definition 9.1.6. For $x \in X$, we define $\|x\| = \sqrt{(x,x)}$.

With this notation, the Cauchy–Schwarz inequality may be written

$$|(x,y)| \le \|x\|\|y\|, \quad x, y \in X.$$

Proposition 9.1.7. *The function $\|\cdot\|$ is a norm on X.*

Proof. Among the properties in Definition 8.1.1, it is only the triangle inequality that really requires a proof. We deduce this from the Cauchy–Schwarz inequality in the following way:

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

The next simple, but useful corollary follows directly from the Cauchy–Schwarz inequality.

Corollary 9.1.8. *The function $(\cdot, z) : X \to \mathbf{K}$ is Lipschitz continuous for every fixed $z \in X$:*

$$|(x,z) - (y,z)| \le \|x - y\|\|z\| \quad \textit{for all} \quad x, y \in X.$$

In vector algebra, the following identity is known as the **parallelogram law**.

Proposition 9.1.9. *For $x, y \in X$, $\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2)$.*

Proof. Expand the left-hand side as in the proof of Proposition 9.1.7. ∎

Thus, if a norm is induced by an inner product, then the norm has to satisfy the parallelogram law. To put it in another way, if a norm does not satisfy the parallelogram law, then it does not come from an inner product.

Example 9.1.10. The standard norm $\|\cdot\|_\infty$ on $C[0,1]$ is not induced by an inner product. Take for instance $f(t) = 1$ and $g(t) = t$ for $0 \leq t \leq 1$. Then

$$\|f+g\|_\infty^2 + \|f-g\|_\infty^2 = 5 \quad \text{but} \quad 2(\|f\|_\infty^2 + \|g\|_\infty^2) = 4. \qquad \square$$

Proposition 9.1.11 (Polarization Identities). *Suppose that $x, y \in X$.*

(a) *If $\mathbf{K} = \mathbf{R}$, then $(x,y) = \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2)$.*

(b) *If $\mathbf{K} = \mathbf{C}$, then $(x,y) = \frac{1}{4}(\|x+y\|^2 + i\|x+iy\|^2 - \|x-y\|^2 - i\|x-iy\|^2)$.*

These identities are proved by expanding the right-hand sides. One can prove that if a norm on a vector space satisfies the parallelogram law, then an inner product on the space can be defined using one of these identities:

Proposition 9.1.12. *If a norm on a vector space satisfies the parallelogram law, then it is possible to define an inner product on the space.*

Exercises

E9.1. Prove Proposition 9.1.12.

Hilbert Spaces

Definition 9.1.13. A **Hilbert space** is a complete inner product space.

Example 9.1.14. According to Example 8.2.2, the spaces in Example 9.1.3 are all Hilbert spaces. □

9.2. Orthogonality

Orthogonality, Orthogonal Complement

Definition 9.2.1.

(a) Two vectors $x, y \in X$ are **orthogonal** if $(x,y) = 0$.

(b) The **orthogonal complement** $Y^\perp$ to a subset Y to X is the set

$$Y^\perp = \{x \in X : (x,y) = 0 \text{ for every } y \in Y\}.$$

Proposition 9.2.2. *For every subset Y to X, the orthogonal complement $Y^\perp$ is a closed subspace to X.*

Proof. Suppose that $u, v \in Y^{\perp}$. Then $(\alpha u + \beta v, y) = \alpha(u, y) + \beta(v, y) = 0$ for all $\alpha, \beta \in \mathbf{K}$ and every $y \in Y$, so $\alpha u + \beta v \in Y^{\perp}$; thus, $Y^{\perp}$ is a subspace to X. To show that $Y^{\perp}$ is closed, suppose that $(x_n) \subset Y^{\perp}$ and that $x_n \to x \in X$. It then follows from Corollary 9.1.8 that $(x, y) = \lim_{n\to\infty}(x_n, y) = 0$ for every $y \in Y$, so $x \in Y^{\perp}$. ∎

The next proposition generalizes Pythagoras' Theorem in classical geometry.

Proposition 9.2.3 (Pythagoras' Theorem). *If $x_1, \dots, x_N$ are pairwise orthogonal, that is, $(x_m, x_n) = 0$ if $m \neq n$, then*

$$\left\|\sum_{n=1}^{N} x_n\right\|^2 = \sum_{n=1}^{N} \|x_n\|^2.$$

Proof. Just expand the left-hand side in the identity using the properties of the inner product and the fact that the vectors are pairwise orthogonal:

$$\left\|\sum_{n=1}^{N} x_n\right\|^2 = \left(\sum_{m=1}^{N} x_m, \sum_{n=1}^{N} x_n\right) = \sum_{m,n=1}^{N} (x_m, x_n) = \sum_{n=1}^{N}(x_n, x_n) = \sum_{n=1}^{N} \|x_n\|^2. \quad ∎$$

Orthonormal Sets

Definition 9.2.4. A subset E to X is **orthonormal** if

$$(e, f) = \begin{cases} 1 & \text{if} \quad e = f \\ 0 & \text{if} \quad e \neq f \end{cases}$$

for all $e, f \in E$. A sequence $(e_n)_{n=1}^{\infty} \subset X$ is orthonormal if the corresponding set $E = \{e_1, e_2, \dots\}$ is orthonormal.

Example 9.2.5. The sequence $(\delta_n)_{n=1}^{\infty} \subset \ell^2$ in Example 5.2.6 is orthonormal:

$$(\delta_m, \delta_n) = \sum_{j=1}^{\infty} \delta_m^j \overline{\delta_n^j} = \begin{cases} 1 \text{ if } m = n \\ 0 \text{ if } m \neq n \end{cases}. \qquad \square$$

Example 9.2.6. The sequence $\left(e^{int}/\sqrt{2\pi}\right)_{-\infty}^{\infty} \subset L^2(-\pi, \pi)$ is orthonormal:

$$\left(\frac{e^{imt}}{\sqrt{2\pi}}, \frac{e^{int}}{\sqrt{2\pi}}\right) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(m-n)t}\, dt = \begin{cases} 1 \text{ if } m = n \\ 0 \text{ if } m \neq n \end{cases}. \qquad \square$$

Lemma 9.2.7. *Suppose that H is a Hilbert space and that $(e_n)_{n=1}^{\infty} \subset H$ is a orthonormal sequence. Then a series $\sum_{n=1}^{\infty} c_n e_n$ is convergent in H if and only if $\sum_{n=1}^{\infty} |c_n|^2 < \infty$.*

Proof. According to Pythagoras' theorem (Theorem 9.2.3),

$$\left\|\sum_{n=N}^{M} c_n e_n\right\|^2 = \sum_{n=N}^{M} |e_n|^2$$

for $M > N$. It follows that the series $\sum_{n=1}^{\infty} c_n e_n$ is convergent in H if and only if $\sum_{n=1}^{\infty} |c_n|^2$ is convergent. ∎

Example 9.2.8. If the sequence $(c_n)_{-\infty}^{\infty} \subset \mathbf{C}$ satisfies $\sum_{-\infty}^{\infty} |c_n|^2 < \infty$, then the function $x(t) = \sum_{-\infty}^{\infty} c_n e^{-int}$, $t \in \mathbf{R}$, belongs to $L^2(-\pi, \pi)$ and has period 2π. Compare this with the following result: If we assume that $\sum_{-\infty}^{\infty} |c_n| < \infty$ (a stronger assumption), then it follows from Weierstrass' theorem that x is continuous on $\mathbf{R}$. □

9.3. Least Distance

Distance to a Convex Subset

In this and the following subsection, H will denote a Hilbert space. A subset K to H is called **convex** if $x, y \in K$ implies that

$$tx + (1-t)y \in H \quad \text{for } 0 \le t \le 1.$$

Theorem 9.3.1. *Let K be a closed, convex subset to H. Then, for every $x \in H$, there exists a unique vector $y \in K$ such that*

$$\|x - y\| = \inf_{z \in K} \|x - z\|.$$

Proof. First choose a sequence $(y_n) \subset K$ such that $\|x - y_n\| \to d = \inf_{z \in K} \|x - z\|$. By the parallelogram law (Theorem 9.1.9),

$$4\left\|x - \frac{y_m + y_n}{2}\right\|^2 + \|y_m - y_n\|^2 = 2(\|x - y_m\|^2 + \|x - y_n\|^2).$$

Since K is convex, $(y_m + y_n)/2 \in K$, so the first term in the left-hand side is at least $4d^2$. On the other hand, the right-hand side tends to $4d^2$, so it follows that $\|y_m - y_n\| \to 0$. If y denotes the limit of the sequence (y_n), then $y \in K$ since K is closed. Moreover, $\|x - y\| = d$ since the norm is continuous. To prove that y is unique, suppose that $\|x - y'\| = d$ for some $y' \in K$. Then, as above,

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

Since the first term in the left member is at least $4d^2$ and the right member is exactly $4d^2$, it follows that $\|y - y'\| = 0$, so $y = y'$. ■

Distance to a Subspace

Theorem 9.3.2. *Suppose that Y is a closed subspace to H. Then*

$$\|x - y\| = \inf_{z \in Y} \|x - z\| \quad \textit{if and only if} \quad (x - y, z) = 0 \quad \textit{for every } z \in Y.$$

Proof. Suppose first that $\|x - y\| = d = \inf_{z \in Y} \|x - z\|$. Given $z \in Y$, choose a scalar $\lambda \in \mathbf{K}$ such that $(x - y, \lambda z) = -|(x - y, z)|$. Then

$$\begin{aligned} d^2 &\le \|(x - y) + t\lambda z\|^2 = \|x - y\|^2 + 2t \operatorname{Re}(x - y, \lambda z) + t^2 |\lambda|^2 \|z\|^2 \\ &= d^2 - 2t|(x - y, z)| + t^2 |\lambda|^2 \|z\|^2 \end{aligned}$$

for every $t \in \mathbf{R}$. This implies that $2|(x - y, z)| \le t|\lambda|^2 \|z\|^2$ for every $t \ge 0$, from which it follows that $(x - y, z) = 0$.

The converse is easier; in fact, by Pythagoras' theorem (Theorem 9.2.3),

$$\|x - z\|^2 = \|(x - y) + (y - z)\|^2 = \|x - y\|^2 + \|y - z\|^2 \geq \|x - y\|^2$$

for every $z \in Y$ since $x - y$ and $y - z$ are orthogonal. ■

9.4. Orthogonal Projections and the Gram–Schmidt Procedure

Orthogonal Projections

Let H denote a Hilbert space. Suppose that Y is a closed subspace to H and let $x \in H$. Then, according to Theorem 9.3.1, there exists a vector $y \in Y$ such that $\|x - y\| = \inf_{z \in Y} \|x - z\|$. Moreover, according to Theorem 9.3.2, this vector satisfies $(x - y, z) = 0$ for every vector $z \in Y$. These two theorems also show that y is uniquely determined by this condition.

Definition 9.4.1. Let Y be a closed subspace to H and let $x \in H$. The unique vector $y \in Y$, that satisfies $(x - y, z) = 0$ for every $z \in Y$, is called the **orthogonal projection** of x on Y. We will denote this vector by $P_Y x$.

Corollary 9.4.2. *If Y is a closed subspace to H, then $H = Y \oplus Y^{\perp}$, that is, for every $x \in H$ there exist unique vectors $y \in Y$ and $z \in Y^{\perp}$ such that $x = y + z$.*

Example 9.4.3. Suppose that $\{e_1, \dots, e_N\} \subset X$ is orthonormal and let Y be the linear span of $\{e_1, \dots, e_N\}$. Then the orthogonal projection of a vector $x \in X$ on Y is $P_Y x = \sum_{n=1}^{N} (x, e_n) e_n$ since $x - P_Y x \perp e_m$ for $m = 1, 2, \dots, N$:

$$(x - P_Y x, e_m) = (x, e_m) - \sum_{n=1}^{N} (x, e_n)(e_n, e_m) = (x, e_m) - (x, e_m) = 0. \qquad \square$$

The Gram–Schmidt Procedure

The **Gram–Schmidt orthogonalization procedure** is probably known to the reader from linear algebra. Given a linearly independent set of vectors $u_1, u_2, \dots$ (finite or infinite) in X, it produces an orthonormal set of vectors $e_1, e_2, \dots$ with the same linear span: Put $e_1 = u_1 / \|u_1\|$ and

$$\begin{cases} f_{n+1} = u_{n+1} - P_{\mathrm{span}\{e_1, \dots, e_n\}} u_{n+1} \\ e_{n+1} = \dfrac{f_{n+1}}{\|f_{n+1}\|} \end{cases} \quad \text{for} \quad n = 1, 2, \dots.$$

Example 9.4.4. It is well-known that if one applies the Gram–Schmidt procedure to the sequence $u_n(t) = t^n$, $0 \leq t \leq 1$, $n = 0, 1, \dots$, in $L^2(0,1)$, one obtains the orthonormal sequence $e_n = \sqrt{n + \frac{1}{2}} P_n$, $n = 0, 1, \dots$, where P_n is the n-th **Legendre polynomial**, given by **Rodriguez' formula**:

$$P_n(t) = \frac{1}{2^n n!} \frac{d^n}{dt^n} (t^2 - 1)^n. \qquad \square$$

Exercises

E9.2. Apply the Gram–Schmidt procedure to the polynomials $1, t, t^2$ in $L^2(0,1)$.

9.5. Complete Orthonormal Systems, Orthonormal Bases

The Finite-Dimensional Case

As the next example shows, it is easy to find the coordinates for a vector in a finite-dimensional inner product space with respect to an orthonormal basis. If one knows the coordinates, the length of the vector can then be calculated by means of Pythagoras' theorem.

Example 9.5.1. Suppose that $\dim(X) = d < \infty$ and that $\{e_1, \dots, e_d\}$ is an orthonormal basis for X. Then every vector $x \in X$ can be written

$$x = x^1 e_1 + \dots + x^d e_d.$$

Taking the inner product of both sides in this identity with $e_n,\ n = 1, \dots, d$, we find that $x^n = (x, e_n)$, so that

$$x = \sum_{n=1}^{d} (x, e_n) e_n.$$

It now follows from Pythagoras' theorem that

$$\|x\|^2 = \sum_{n=1}^{d} |(x, e_n)|^2. \qquad \square$$

We will next investigate to what extent this example can be generalized to infinite-dimensional spaces.

Bessel's Inequality

Theorem 9.5.2 (Bessel's Inequality). *If $(e_n)_{n=1}^{\infty} \subset X$ is orthonormal, then, for every $x \in X$,*

$$\sum_{n=1}^{\infty} |(x, e_n)|^2 \leq \|x\|^2.$$

Proof. According to Example 9.4.3, the orthogonal projection of x on the subspace $\operatorname{span}\{e_1, \dots, e_N\}$ to X is the vector $\sum_{n=1}^{N} (x, e_n) e_n$. Two applications of Pythagoras' theorem now shows that

$$\begin{aligned} \|x\|^2 &= \left\| x - \sum_{n=1}^{N} (x, e_n) e_n \right\|^2 + \left\| \sum_{n=1}^{N} (x, e_n) e_n \right\|^2 \\ &= \left\| x - \sum_{n=1}^{N} (x, e_n) e_n \right\|^2 + \sum_{n=1}^{N} |(x, e_n)|^2 \geq \sum_{n=1}^{N} |(x, e_n)|^2. \end{aligned}$$

Since this inequality holds for any N, Bessel's inequality follows. ■

Example 9.5.3. The Fourier coefficients of $f \in L^2(-\pi,\pi)$ is defined by

$$F(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\,dt, \quad n = 0, \pm 1, \pm 2, \dots .$$

Notice that $F(n) = (f, e_n)/\sqrt{2\pi}$, where e_n is defined in Example 9.2.6. It now follows from Bessel's inequality that

$$\sum_{-\infty}^{\infty} |F(n)|^2 \le \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t)|^2\,dt. \qquad \square$$

Corollary 9.5.4. *If* $(e_n)_{n=1}^{\infty} \subset X$ *is orthonormal, then the series* $\sum_{n=1}^{\infty}(x, e_n)e_n$ *is convergent for every* $x \in X$*.*

Proof. Let $S_N = \sum_{n=1}^{N}(x, e_n)e_n$, $N = 1, 2, \dots$, denote the N-th partial sum to the series $\sum_{n=1}^{\infty}(x, e_n)e_n$. If $M > N$, then

$$\|S_M - S_N\|^2 = \left\| \sum_{n=N+1}^{M} (x, e_n)e_n \right\|^2 = \sum_{n=N+1}^{M} |(x, e_n)|^2,$$

so it follows from Bessel's inequality that $(S_N)_{n=1}^{\infty}$ is a Cauchy sequence and hence convergent. ■

Unconditional Convergence

In this subsection, Z denotes a normed space. A **permutation** of $\mathbf{N}$ is a bijection from $\mathbf{N}$ to $\mathbf{N}$.

Definition 9.5.5. A series $\sum_{n=1}^{\infty} x_n$, where every $x_n \in Z$, is **unconditionally convergent** if $\sum_{n=1}^{\infty} x_{\sigma(n)}$ is convergent in Z with the same sum for every permutation $\sigma : \mathbf{N} \to \mathbf{N}$.

A series $\sum_{n=1}^{\infty} x_n$ is thus unconditionally convergent if it converges to the same sum however the terms are arranged. A classical theorem by Riemann states that for series with real or complex terms, unconditional convergence is the same as absolute convergence.

Theorem 9.5.6 (B. Riemann). *A series with real or complex terms is unconditionally convergent if and only if it is absolutely convergent.*

The corresponding result holds true in every finite-dimensional normed space Z. In general, absolute convergence implies unconditional convergence:

Theorem 9.5.7. *In a normed space, every absolutely convergent series is unconditionally convergent.*

Since we will not use this result, we omit its proof. The converse to this theorem is false: A. Dvoretzky and C.A. Rogers showed in 1950 that in every infinite dimensional normed space, there exists a unconditionally convergent series that is not absolutely convergent. Thus, the property that unconditional and absolute convergence coincide completely characterizes finite-dimensional spaces.

Complete Orthonormal Systems, Parseval's Identity

Let H be a Hilbert space.

Lemma 9.5.8. *Suppose that $E \subset H$ is an orthonormal set and let $x \in H$. Then the set $E_x = \{e \in E : (x, e) \neq 0\}$ is countable.*

Proof. Put $E_m = \{e \in E : |(x, e)| \geq \frac{1}{m}\}$, $m = 1, 2, \dots$. If $e_1, \dots, e_k \in E_m$, it follows from Bessel's inequality that

$$\frac{k}{m^2} \leq \sum_{n=1}^{k} |(x, e_n)|^2 \leq \|x\|^2.$$

Thus, $k \leq m^2\|x\|^2$, so E_m is finite. This implies that the set $E = \bigcup_{m=1}^{\infty} E_m$ is countable. ■

Theorem 9.5.9. *For an orthonormal subset E to H, the following conditions are equivalent.*

(i) *For every $x \in H$, $x = \sum_{e\in E}(x, e)e$, where the series is unconditionally convergent*;

(ii) *For every $x \in H$, $\|x\|^2 = \sum_{e\in E} |(x, e)|^2$*;

(iii) *If $(x, e) = 0$ for every $e \in E$, then $x = 0$.*

Notice that, by Lemma 9.5.8, the series in (i) and (ii) contain only countably many terms. If $E \subset H$ satisfies (i)–(iii), we say that E is a **complete orthonormal system** in H. The identity in (ii) is known as **Parseval's identity**.

Proof. Given $x \in H$, let $(e_n)_{n=1}^{\infty}$ be any enumeration of E_x. We first assume that (i) holds true and deduce (ii). As in the proof of Bessel's inequality,

$$\|x\|^2 - \sum_{n=1}^{N} |(x, e_n)|^2 = \left\| x - \sum_{n=1}^{N}(x, e_n)e_n \right\|^2.$$

The right-hand side tends to 0 as $N \to \infty$, so Parseval's identity holds.

The fact that (ii) implies (iii) is self-evident.

Finally, suppose that (iii) holds. Then, according to Corollary 9.5.4, the series $\sum_{n=1}^{\infty}(x, e_n)e_n$ is convergent; denote the sum by y. Since

$$(x - y, e) = (x, e) - (x, e) = 0$$

for every $e \in E$, we have $y = x$, and hence $x = \sum_{n=1}^{\infty}(x, e_n)e_n$. The series is unconditionally convergent since the sum is x for any enumeration of E_x. ■

Orthonormal Bases

In the case when the set E in Theorem 9.5.9 itself is countable, (i) means that E is an orthonormal basis for H.

Example 9.5.10. The sequence $(\delta_n)_{n=1}^{\infty}$ in Example 5.2.6 is an orthonormal basis for ℓ^2. □

Theorem 9.5.11. *A Hilbert space H has an orthonormal basis if and only if H is separable.*

Proof. If $\dim(X) < \infty$, there is nothing to prove, so suppose that $\dim(X) = \infty$. The necessity part of the theorem then follows directly from Proposition 8.4.4.

For the sufficiency part, suppose that $(e_n)_{n=1}^{\infty}$ is dense in H. After successively removing linearly dependent elements in the sequence, we can assume that $(e_n)_{n=1}^{\infty}$ is linearly independent with span dense in H. Using the Gram–Schmidt procedure, we can also assume that $(e_n)_{n=1}^{\infty}$ is orthonormal. Let $x \in H$ and let $\varepsilon > 0$ be arbitrary. One can then find a vector $y = \sum_{n=1}^{m_\varepsilon} a_n e_n$ such that $\|x - y\| < \varepsilon$. It now follows from the properties of the orthogonal projection that if $N \geq m_\varepsilon$, then

$$\left\| x - \sum_{n=1}^{N} (x, e_n) e_n \right\| \leq \left\| x - \sum_{n=1}^{m_\varepsilon} (x, e_n) e_n \right\| \leq \|x - y\| < \varepsilon.$$

Since ε was arbitrary, this shows that $x = \sum_{n=1}^{\infty} (x, e_n) e_n$. ■

Example 9.5.12. We will show that the sequence $\left(e^{int}/\sqrt{2\pi}\right)_{-\infty}^{\infty}$ is an orthonormal basis for $L^2(-\pi, \pi)$ by verifying (iii) in Theorem 9.5.9. Therefore, suppose that $f \in L^2(-\pi, \pi)$ and that $(f(t), e^{int}) = 0$ for every integer n. Using the fact that continuous functions are dense in $L^2(-\pi, \pi)$, it follows from Corollary 5.7.8 that trigonometric polynomials are also dense in this space. Thus, for an arbitrary $\varepsilon > 0$, there exists a trigonometric polynomial ϕ such that $\|f - \phi\|_2 < \varepsilon$. From Hölder's inequality, we now obtain that

$$\|f\|_2^2 = \int_{-\pi}^{\pi} |f|^2 \, dt = \int_{-\pi}^{\pi} (f - \phi) \overline{f} \, dt \leq \|f - \phi\|_2 \|f\|_2 < \varepsilon \|f\|_2,$$

which implies that $\|f\|_2 < \varepsilon$. Since ε was arbitrary, it follows that $\|f\|_2 = 0$ and hence that $f = 0$.

The properties (i) and (ii) in Theorem 9.5.9 show that, if $f \in L^2(-\pi, \pi)$, then

$$f(t) = \sum_{-\infty}^{\infty} F(n) e^{int} \quad \text{and} \quad \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(t)|^2 \, dt = \sum_{-\infty}^{\infty} |F(n)|^2,$$

where the first series converges unconditionally in $L^2(-\pi, \pi)$. □

Let us point out that there are Hilbert spaces that are non-separable.

Example 9.5.13. The space $\ell^2(\mathbf{R})$ consists of all complex-valued functions x on $\mathbf{R}$ that are 0 outside a countable subset to $\mathbf{R}$ and that satisfies $\sum_{t \in \mathbf{R}} |x(t)|^2 < \infty$. One can show that $\ell^2(\mathbf{R})$ is a Hilbert space with the inner product

$$(x, y) = \sum_{t \in \mathbf{R}} x(t) \overline{y(t)}, \quad x, y \in \ell^2(\mathbf{R});$$

see Exercise E9.3. This space is, however, not separable. In fact, if $\delta_u(t) = 1$ for $t = u$ and $\delta_u(t) = 0$ for $t \neq u$, then $\|\delta_u - \delta_v\| = \sqrt{2}$ if $u \neq v$. Since $(\delta_t)_{t \in \mathbf{R}}$ is uncountable, this implies that $\ell^2(\mathbf{R})$ is non-separable. It is easy to see that $(\delta_t)_{t \in \mathbf{R}}$ is a complete orthonormal system in $\ell^2(\mathbf{R})$. □

The Riesz–Fischer Theorem

Theorem 9.5.14 (Riesz–Fischer). *If H is a separable, infinite-dimensional Hilbert space over the complex numbers, then H is isometrically isomorphic to ℓ^2, i.e., there exists a bijection from H to ℓ^2 that preserves the norm.*

Similarly, if H is a separable, infinite-dimensional Hilbert space over the real numbers, then H is isometrically isomorphic to $\ell^2(\mathbf{N}, \mathbf{R})$. If $\dim(H) = d$, then H is isometrically isomorphic to either $\mathbf{C}^d$ or $\mathbf{R}^d$.

Proof. According to Theorem 9.5.11, H has an orthonormal basis $(e_n)_{n=1}^{\infty}$. Bessel's inequality shows that the mapping $T(x) = ((x, e_n))_{n=1}^{\infty}$, $x \in H$, maps H into ℓ^2. To prove that T is injective, suppose that $T(x) = T(y)$. Then $(x - y, e_n) = 0$ for every n, and hence, by Theorem 9.5.9, $x = y$. The mapping T is also surjective, since if $(c_n)_{n=1}^{\infty} \in \ell^2$, then $T(x) = (c_n)_{n=1}^{\infty} \in \ell^2$ for $x = \sum_{n=1}^{\infty} c_n e_n$, where $x \in H$ by Lemma 9.2.7. Finally, T is isometric since, by Parseval's identity,

$$\|T(x)\|_{\ell^2}^2 = \sum_{n=1}^{\infty} |(x, e_n)|^2 = \|x\|^2. \qquad \blacksquare$$

Exercises

E9.3. Show that $\ell^2(\mathbf{R})$ is a Hilbert space with the inner product $(x, y) = \sum_{t \in \mathbf{R}} x(t)\overline{y(t)}$.

Chapter 10

Bounded Linear Operators

In this chapter, X and Y will denote two normed spaces over a field $\mathbf{K}$ which is either $\mathbf{R}$ or $\mathbf{C}$. The norms in X and Y will both be denoted by $\|\cdot\|$.

10.1. Linear Operators

Definition 10.1.1. A **linear operator** from X to Y is a function $T : X \to Y$ such that

$$T(\alpha x + \beta y) = \alpha T(x) + \beta T(y)$$

for all $x, y \in X$ and all $\alpha, \beta \in \mathbf{K}$. Let $L(X, Y)$ denote the class of linear operators from X to Y. In the case $Y = \mathbf{K}$, we call the elements in $L(X, \mathbf{K})$ **linear functionals**.

We will often write Tx instead of $T(x)$ etc. Notice that if $T \in L(X, Y)$, then $T0 = 0$ (where the first 0 is the zero in X and the second is the zero in Y).

Example 10.1.2. The following operators are both linear.

(a) The operator $T : C[a, b] \to C[a, b]$ is defined by $Tx(t) = \int_a^t x(\tau)\,d\tau,\ a \le t \le b$, for $x \in C[a, b]$.

(b) The subspace $C^1[0, 1]$ to $C[0, 1]$ consists of all continuously differentiable functions on the interval $[0, 1]$. The operator $T : C^1[0, 1] \to C[0, 1]$ is defined by $Tx(t) = x'(t),\ 0 \le t \le 1$, for $x \in C^1[0, 1]$. □

10.2. Bounded operators

Theorem 10.2.1. *The following conditions are equivalent for $T \in L(X, Y)$:*

(i) T *is continuous at* 0;

(ii) T *is continuous*;

(iii) T *is uniformly continuous*;

(iv) T *Lipschitz continuous*;

(v) *there exists a constant $C \ge 0$ such that $\|Tx\| \le C\|x\|$ for every $x \in X$.*

According to Remark 7.2.2, condition (iv) means that there exists a constant $C \ge 0$ such that

$$\|Tx - Ty\| \le C\|x - y\| \quad \text{for all } x, y \in X.$$

Proof. It is easy to see that every condition implies the one above. We therefore assume that (i) holds and deduce (v). By the definition of continuity at 0, using the fact that $T0 = 0$, there exists a number $\delta > 0$ such that $\|Tx\| \le 1$ if $\|x\| \le \delta$. For an arbitrary $x \ne 0$, this implies that

$$\left\| T\left(\delta \frac{x}{\|x\|}\right) \right\| = \frac{\delta}{\|x\|}\|Tx\| \le 1 \quad \text{and hence that} \quad \|Tx\| \le \frac{1}{\delta}\|x\|.$$ ■

Definition 10.2.2.

(a) We say that an operator $T \in L(X, Y)$ is **bounded** if there exists a constant $C \geq 0$ such that $\|Tx\| \leq C\|x\|$ for every $x \in X$. Otherwise, T is said to be **unbounded**.

(b) The class of bounded operators from X to Y is denoted $B(X, Y)$. In the case $X = Y$, we write $B(X)$ instead of $B(X, X)$.

(c) For $T \in B(X, Y)$, the **norm** $\|T\|$ of T is defined by

$$\|T\| = \inf\{C \geq 0 : \|Tx\| \leq C\|x\| \text{ for every } x \in X\}.$$

In Proposition 10.4.1 below, we show that $\|\cdot\|$ really is a norm on $B(X, Y)$. Notice that if $T \in B(X, Y)$, then $\|Tx\| \leq \|T\|\|x\|$ for every $x \in X$.

We leave the proof of the following result to the reader.

Proposition 10.2.3. *For every* $T \in B(X, Y)$,

$$\|T\| = \sup_{x \neq 0} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\|=1} \|Tx\| = \sup_{\|x\| \leq 1} \|Tx\|.$$

Exercises

E10.1. Supply a proof for Proposition 10.2.3.

10.3. Examples

Example 10.3.1 (Continuation of Example 10.1.2).

(a) If $x \in C[a, b]$, then

$$|Tx(t)| \leq \int_a^b |x(\tau)|\, d\tau \leq (b - a)\|x\|_\infty$$

for every $t \in [a, b]$, which implies that $\|Tx\|_\infty \leq (b - a)\|x\|_\infty$, from which it follows that $\|T\| \leq b - a$. On the other hand, if $x(t) = 1$, $a \leq t \leq b$, then $Tx(t) = t - a$, so $\|Tx\|_\infty = b - a = (b - a)\|x\|_\infty$. This shows that T is bounded with $\|T\| = b - a$.

(b) In this case, the operator T is unbounded. Indeed, if $x(t) = t^n$, $0 \leq t \leq 1$, for $n = 1, 2, \ldots$, then $Tx(t) = nt^{n-1}$, which implies that $\|Tx\|_\infty = n = n\|x\|_\infty$, so T cannot be bounded. □

Example 10.3.2. Instead of considering $C^1[0, 1]$ as a subspace to $C[0, 1]$, we can equip $C^1[0, 1]$ with the norm $\|x\|_{C^1[0,1]} = \|x\|_\infty + \|x'\|_\infty$. With this norm, the differentiation operator in Example 10.1.2 (now defined on another domain) is in fact bounded: If $x \in C^1[0, 1]$, then

$$|Tx(t)| = |x'(t)| \leq \|x'\|_\infty \leq \|x\| \quad \text{for every } t \in [0, 1],$$

which implies that $\|Tx\|_\infty \leq \|x\|_{C^1[0,1]}$, so T is bounded with $\|T\| \leq 1$. □

Example 10.3.3. The linear operator $T : C[a,b] \to \mathbf{K}$ is defined by $Tx = x(t_0)$, where $a \le t_0 \le b$, for $x \in C[a,b]$. Since

$$|Tx| = |x(t_0)| \le \|x\|_\infty \quad \text{for every } x \in C[a,b],$$

we see that T is bounded with $\|T\| \le 1$. But $T1 = 1 = \|1\|_\infty$, so $\|T\| = 1$. □

Example 10.3.4. Let $1 \le p \le \infty$. The **left shift operator** $L : \ell^p \to \ell^p$ and the **right shift operator** $R : \ell^p \to \ell^p$ are defined by

$$L(x^1, x^2, ...) = (x^2, x^3, ...) \quad \text{and} \quad R(x^1, x^2, ...) = (0, x^1, x^2, ...)$$

for $x = (x^1, x^2, ...) \in \ell^p$, respectively. We leave it to the reader to show that both L and R are bounded with $\|L\| = \|R\| = 1$ (see Exercise E10.2). □

Example 10.3.5. Given $k \in C[a,b]$, we define the operator $T : C[a,b] \to \mathbf{K}$ by $Tx = \int_a^b k(t)x(t)\,dt$ for $x \in C[a,b]$. Then T is bounded since

$$|Tx| = \left| \int_a^b k(t)x(t)\,dt \right| \le \int_a^b |k(t)|\,dt\,\|x\|_\infty,$$

for every $x \in C[a,b]$, and this also shows that $\|T\| \le \int_a^b |k(t)|\,dt$. To show that the last inequality in fact is an equality, we consider the family of functions

$$x_\varepsilon(t) = \frac{\overline{k(t)}}{|k(t)| + \varepsilon}, \quad a \le t \le b,$$

where $\varepsilon > 0$. Notice that $x_\varepsilon \in C[a,b]$ and $\|x_\varepsilon\|_\infty \le 1$ for every $\varepsilon > 0$. Moreover,

$$Tx_\varepsilon = \int_a^b \frac{|k(t)|^2}{|k(t)| + \varepsilon}\,dt \ge \int_a^b \frac{|k(t)|^2 - \varepsilon}{|k(t)| + \varepsilon}\,dt = \int_a^b |k(t)|\,dt - \varepsilon.$$

From the last inequality, we obtain that

$$\|T\| = \sup_{\|x\|_\infty \le 1} |Tx| \ge \sup_{\varepsilon > 0} |Tx_\varepsilon| \ge \int_a^b |k(t)|\,dt.$$ □

Example 10.3.6. The **Fredholm operator** $T : C[a,b] \to C[a,b]$ is defined by

$$Tx(t) = \int_a^b K(t,\tau)x(\tau)\,d\tau, \quad a \le t \le b,$$

for $x \in C[a,b]$, where the kernel $K \in C([a,b] \times [a,b])$ (see also Example 7.5.1). As in Example 10.3.5, $\|T\| \le \max_{a \le t \le b} \int_a^b |K(t,\tau)|\,d\tau$. This inequality is also an equality. In fact, if $a \le t \le b$ is fixed, then, by the same example,

$$\begin{aligned}\int_a^b |K(t,\tau)|\,d\tau &= \sup_{\|x\|_\infty \le 1} \left| \int_a^b K(t,\tau)x(\tau)\,d\tau \right| \le \sup_{\|x\|_\infty \le 1} \max_{a \le s \le b} \left| \int_a^b K(s,\tau)x(\tau)\,d\tau \right| \\ &= \sup_{\|x\|_\infty \le 1} \|Tx\|_\infty = \|T\|,\end{aligned}$$

from which it follows that $\max_{a \le t \le b} \int_a^b |K(t,\tau)|\,d\tau \le \|T\|$. □

Example 10.3.7. Suppose that $K \in L^2((a,b) \times (a,b))$ and define Tx as in Example 10.3.6, but now for $x \in L^2(a,b)$. It then follows from Hölder's inequality that

$$\begin{aligned}\|Tx\|_2 &= \left(\int_a^b \left|\int_a^b K(t,\tau)x(\tau)\,d\tau\right|^2 dt\right)^{1/2} \\ &\leq \left(\int_a^b \int_a^b |K(t,\tau)|^2\,d\tau dt\right)^{1/2} \left(\int_a^b |x(\tau)|^2\,d\tau\right)^{1/2}\end{aligned}$$

Thus, T maps $L^2(a,b)$ into $L^2(a,b)$ and $\|T\| \leq \|K\|_2$. One can show that this – in general – is not an equality. □

Proposition 10.3.8.

(a) *If* $\dim(X) < \infty$*, then every operator* $T \in L(X,Y)$ *is bounded.*

(b) *If* $\dim(X) = \infty$*, then there exist an unbounded operator* $T : X \to Y$*.*

Proof.

(a) Suppose that $\dim(X) = d$ and let $e_1, \dots, e_d$ be a basis for X. If $x \in X$ and $x = \sum_{j=1}^d x^j e_j$, then, using the triangle inequality and Theorem 8.5.3, it follows that

$$\|Tx\| = \left\|\sum_{j=1}^d x^j Te_j\right\| \leq \max_{1\leq j\leq d} \|Te_j\| \sum_{j=1}^d |x^j| \leq C \max_{1\leq j\leq d} \|Te_j\|\|x\|.$$

(b) Let $\{e_n\}_{n=1}^\infty$ be a linearly independent subset to the unit sphere S_X in X and let $\mathcal{B}$ be a basis for X containing $\{e_n\}_{n=1}^\infty$. For a non-zero element y in Y, define T by $Te_n = ny$, $n = 1, 2, \dots$, and $Tx = 0$ for all other elements x of $\mathcal{B}$, and then extend T to X by linearity. Then T is unbounded since

$$\sup_{\|x\|=1} \|Tx\| \geq \sup_{n=1,2,\dots} \|Te_n\| = \sup_{n=1,2,\dots} n\|y\| = \infty.$$ ■

Exercises

E10.2. Show that the left shift and right shift operators are bounded on ℓ^p, $1 \leq p \leq \infty$, and verify that both have norm 1.

E10.3. Prove that if $K \in C([a,b] \times [a,b])$, then the function $[a,b] \ni t \mapsto \int_a^b |K(t,\tau)|\,d\tau$ is continuous (this was used in Example 10.3.6).

10.4. The Spaces $L(X, Y)$ and $B(X, Y)$

Proposition 10.4.1.

(a) *The space* $L(X,Y)$ *is a vector space.*

(b) *The space* $B(X,Y)$ *is a subspace to* $L(X,Y)$*.*

(c) *The function* $\|\cdot\|$ *is a norm on* $B(X,Y)$*.*

Proof.

(a) For $S, T \in L(X,Y)$ and $\alpha, \beta \in \mathbf{K}$, the operator $\alpha S + \beta T$ is defined by

$$(\alpha S + \beta T)(x) = \alpha Sx + \beta Tx \quad \text{for } x \in X.$$

We leave it to the reader to show that $\alpha S + \beta T \in L(X,Y)$.

(b) Let $S, T \in B(X,Y)$ och $\alpha, \beta \in \mathbf{K}$. Since

$$\|(\alpha S + \beta T)(x)\| \leq |\alpha|\|S\|\|x\| + |\beta|\|T\|\|x\| = (|\alpha|\|S\| + |\beta|\|T\|)\|x\|$$

for every $x \in X$, it follows that the operator $\alpha S + \beta T$ is bounded, that is, $\alpha S + \beta T \in B(X,Y)$.

(c) Recall that $\|T\| = \inf\{C \geq 0 : \|Tx\| \leq C\|x\| \text{ for every } x \in X\}$ for an operator $T \in B(X,Y)$. The norm is obviously positive. It is also definite, since if $\|T\| = 0$, then $\|Tx\| = 0$, i.e., $Tx = 0$ for every $x \in X$, which by definition means that $T = 0$. The homogeneity of the norm is proved in the following way for $\alpha \neq 0$ (the case $\alpha = 0$ being self-evident):

$$\|\alpha T\| = |\alpha| \inf\{\tfrac{C}{|\alpha|} \geq 0 : \|Tx\| \leq \tfrac{C}{|\alpha|}\|x\| \text{ for every } x \in X\} = |\alpha|\|T\|.$$

To prove the triangle inequality, suppose that $S, T \in B(X,Y)$. Then, since

$$\|(S+T)(x)\| \leq (\|S\| + \|T\|)\|x\| \quad \text{for every } x \in X,$$

it follows that $\|S + T\| \leq \|S\| + \|T\|$. ■

Definition 10.4.2. Let $T_n \in B(X,Y)$, $n = 1, 2, \dots$, and $T \in B(X,Y)$. We say that

(i) T_n **converges strongly** to T if $T_n \to T$ in $B(X,Y)$, i.e., $\|T - T_n\| \to 0$;

(ii) T_n **converges pointwise** to T if $T_n x \to Tx$ in Y, i.e., $\|Tx - T_n x\| \to 0$, for every $x \in X$.

Proposition 10.4.3. *Strong convergence implies pointwise convergence.*

Proof. If $T_n \to T$ in $B(X,Y)$ and $x \in X$, then

$$\|Tx - T_n x\| \leq \|T - T_n\|\|x\| \longrightarrow 0.$$

■

Example 10.4.4. In general, pointwise convergence does not imply strong convergence. For instance, if T_n is defined by $T_n x = x^n$ for $x \in \mathbf{c}_0$, then $T_n x \to 0$ for every $x \in \mathbf{c}_0$. On the other hand, $\|T_n\| = \sup_{\|x\| \leq 1} |T_n x| = 1$, from which it follows that T_n does not converge to 0 strongly. □

Theorem 10.4.5. *If* Y *is complete, then* $B(X,Y)$ *is a Banach space.*

Proof. Suppose that $(T_n)_{n=1}^{\infty}$ is a Cauchy sequence in $B(X,Y)$. Then, as in the proof of Proposition 10.4.3, $(T_n x)_{n=1}^{\infty}$ is a Cauchy sequence in Y for every $x \in X$, and therefore $Tx = \lim_{n\to\infty} T_n x$ exists for every $x \in X$. Since limits are linear, we have $T \in L(X,Y)$. For a given $\varepsilon > 0$, choose N so large that $\|T_m - T_n\| < \varepsilon$ if $m, n \geq N$. For $x \in X$, it then follows that

$$\|T_m x - T_n x\| \leq \|T_m - T_n\| \|x\| < \varepsilon \|x\|$$

if $m, n \geq N$. We now let $m \to \infty$ in this inequality using the continuity of the norm (Proposition 8.1.4), to obtain

$$\|Tx - T_n x\| \leq \varepsilon \|x\|, \quad \text{whence} \quad \|T - T_n\| \leq \varepsilon$$

if $n \geq N$. Thus, $T_n \to T$ in $B(X,Y)$. Finally, since

$$\|Tx\| \leq \|Tx - T_N x\| + \|T_N x\| < (\varepsilon + \|T_N\|)\|x\|$$

for every $x \in X$, we see that $T \in B(X,Y)$. ∎

Chapter 11
Duality

In this chapter, X will denote a normed space over a field $\mathbf{K}$ which is either $\mathbf{R}$ or $\mathbf{C}$.

11.1. Definition of the Dual Space

Definition 11.1.1. The space $X' = B(X, \mathbf{K})$ is called the **dual space** to X.

According to our notation, the elements in X' are called **bounded linear functionals** on X. The norm in X' is

$$\|f\| = \sup_{\|x\| \leq 1} |f(x)|, \quad f \in X'.$$

According to Theorem 10.4.5, X' is a Banach space with this norm.

11.2. Examples

Example 11.2.1. According to Examples 10.3.1, 10.3.3, and 10.3.5, the following operators belong to the dual of $C[a, b]$:

(a) the operator $f(x) = \int_a^b x(\tau)\, d\tau,\ x \in C[a, b]$, with norm $b - a$;

(b) the operator $f(x) = x(t_0),\ x \in C[a, b]$, where $a \leq t_0 \leq b$ is fixed, with norm 1;

(c) the operator $f(x) = \int_a^b x(\tau) y(\tau)\, d\tau,\ x \in C[a, b]$, where $y \in C[a, b]$, with norm $\int_a^b |y(\tau)|\, d\tau$. □

11.3. Finite Dimensional Spaces

Theorem 11.3.1. *If* $\dim(X) = d < \infty$, *then* $\dim(X') = d$.

Proof. Let $e_1, \ldots, e_d$ be a basis for X. We define operators $e'_1, \ldots, e'_d$ from X to $\mathbf{K}$ in the following way: if $x = \sum_{j=1}^d x^j e_j$, then $e'_k(x) = x^k$ for $k = 1, \ldots, k$. In particular, $e'_k(e_j) = \delta_{jk},\ 1 \leq j, k \leq d$. It now follows from Lemma 8.5.4 that

$$|e'_k(x)| = |x^k| \leq \left(\sum_{j=1}^d |x^j|^2 \right)^{1/2} \leq C\|x\|$$

for every $x \in X$, showing that $e'_k \in X'$ for $k = 1, \ldots, d$. The operators $e'_1, \ldots, e'_d$ are linearly independent, since if

$$\lambda_1 e'_1 + \ldots \lambda_d e'_d = 0,$$

then, by applying the left-hand side to e_j, we obtain that $\lambda_j = 0$ for $j = 1, \ldots, d$. These operators also span X'. Indeed, if $f \in X'$, then

$$f = \sum_{k=1}^d y^k e'_k, \tag{1}$$

where $y^k = f(e_k)$, $k = 1, \dots, d$; this identity is verified by applying both sides to the basis of X. Thus, $e'_1, \dots, e'_d$ is a basis for f, so $\dim(X') = d$. ■

Notice that it follows from (1) that if $f \in X'$, then

$$f(x) = \sum_{k=1}^{d} x^k y^k, \quad x \in X, \tag{2}$$

where $y^k = f(e_k)$, $k = 1, \dots, d$. It follows conversely from the Cauchy–Schwarz inequality (Theorem 9.1.4) that, for fixed numbers $y^1, \dots, y^d \in \mathbf{K}$, the operator, defined by (2), belongs to X'.

11.4. ℓ^p-spaces

Theorem 11.4.1. *Suppose that $1 \le p < \infty$. Then $f \in (\ell^p)'$ if and only if there exists a unique sequence $y \in \ell^{p'}$ such that*

$$f(x) = \sum_{k=1}^{\infty} x^k y^k \quad \text{for } x \in \ell^p. \tag{3}$$

If $f \in (\ell^p)'$ is given by (3) *for some $y \in \ell^{p'}$, then $\|f\| = \|y\|_{p'}$.*

Proof. We prove the theorem for $1 < p < \infty$ and leave the case $p = 1$ as an exercise. First, suppose that $f \in (\ell^p)'$ and let $(e_k)_{k=1}^{\infty}$ be the standard Schauder basis for ℓ^p. Then

$$f(x) = \sum_{k=1}^{\infty} x^k f(e_k) = \sum_{k=1}^{\infty} x^k y^k$$

for $x \in \ell^p$, where $y_k = f(e_k)$, $k = 1, 2, \dots$. This shows proves that y is uniquely determined. To show that $y \in \ell^{p'}$, put $x^k = |y^k|^{p'}/y^k$ if $y^k \ne 0$ and $x^k = 0$ otherwise. Then, for $1 \le N < \infty$,

$$\sum_{k=1}^{N} |x^k|^p = \sum_{k=1}^{N} |y^k|^{(p'-1)p} = \sum_{k=1}^{N} |y^k|^{p'}. \tag{4}$$

Using this identity and the fact that f is bounded, it now follows that

$$\sum_{k=1}^{N} |y^k|^{p'} = \sum_{k=1}^{N} x^k y^k \le \|f\| \left(\sum_{k=1}^{N} |x^k|^p \right)^{1/p} = \|f\| \left(\sum_{k=1}^{N} |y^k|^{p'} \right)^{1/p},$$

which implies that $y \in \ell^{p'}$ with $\|y\|_{p'} \le \|f\|$. It also follows from (4) that $x \in \ell^p$. Hölder's inequality (Corollary 1.3.2) now shows that

$$|f(x)| \le \sum_{k=1}^{\infty} |x^k y^k| \le \|x\|_p \|y\|_{p'}$$

for $x \in \ell^p$, and hence that $\|f\| \le \|y\|_{p'}$.

For the converse, let f be defined by (3), where $y \in \ell^{p'}$. Hölder's inequality then shows that $f \in (\ell^p)'$ with $\|f\| \le \|y\|_{p'}$. Using the same argument as above, we finally see that $\|y\|_{p'} \le \|f\|$. ■

It follows from Theorem 11.4.1 that the mapping $(\ell^p)' \ni f \mapsto y \in \ell^{p'}$ is a isometric isomorphism. Thus, $(\ell^p)' \cong \ell^{p'}$ for $1 \le p < \infty$. It is often practical to identify f with y. We will adapt this convention in what follows and — when there is no risk of confusion — write $(\ell^p)' = \ell^{p'}$.

Example 11.4.2. Suppose that $y \in \ell^1$ and let

$$f(x) = \sum_{k=1}^{\infty} x^k y^k \quad \text{for } x \in \ell^\infty.$$

This series is absolutely convergent, and since $|f(x)| \le \|x\|_\infty \|y\|_1$ for every $x \in \ell^\infty$, we have that $f \in (\ell^\infty)'$ with $\|f\| \le \|y\|_1$. Now, if $x^k = \overline{y^k}/|y^k|$ if $y^k \ne 0$ and $x^k = 0$ otherwise, then $x \in \ell^\infty$ and $f(x) = \|y\|_1$. Thus, $\|f\| = \|y\|_1$. This shows that $\ell^1 \subset (\ell^\infty)'$ (or more precisely, that ℓ^1 is isometrically isomorphic to a subspace of $(\ell^\infty)'$). □

The following theorem will be proved later.

Theorem 11.4.3. *If X' is separable, then X is also separable.*

Since ℓ^1 is separable, but ℓ^∞ is not, it follows from this theorem that ℓ^1 is a proper subset to $(\ell^\infty)'$. We will not give an exact description of $(\ell^\infty)'$ here.

11.5. L^p-spaces

Example 11.5.1. Suppose that $y \in L^{p'}(a,b)$, where $1 \le p < \infty$, and let

$$f(x) = \int_a^b x(t)y(t)\,dt \quad \text{for } x \in L^p(a,b). \tag{5}$$

It follows directly from Hölder's inequality that

$$|f(x)| \le \|x\|_p \|y\|_{p'}$$

for every $x \in L^p(a,b)$, so $f \in (L^p(a,b))'$ with $\|f\| \le \|y\|_{p'}$. Using a similar argument as in the proof of Theorem 11.4.1, one can actually show that $\|f\| = \|y\|_{p'}$. This shows that $L^{p'}(a,b) \subset (L^p(a,b))'$ for $1 \le p < \infty$. □

The following theorem is a special case of **Riesz' representation theorem**, usually proved in courses on Lebesgue integration. Together with Example 11.5.1, it shows that $(L^p(a,b))' = L^{p'}(a,b)$ for $1 \le p < \infty$.

Theorem 11.5.2. *Let $1 \le p < \infty$. Then, for every $f \in (L^p(a,b))'$, there exists a unique function $y \in L^{p'}(a,b)$ such that*

$$f(x) = \int_a^b x(t)y(t)\,dt \quad \textit{for } x \in L^p(a,b)$$

and $\|f\| = \|y\|_{p'}$.

11.6. The Riesz Representation Theorem for Hilbert Spaces

Let H denote a Hilbert space. Given a fixed element $u \in H$, we define $f(x) = (x, u)$ for $x \in H$. According to the properties of the inner product, f is a linear operator from H to $\mathbf{C}$. The Cauchy–Schwarz inequality (Theorem 9.1.4) also shows that

$$|f(x)| = |(x, u)| \leq \|x\|\|u\|$$

for every $x \in H$, so $f \in H'$ with $\|f\| \leq \|u\|$. Since $f(u) = \|u\|^2$, we actually have $\|f\| = \|u\|$. The following representation theorem, proved by F. Riesz in 1934, shows that every element H' is given in this way as the inner product with some vector in H.

Theorem 11.6.1. *For every $f \in H'$ there exists a unique element $u \in H$ such that*

$$f(x) = (x, u) \quad \textit{for } x \in H.$$

This element satisfies $\|u\| = \|f\|$.

Together with the discussion above, this theorem shows that H' is isometrically isomorphic to H.

Proof. The set $Y = \ker(f) = \{y \in X : f(x) = 0\}$ is a closed subset to H. If $Y = H$, $u = 0$ works in the statement of the theorem. Otherwise, choose an element $z \in Y^{\perp}$ such that $\|z\| = 1$. Every element $x \in H$ may be decomposed as $x = y + \lambda z$, where $y \in Y$ and $\lambda = f(x)/f(z)$ (notice that $f(z) \neq 0$). Since

$$0 = (y, z) = (x - \lambda z, z) = (x, z) - \lambda = (x, z) - \frac{f(x)}{f(z)}$$

it follows that

$$f(x) = (x, \overline{f(z)}z).$$

We therefore take $u = \overline{f(z)}z$. As above, we then have $\|f\| = \|u\|$. To prove uniqueness, suppose that there exist elements $u_1, u_2 \in H$ so that $f(x) = (x, u_1) = (x, u_2)$ for every $x \in H$. Then $(x, u_1 - u_2) = 0$ for every $x \in H$, from which it follows that $u_1 - u_2 = 0$, i.e., $u_1 = u_2$. ■

Example 11.6.2. We state two consequences of Theorem 11.6.1. The reader should compare these with Theorem 11.4.1 and Theorem 11.5.2, respectively.

(a) Every $f \in (\ell^2)'$ may be written

$$f(x) = \sum_{j=1}^{\infty} x^j \overline{y^j}, \quad x \in \ell^2,$$

for some $y \in \ell^2$ with $\|y\|_2 = \|f\|$.

(b) Every $f \in (L^2(a, b))'$ may be written

$$f(x) = \int_a^b x(t)\overline{y(t)}\, dt, \quad x \in L^2(a, b),$$

for some $y \in L^2(a, b)$ with $\|y\|_2 = \|f\|$. □

Exercises

E11.1. Prove Theorem 11.4.1 in the case $p = 1$.

E11.2. Show that $\mathbf{c}' = \mathbf{c}_0' = \ell^1$.

E11.3. Show that if f is defined by (5), where $1 \leq p < \infty$ and $y \in L^{p'}(a,b)$, then $\|f\| = \|y\|_{p'}$.

Chapter 12

The Hahn–Banach Extension Theorem

Let X denote a vector space over a field $\mathbf{K}$, which is either $\mathbf{R}$ or $\mathbf{C}$. Let also Y be a non-trivial subspace to X and f a linear functional on Y. A linear **extension** of f to X is a linear functional F on X such that $F = f$ on Y.

12.1. Semi-norms

Definition 12.1.1. A **semi-norm** on a vector space X is a function $\rho : X \to \mathbf{R}$ such that for all $x, y \in X$ the following properties hold:

(i) $\rho(x) \geq 0$;

(ii) $\rho(\alpha x) = |\alpha|\rho(x)$ for every $\alpha \in \mathbf{K}$;

(iii) $\rho(x + y) \leq \rho(x) + \rho(y)$.

The difference between a semi-norm and a norm is thus that a semi-norm does not have to be definite: it is possible that $\rho(x) = 0$ even if x is not 0.

Example 12.1.2. The function $\rho(x) = \int_0^1 |x(t)|\,dt$, $x \in L^1(\mathbf{R})$, is a semi-norm on $L^1(\mathbf{R})$, but not a norm since $\rho(x) = 0$ only implies that $x = 0$ on $(0, 1)$. □

In what follows, ρ denotes a semi-norm on X.

12.2. The Hahn–Banach Theorem

The following extension theorem for real vector spaces was proved independently by Hahn in 1927 and by Banach in 1929.

Theorem 12.2.1. *Let X be a real vector space. If f is a linear functional, defined on a subspace Y to X, such that $|f| \leq \rho$ on Y, then there exists a linear extension F of f to X such that $|F| \leq \rho$ on X.*

The extension F is in general not uniquely determined (not even in finite dimension).

Example 12.2.2. Let X denote the space of real-valued functions in $L^1(-1, 1)$ and let Y be the subspace to X, consisting of functions that are 0 on $(-1, 0)$. A semi-norm on X is the ordinary norm $\rho(x) = \|x\|_1$ and a linear functional on Y is $f(x) = \int_0^1 x(t)\,dt$. Obviously, $|f| \leq \rho$ on Y. The functional

$$F(x) = a \int_{-1}^{0} x(t)\,dt + \int_0^1 x(t)\,dt$$

extends f to X and satisfies $|F| \leq \rho$ for $|a| \leq 1$. □

The proof of this theorem relies on the following lemma.

Lemma 12.2.3. *Let X be a real vector space and Y a non-trivial subspace to X. If f is a linear functional on Y such that $|f| \le \rho$ on Y and $x \in X \setminus Y$, then there exists a linear extension F of f to* $\operatorname{span}(Y \cup \{x\})$ *such that $|F| \le \rho$ on X.*

Notice that $\operatorname{span}(Y \cup \{x\})$ consists of elements $u = y \pm tx$, where $y \in Y$ and $t \ge 0$.

Proof. Using the linearity of f and the fact that $|f| \le \rho$ on Y, it follows that

$$f(y) + f(z) = f(y + x) + f(z - x) \le \rho(y + x) + \rho(z - x),$$

and hence that

$$f(z) - \rho(z - x) \le \rho(y + x) - f(y)$$

for all $y, z \in Y$. Taking the supremum over z and the infimum over y, it follows that there exists a real number c such that

$$f(z) - \rho(z - x) \le c \le \rho(y + x) - f(y)$$

for all $y, z \in Y$, which in turn implies that $f(y) \pm c \le \rho(y \pm x)$ for every $y \in Y$. Now, for $t > 0$,

$$f(y) \pm tc = t(f(\tfrac{y}{t}) \pm c) \le t(\rho(\tfrac{y}{t}) \pm c) = \rho(y \pm tx).$$

If we define $F(y \pm tx) = f(y) \pm tc$ for $y \in Y$ and $t \ge 0$, we thus obtain an extension F of f to $\operatorname{span}(Y \cup \{x\})$ that satisfies $F(u) \le \rho(u)$ for every $u \in \operatorname{span}(Y \cup \{x\})$. Finally, since

$$-F(u) = F(-u) \le \rho(-u) = \rho(u),$$

we have $|F(u)| \le \rho(u)$ for every $u \in \operatorname{span}(Y \cup \{x\})$. ∎

Proof (Theorem 12.2.1). Let $\mathcal{A}$ denote the class of pairs (Z, F), where Z is a subspace to X such that $Y \subset Z$ and F is an extension of f to Z such that $|F| \le \rho$. Notice that $\mathcal{A} \ne \varnothing$ since $(Y, f) \in \mathcal{A}$. We then introduce a partial order on $\mathcal{A}$ in the following manner: $(Z_1, F_1) \le (Z_2, F_2)$ if $Z_1 \subset Z_2$ and F_2 is an extension of F_1 to Z_2.

Now, let $\mathcal{K}$ be a completely ordered chain in $\mathcal{A}$. Then $V = \bigcup_{(Z,F)\in\mathcal{K}} Z$ is a subspace to X due to the fact that $\mathcal{K}$ is completely ordered. Moreover, the functional G on V, defined by $G(x) = F(x)$ if $x \in Z$, where $(Z, F) \in \mathcal{K}$, is an extension of f such that $G \le \rho$. We also see that (V, G) is an upper bound to $\mathcal{K}$. It thus follows from Zorn's lemma that $\mathcal{A}$ has a maximal element (Z, F). Then $Z = X$, since otherwise (Z, F) would not be maximal according to Lemma 12.2.3. ∎

The Hahn–Banach theorem was generalized to complex spaces by Bohnenblust and Sobczyk in 1938. We shall also call this version the Hahn–Banach theorem.

Theorem 12.2.4. *Let X be a complex vector space. If f is a linear functional, defined on a subspace Y to X, such that $|f| \le \rho$ on Y, then there exists a linear extension F of f to X such that $|F| \le \rho$ on X.*

Proof. We start by writing $f = g + ih$, where g and h are real-valued functionals on Y. Then

$$f(iy) = g(iy) + ih(iy) = ig(y) - h(y),$$

so it follows that $h(y) = -g(iy)$ and hence that $f(y) = g(y) - ig(iy)$ for every $y \in Y$. Denote by $X_{\mathbf{R}}$ the real vector space associated to X. Then, according to the real Hahn–Banach theorem, g has an extension G to $X_{\mathbf{R}}$ that satisfies $|G| \leq \rho$. We now let $F(x) = G(x) - iG(ix)$ for $x \in X$. Then F extends f to X. We leave it to the reader to verify that F is linear. Finally, given $x \in X$, choose $\theta \in \mathbf{R}$ so that $F(x) = e^{i\theta}|F(x)|$. Then

$$|F(x)| = e^{-i\theta}F(x) = F(e^{-i\theta}x) = G(e^{-i\theta}x) \leq \rho(e^{-i\theta}x) = \rho(x). \qquad \blacksquare$$

Exercises

E12.1. Give an example of a linear functional f from $\mathbf{R}$ to $\mathbf{R}$ such that $|f(x)| \leq \|x\|_2$ for every $x \in \mathbf{R}$ which has no unique extension to $\mathbf{R}^2$.

E12.2. Show that the functional F in the proof of Theorem 12.2.4 is linear.

12.3. Extension Theorems for Normed Spaces

In this and the next section, X denotes a normed space over $\mathbf{K}$ and Y a subspace to X. We also assume that $f \in Y'$. We first prove that f can be extended to X without increasing the norm.

Theorem 12.3.1. *There exists an extension $F \in X'$ of f to X which satisfies $\|F\| = \|f\|$.*

Proof. We first define a new norm ρ on X by $\rho(x) = \|f\|\|x\|$, $x \in X$. By the definition of the norm of f, we then have $|f(y)| \leq \rho(y)$ for every $y \in Y$. The Hahn–Banach theorem now shows that f has a linear extension F to X such that

$$|F(x)| \leq \rho(x) = \|f\|\|x\| \quad \text{for every } x \in X.$$

From the last inequality, it follows that F is bounded and $\|F\| \leq \|f\|$. But since F is an extension of F, we also have

$$\|f\| = \sup_{y \in S_Y} |f(y)| = \sup_{y \in S_Y} |F(y)| \leq \sup_{x \in S_X} |F(x)| = \|F\|. \qquad \blacksquare$$

It is possible to prove various versions of Theorem 12.3.1 without the use of the Hahn-Banach theorem by making stronger assumptions. The following result is useful when one has a bounded functional, defined on a dense subspace to a larger space, and wants to extend the operator to the whole space. The procedure is often described by saying that the operator is *extended by continuity.*

Proposition 12.3.2. *If Y is dense in X, then f has a unique extension $F \in X'$ to X such that $\|F\| = \|f\|$.*

Proof. For $x \in X$, choose $(y_n)_{n=1}^{\infty} \subset Y$ such that $y_n \to x$. Then, since

$$|f(y_m) - f(y_n)| \leq \|f\|\|y_m - y_n\| \longrightarrow 0 \quad \text{as } m, n \to \infty$$

and **K** is complete, we see that $F(x) = \lim_{n\to\infty} f(y_n)$ exists. It is easy to see that the limit is independent of which sequence we choose and that the operator F is linear. Moreover,

$$|F(x)| = \lim_{n\to\infty} |f(y_n)| \le \lim_{n\to\infty} \|f\|\|y_n\| = \|f\|\|x\|$$

for every $x \in X$, which shows that F is bounded with $\|F\| \le \|f\|$. As in the last part of the proof of Theorem 12.3.1, one can also show that $\|f\| \le \|F\|$, so we in fact have $\|F\| = \|f\|$. From the independence of the approximating sequence, it also follows that $F = f$ on Y, so F is an extension of f. Finally, if $G \in X'$ is another extension of f, then

$$G(x) = \lim_{n\to\infty} G(y_n) = \lim_{n\to\infty} f(y_n) = F(x)$$

for every $x \in X$, showing that $G = F$. ■

Exercises

E12.3. Prove Theorem 12.3.1 with the extra assumption that X is separable.

E12.4. Prove Theorem 12.3.1 in the case when X is a Hilbert space.

12.4. Consequences of the Hahn–Banach Theorem

In the following corollary, $S_{X'} = \{f \in X' : \|f\| = 1\}$ denotes the unit sphere in X'.

Corollary 12.4.1. *For every $x \in X$,*

$$\|x\| = \sup_{f \in S_{X'}} |f(x)|. \tag{1}$$

Moreover, the supremum is attained for some functional $f \in S_{X'}$.

The supremum in (1) may thus be replaced by maximum. By definition of the norm in the dual space,

$$\|f\| = \sup_{x \in S_X} |f(x)| \quad \text{for } f \in X'.$$

In general, however, is false that $\|f\| = \max_{x\in S_X} |f(x)|$ for $f \in X'$.

Notice that if $x \ne 0$, then it follows from (1) that $f(x) \ne 0$ for some $f \in S_{X'}$, and hence that the dual space to X contains other elements than 0.

Proof. It is clear that the right-hand side of (1) is less than or equal $\|x\|$ for every $x \in X$. To prove the reverse inequality, let $x \in X$ and define g on $\operatorname{span}\{x\}$ by $g(\alpha x) = \alpha\|x\|$ for $\alpha \in \mathbf{K}$. Then, $|g(\alpha x)| = |\alpha|\|x\| = \|\alpha x\|$ for every $\alpha \in \mathbf{K}$, so we have $\|g\| = 1$. Theorem 12.3.1 now shows that there exists an extension $G \in X'$ of g to X such that $\|G\| = 1$. Then

$$\|x\| = |g(x)| = |G(x)| \le \sup_{f\in S_{X'}} |f(x)|,$$

which proves (1). It also follows that $|G(x)| = \|x\|$. ■

Example 12.4.2. For $f \in L^p(E)$, where $1 \leq p < \infty$,

$$\|f\|_p = \max_{\|g\|_{p'}=1} \left| \int_E f(x)g(x)\,dx \right|. \qquad \square$$

The next corollary shows that X' is large enough to *separate points* in X.

Corollary 12.4.3. *If $x_1, x_2 \in X$ and $x_1 \neq x_2$, then there exists a functional $f \in X'$ such that $f(x_1) \neq f(x_2)$.*

Proof. Take $x = x_1 - x_2$ in Corollary 12.4.1. ■

Corollary 12.4.4. *Let Y be a closed subspace of X. Then, for every $x \in X \setminus Y$, there exists a functional $f \in S_{X'}$ such that $f(x) = \operatorname{dist}(x, Y)$ and $f = 0$ on Y.*

Notice that $\operatorname{dist}(x, Y) = \inf_{y \in Y} \|x - y\| > 0$ since Y is closed.

Proof. Define g on $\operatorname{span}(Y \cup x)$ by $g(y + tx) = t \operatorname{dist}(x, Y)$ for $y \in Y$and $t \in \mathbf{R}$. Then $g = 0$ on Y and $g(x) = \operatorname{dist}(x, Y)$. Moreover,

$$|g(y + tx)| = |t| \operatorname{dist}(x, Y) \leq |t| \|x - \left(-\tfrac{y}{t}\right)\| = \|y + tx\|,$$

from which it follows that $\|g\| \leq 1$. We also have

$$\operatorname{dist}(x, Y) = \inf_{y \in Y} |g(x - y)| \leq \|g\| \inf_{y \in Y} \|x - y\| = \|g\| \operatorname{dist}(x, Y),$$

so $\|g\| \geq 1$, and hence $\|g\| = 1$. Let finally $f \in S_{X'}$ be any extension of g to X. ■

Corollary 12.4.5. *For a subspace Z, the following conditions are equivalent X:*

(i) *Z is dense in X;*

(ii) *if $f \in X'$ and $f = 0$ on Z, then $f = 0$.*

Proof. Suppose first that (i) holds. For $x \in X$, take a sequence $(z_n) \subset Z$ such that $z_n \to x$. Then $f(x) = \lim_{n \to \infty} f(z_n) = 0$, which proves that (ii) holds.

For the converse, suppose that (i) does not hold and let $Y = \overline{Z}$. Then there exists $x \in X \setminus Y$ and $f \in S_{X'}$ such that $f(x) \neq 0$ and $f = 0$ on Y. But then (ii) cannot hold. ■

Chapter 13

The Second Dual, Reflexive Spaces

In what follows, X will denote a normed space over a field $\mathbf{K}$, which is either $\mathbf{R}$ or $\mathbf{C}$.

13.1. The Definition

Definition 13.1.1. The space $X'' = (X')'$ is called the **second dual** to X.

Notice that $X'' = B(X', \mathbf{K})$ is complete since $\mathbf{K}$ is complete.

13.2. Reflexive Spaces

Given $x \in X$, we define $Tx \in L(X', \mathbf{K})$ through $Tx(f) = f(x)$ for $f \in X'$. Since

$$|Tx(f)| = |f(x)| \leq \|x\| \|f\|,$$

we see that Tx is bounded on X', i.e., $Tx \in X''$. By varying x, we get a linear mapping T from X to X''. This mapping is in fact isometric (and therefore bounded):

$$\|Tx\| = \sup_{f \in S_{X'}} |Tx(f)| = \sup_{f \in S_{X'}} |f(x)| = \|x\|,$$

where the last equality follows from Corollary 12.4.1. This fact implies that T is injective: If $Tx = 0$, then $\|x\| = \|Tx\| = 0$, and therefore $x = 0$. The mapping T is known as the **canonical embedding** of X into X''. Notice that X is isometrically isomorphic to the image $T(X)$ of X under the canonical embedding: $X \cong T(X)$.

Definition 13.2.1. The space X is called **reflexive** if the canonical embedding T is surjective.

If X is reflexive, then X is thus isometrically isomorphic to X'' via the canonical embedding. R.C. James showed in 1950–1952 that the converse is false: It is possible that $X \cong X''$ without X being reflexive; the bijection from X to X'' may be given by some other mapping than the canonical embedding.

13.3. Examples

Example 13.3.1. If X is not complete, then X cannot be reflexive. □

Theorem 13.3.2. *If X is finite-dimensional, then X is reflexive.*

Proof. According to Theorem 11.3.1, $\dim(X'') = \dim(X') = \dim(X)$. From this fact and the fact that $T : X \to X''$ is injective, it follows that T is surjective. ■

Theorem 13.3.3. *Every Hilbert space H is reflexive.*

Proof. Suppose that $f \in H'$. It then follows from Riesz' representation theorem (Theorem 11.6.1) that there exists a unique vector $y \in H$ such that $f(x) = (x, y)$ for $x \in H$ and $\|y\| = \|f\|$. The mapping $A : H' \to H$, given by $Af = y$ is then a isometric bijection, which is also conjugate linear (additive and conjugate homogeneous). Moreover, H' is a Hilbert space with the inner product

$$(f,g) = (Ag, Af), \quad f, g \in H'.$$

Now, suppose that $F \in H''$. It then follows from Riesz' representation theorem, applied to H', that there exists a unique functional $g \in H'$ such that $F(f) = (f,g)$ for $f \in H'$. If we let $x = Ag$, then

$$F(f) = (f,g) = (Ag, Af) = (x, Af) = f(x) = Tx(f) \quad \text{for every } f \in H',$$

thus showing that $Tx = F$. It follows that T is surjective. ■

Theorem 13.3.4. *For $1 < p < \infty$, ℓ^p is reflexive.*

Proof. In Theorem 11.4.1, we saw that for every $f \in (\ell^p)'$, there exists a unique sequence $y \in \ell^{p'}$ such that

$$f(x) = \langle x, y\rangle \quad \text{for } x \in \ell^p, \quad \text{where} \quad \langle x, y\rangle = \sum_{k=1}^{\infty} x^k y^k.$$

Define the operator $A_p : (\ell^p)' \to \ell^{p'}$ by $A_p f = y$ for $f \in (\ell^p)'$. If $F \in (\ell^p)''$, then we have $x = A_{p'} \circ F \circ A_p^{-1} \in \ell^p$. We now see that

$$F(f) = F(A_p^{-1}y) = F \circ A_p^{-1}(y) = A_{p'}^{-1}x(y) = \langle x, y\rangle = f(x) = Tx(f),$$

which shows that T is surjective. ■

Example 13.3.5. According to Theorem 11.4.1, $(\ell^1)' \cong \ell^\infty$. Now, if ℓ^1 were reflexive, then $(\ell^\infty)' \cong \ell^1$. But this is impossible since ℓ^1 is separable, but ℓ^∞ is not. Thus, ℓ^1 is not reflexive. A similar argument shows that $\mathbf{c}$ och $\mathbf{c}_0$ are not reflexive. □

The proof of the following theorem is almost identical to that of Theorem 13.3.4. As in Exampe 13.3.5, one also shows that $L^1(E)$ is not reflexive.

Theorem 13.3.6. *For $1 < p < \infty$, $L^p(E)$ is reflexive.*

Chapter 14

The Uniform Boundedness Principle, Banach–Steinhaus' Theorem

In the present chapter, X let denote a Banach space and Y a normed space over the real or the complex numbers. Let also $T_n,\ n = 1, 2, ...,$ denote a sequence of operators in $B(X, Y)$ and T an operator in $L(X, Y)$.

14.1. The Uniform Boundedness Principle

Different versions of the following theorem was proved by Hahn 1922, Banach 1922, Hildebrandt 1923, and Banach and Steinhaus 1927.

Theorem 14.1.1 (The Uniform Boundedness Principle). *Suppose that the sequence $(T_n)_{n=1}^{\infty}$ is bounded pointwise.Then $(T_n)_{n=1}^{\infty}$ is also bounded.*

Remark 14.1.2.

(i) The assumption means that $\sup_{n\geq 1} \|T_n x\| < \infty$ for every $x \in X$.

(ii) The conclusion means that $\sup_{n\geq 1} \|T_n\| < \infty$.

Proof. By the assumption, $X = \bigcup_{m=1}^{\infty} X_m$, where

$$X_m = \{x \in X : \|T_n x\| \leq m \text{ for evey } n\}.$$

Notice that X_m is closed since the norm in X and every operator T_n is continuous. Since X is complete, it follows from Baire's theorem (Theorem 6.2.4) that some set X_m has nonempty interior. Thus, there exists a element $x_0 \in X_m$ and a number $r > 0$ such that $x \in X_m$ if $\|x - x_0\| \leq r$. If $\|y\| \leq 1$, then $x = x_0 + ry \in X_m$, so

$$\|T_n y\| = \left\| T\left(\frac{x - x_0}{r}\right) \right\| \leq \frac{2m}{r} \quad \text{for every } n.$$

This shows that $\|T_n\| \leq 2m/r$ for every n. ■

14.2. Banach–Steinhaus sats

Theorem 14.2.1. *If T_n converges to T pointwise, then*

(i) $C = \sup_{n\geq 1} \|T_n\| < \infty$;

(ii) $T \in B(X, Y)$;

(iii) $\|T\| \leq \liminf_{n\to\infty} \|T_n\|$.

Remark 14.2.2. In Example 10.4.4 we saw that if T_n converges to T pointwise, it does not follow that T_n converges to T strongly, i.e., with respect to the norm. This example also shows that the inequality in (iii) may be strict.

Proof.

(i) If the sequence $(T_n x)_{n=1}^{\infty}$ is convergent for every $x \in X$, then $(T_n x)_{n=1}^{\infty}$ is obviously bounded. The conclusion now follows from Theorem 14.1.1.

(ii) Since $\|T_n x\| \le C\|x\|$ for every $x \in X$ and every $n \ge 1$, it follows using the fact that T_n converges to T that $\|Tx\| \le C\|x\|$ for every $x \in X$, and hence that $T \in B(X,Y)$.

(iii) We have

$$\|Tx\| = \lim_{n\to\infty} \|T_n x\| = \liminf_{n\to\infty} \|T_n x\| \le (\liminf_{n\to\infty} \|T_n\|)\|x\|,$$

from which it follows that $\|T\| \le \liminf \_n \to \infty \|T_n\|$. ■

14.3. An Application to Fourier Series

We next present a classic application of Banach–Steinhaus' theorem.

Example 14.3.1. Let $C_{2\pi}(\mathbf{R})$ denote the subspace to $C(\mathbf{R})$ consisting of functions with period 2π. We will show that there exists a function $f \in C_{2\pi}(\mathbf{R})$ whose Fourier series is divergent in 0. For $f \in C_{2\pi}(\mathbf{R})$, the Nth partial sum to the Fourier series of f is

$$S_N f(t) = \sum_{-N}^{N} F(n)e^{int} = \frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(t-\tau)f(\tau)\,d\tau, \quad N = 1, 2, \dots,$$

where D_N is the **Dirichlet kernel**:

$$D_N(\tau) = \sum_{-N}^{N} e^{in\tau} = \frac{\sin(N+\frac{1}{2})\tau}{\sin(\tau/2)}, \quad \tau \in \mathbf{R}.$$

Define $T_N : C_{2\pi}(\mathbf{R}) \to \mathbf{C}$ by $T_N f = S_N f(0)$, $N = 1, 2, \dots$. To prove that there exists a function f such that the Fourier series of f is divergent in 0, it suffices to show that $\sup_{n\ge1} \|T_N\| = \infty$. According to Example 10.3.5, $\|T_N\| = \frac{1}{2\pi}\|D_N\|_1$. Now,

$$\|D_N\|_1 \ge 4\int_0^{\pi} \left|\sin(N+\frac{1}{2})\tau\right| \frac{d\tau}{\tau} = 4\int_0^{(N+\frac{1}{2})\pi} |\sin u|\,\frac{du}{u} \longrightarrow \infty \quad \text{as } N \to \infty.$$

This shows that $\sup_{n\ge1} \|T_N\| = \infty$. □

Remark 14.3.2. The result in Example 14.3.1 was proved by Paul du Bois-Reymond frn 1873. In 1915, Nikolai Luzin conjectured that the Fourier series of a function $f \in L^2(-\pi,\pi)$ converges a.e. This was proved by Lennart Carleson as late as 1966. Carleson's result was generalized in 1968 by Richard A. Hunt to $L^p(-\pi,\pi)$, where $1 < p < \infty$. Much earlier, in 1923, Andrey Kolmogorov had proved that there exists a function in $L^1(-\pi,\pi)$ whose Fourier series diverges a.e.

Chapter 15

Weak and Weak* Convergence

Let X denote a normed space over a field $\mathbf{K}$, which is either $\mathbf{R}$ or $\mathbf{C}$. We say that $x_n \in X$ **converges strongly** to $x \in X$ if x_n converges to x in X, that is if $\|x - x_n\| \to 0$. This type of convergence is as usual denoted $x_n \to x$.

15.1. Weak Convergence

Definition 15.1.1. We say that $x_n \in X$ **converges weakly** to $x \in X$ and write $x_n \rightharpoonup x$ if

$$f(x_n) \to f(x) \quad \text{for every } f \in X'.$$

Example 15.1.2. According to Theorem 11.4.1, $x_n \rightharpoonup x$ in ℓ^p, $1 \le p < \infty$, if $\sum_{j=1}^{\infty} x_n^j y^j \to \sum_{j=1}^{\infty} x^j y^j$ for every sequence $y \in \ell^{p'}$. □

Example 15.1.3. According to Theorem 11.5.1, $f_n \rightharpoonup f$ in $L^p(E)$, $1 \le p < \infty$, if $\int_E f_n g \, dx \to \int_E f g \, dx$ for every function $g \in L^{p'}(E)$. □

The following theorem shows that strong convergence implies weak convergence and that the usual rules for limits also hold for weak limits.

Theorem 15.1.4.

(a) *If $x_n \to x$, then $x_n \rightharpoonup x$.*

(b) *If $x_n \rightharpoonup x$ and $y_n \rightharpoonup y$, then $\alpha x_n + \beta y_n \rightharpoonup \alpha x + \beta y$ for all $\alpha, \beta \in \mathbf{K}$.*

Proof.

(a) If $f \in X'$, then $|f(x) - f(x_n)| \le \|f\| \|x - x_n\| \to 0$.

(b) Again if $f \in X'$, then

$$f(\alpha x_n + \beta y_n) = \alpha f(x_n) + \beta f(y_n) \to \alpha f(x) + \beta f(y) = f(\alpha x + \beta y). \quad \blacksquare$$

15.2. Examples

For finite-dimensional spaces, weak convergence coincides with strong convergence.

Theorem 15.2.1. *If $\dim(X) < \infty$, then every weakly convergent sequence in X is strongly convergent.*

Proof. Let $e_1, \dots, e_d$ be a basis for X and let $e_1', \dots, e_d'$ be the corresponding dual basis, defined by the condition $e_k'(x) = x^k$ for $k = 1, \dots, d$, where x^k is the kth coordinate of $x \in X$ with respect to $e_1, \dots, e_d$ (see Theorem 11.3.1). If x_n converges weakly to x in X, then

$$x_n^k = e_k'(x_n) \to e_k'(x) = x^k \quad \text{for } k = 1, \dots, d.$$

Since all norms on X are equivalent (see Theorem 8.5.3), we now obtain that

$$\|x - x_n\| \leq C\left(\sum_{j=1}^{d} |x^j - x_n^j|^2\right)^{1/2} \longrightarrow 0,$$

showing that x_n converges strongly to x. ■

The following two examples show that weak convergence in general does not imply strong convergence.

Example 15.2.2. We will show that the sequence $(e^{int})_{-\infty}^{\infty} \subset L^2(-\pi,\pi)$ converges weakly to 0. If $f \in (L^2(-\pi,\pi))'$ has the representation $f(x) = \int_{-\pi}^{\pi} x(t)y(t)\,dt$ for $x \in L^2(-\pi,\pi)$, where $y \in L^2(-\pi,\pi)$, then

$$f(e^{int}) = \int_{-\pi}^{\pi} y(t)e^{int}\,dt = 2\pi Y(-n),$$

where $Y(n)$ denote the Fourier coefficients of y. Since $Y(-n) \to 0$ as $|n| \to \infty$ (which, for instance, follows from Bessel's inequality), we see that $e^{int} \rightharpoonup 0$. The sequence does not, however, converge strongly to 0 since $\|e^{int}\|_2 = \sqrt{2\pi}$. □

Example 15.2.3. We next show that the standard basis $(e_n)_{n=1}^{\infty}$ for ℓ^p, where $1 < p < \infty$ (see Example 8.4.2), converges weakly to 0. Suppose that $f \in (\ell^p)'$ is given by $f(x) = \sum_{j=1}^{\infty} x^j y^j$ for $x \in \ell^p$, where $y \in \ell^{p'}$. Then $f(e_n) = y^n \to 0$ since $\sum_{j=1}^{\infty} |y^j|^{p'} < \infty$. The sequence does not converge strongly to 0 since, for every n, $\|e_n\|_p = 1$. □

The following theorem, proved by I. Schur in 1920, shows that the result in the previous example is false in the case $p = 1$. Since we will not use this result, we omit its proof.

Theorem 15.2.4. *In ℓ^1, every weakly convergent sequence is strongly convergent.*

15.3. Further Properties of Weak Convergence

The next theorem shows that weak limits are unique and weakly convergent sequences are bounded.

Theorem 15.3.1. *Suppose that X is complete.*

(a) *If $x_n \rightharpoonup x$ and $x_n \rightharpoonup y$, then $x = y$.*

(b) *If $x_n \rightharpoonup x$, then $(x_n)_{n=1}^{\infty}$ is bounded and $\|x\| \leq \liminf_{n\to\infty} \|x_n\|$.*

In the proof of (a), we use the following lemma which is a consequence of Corollary 12.4.1. It shows that the dual space X' of X is large enough to separate the points of X.

Lemma 15.3.2. *If $x, y \in X$ and $f(x) = f(y)$ for every $f \in X'$, then $x = y$.*

Proof. We have $f(x-y)=0$ for every $f \in X'$ and hence

$$\|x-y\| = \sup_{\|f\|\le 1} |f(x-y)| = 0. \qquad \blacksquare$$

Proof (Theorem 15.3.1).

(a) This follows directly from Lemma 15.3.2 since it follows from the assumption that $f(x)=f(y)$ for every $f \in X'$.

(b) If T denotes the canonical embedding of X into X'' (see Section 13.2), then it follows that $Tx_n(f) \to Tx(f)$ for every $f \in X'$. The Banach–Steinhaus theorem (Theorem 14.2.1) then shows that

$$\sup_{n\ge 1} \|Tx_n\| < \infty \quad \text{and} \quad \|Tx\| \le \liminf_{n\to\infty} \|Tx_n\|.$$

The statement now follows from the fact that T is an isometry. ■

A necessary condition for a sequence to be weakly convergent is thus that the sequence is bounded. In some cases it is also possible to find sufficient conditions for weak convergence.

Example 15.3.3. Let $(x_n)_{n=1}^{\infty}$ be a bounded sequence in ℓ^p, $1<p<\infty$. If $x_n \rightharpoonup x \in \ell^p$ and $f_j \in \ell^p$, $j=1,2,\dots$, is defined by $f_j(x)=x^j$ for $x \in \ell^p$, then $x_n^j = f_j(x_n) \to f_j(x) = x^j$ for every j.

Suppose conversely $x_n^j \to x^j$ for every j. Let $f(x)=\sum_{j=1}^{\infty} x^j y^j$ for $x \in \ell^p$, where $y \in \ell^{p'}$. Then

$$\begin{aligned}|f(x)-f(x_n)| &\le \sum_{j=1}^{\infty} |x^j - x_n^j||y^j| \le \sum_{j=1}^{N} |x^j - x_n^j||y^j| + \sum_{j=N+1}^{\infty} (|x^j|+|x_n^j|)|y^j| \\ &\le \sum_{j=1}^{N} |x^j - x_n^j||y^j| + (\|x\|_p + \|x_n\|_p)\Bigl(\sum_{j=N+1}^{\infty} |y^j|^{p'}\Bigr)^{1/p'}.\end{aligned}$$

According to Theorem 15.3.1, there exists a constant C such that $\|x_n\| \le C$ for every n. We now see that if we first choose N and after that n large enough, then $|f(x)-f(x_n)|$ can be made arbitrarily small. Thus, $x_n \rightharpoonup x$. □

Example 15.3.4. Lt $(f_n)_{n=1}^{\infty} \subset C[a,b]$ be bounded. If $f_n \rightharpoonup f \in C([a,b])$, then f_n converges pointwise to f, i.e., $f_n(t) \to f(t)$ for every $t \in [a,b]$. This follows if we apply the functional δ_t, $a \le t \le b$, defined by $\delta_t(f) = f(t)$, to the sequence. It is also true, but harder to show, that pointwise convergence of bounded sequences in $C[a,b]$ implies weak convergence. □

15.4. Weak* Convergence

Below, $(f_n)_{n=1}^{\infty}$ will denote a sequence in X'. We say that f_n **converges strongly** to $f \in X'$, and write $f_n \to f$, if $\|f - f_n\| \to 0$, where $\|\cdot\|$ is the norm in X'.

Definition 15.4.1. The sequence $(f_n)_{n=1}^{\infty}$ **converges weak*** to $f \in X'$ if

$$f_n(x) \to f(x) \quad \text{for every } x \in X.$$

This convergence is denoted $f_n \overset{*}{\rightharpoonup} f$.

Thus, weak* convergence in X' coincides with pointwise convergence. The following theorem summarizes some properties of weak* convergence; we leave the proof to the reader.

Theorem 15.4.2.

(a) *If $f_n \to f$, then $f_n \overset{*}{\rightharpoonup} f$.*

(b) *If $f_n \overset{*}{\rightharpoonup} f$ and $g_n \overset{*}{\rightharpoonup} g$, then $\alpha f_n + \beta g_n \overset{*}{\rightharpoonup} \alpha f + \beta g$ for all $\alpha, \beta \in \mathbf{K}$.*

(c) *If $f_n \overset{*}{\rightharpoonup} f$ and $f_n \overset{*}{\rightharpoonup} g$, then $f = g$.*

(d) *If $f_n \overset{*}{\rightharpoonup} f$, then $(f_n)_1^{\infty}$ is bounded and $\|f\| \leq \liminf_{n\to\infty} \|f_n\|$.*

Remark 15.4.3. In general, weak* convergence is not the same as weak convergence in X'. First of all, $f_n \rightharpoonup f$ in X' if $F(f_n) \to F(f)$ for every $F \in X''$. Now, every $x \in X$ can be identified with an element of X'', so if $f_n \rightharpoonup f$, then $f_n \overset{*}{\rightharpoonup} f$, i.e., weak convergence implies weak* convergence. If X is reflexive, these concepts, however, coincide.

Chapter 16

The Open Mapping Theorem and the Closed Graph Theorem

Let X and Y be Banch spaces over a field $\mathbf{K}$, which is either $\mathbf{R}$ or $\mathbf{C}$, and let T be a linear operator from X to Y. Consider the equation

$$Tx = y, \tag{1}$$

where $x \in X$ is unknown and the right-hand side $y \in Y$ is known. If T is invertible with inverse $T^{-1} : Y \to X$, then the equation has exactly one solution $x = T^{-1}y$ for every y. Moreover, if T^{-1} is bounded, then

$$\|x\| = \|T^{-1}y\| \leq \|T^{-1}\|\|y\|.$$

It follows that the problem (1) is numerically **stable** in the sense that small errors in the right-hand side of (1) give rise to small errors in the solution.

16.1. Open Mappings

Suppose that T is bounded and invertible; we will investigate if the same holds for the operator T^{-1}. Notice that T^{-1} is bounded, i.e., continuous, if and only if $(T^{-1})^{-1}(G) = T(G)$ is open for every open subset G to X.

Definition 16.1.1. The operator T is said to be **open** if $T(G)$ is open for every open subset G to X.

Example 16.1.2. Every bounded operator is not open. For instance, the operator $T : \ell^\infty \to \mathbf{c}_0$, which maps $x = (x^j)_{j=1}^\infty \in \ell^\infty$ to $Tx = (x^j/j)_{j=1}^\infty \in \mathbf{c}_0$, is not open since $T(B_1(0)) = \{y \in \mathbf{c}_0 : |y^j| \leq c/j \text{ for } j = 1, 2, \dots, \text{ for some } c < 1\}$ does not contain an open ball with center 0. □

Theorem 16.1.3. *Every open mapping T is surjective.*

Proof. Since T is open and $T0 = 0$, $B_\varepsilon(0) \subset T(B_1(0))$ for some number $\varepsilon > 0$. From the linearity of T, it now follows that $B_R(0) \subset T(B_{R/\varepsilon}(0))$ for every $R > 0$, which shows that T is surjective. ■

16.2. The Open Mapping Theorem

Lemma 16.2.1. *If $T \in B(X, Y)$ is surjective, then there exists a number $\varepsilon > 0$ such that $B_\varepsilon(0) \subset T(B_1(0))$.*

Proof. We begin by showing that

$$B_\varepsilon(0) \subset \overline{T(B_{1/2}(0))} \quad \text{for some} \quad \varepsilon > 0. \tag{2}$$

Since T is surjective, $Y = \bigcup_{n=1}^\infty T(B_{n/2}(0))$. Baire's category theorem (Theorem 6.2.4) now shows that some set $\overline{T(B_{n/2}(0))}$ has non-empty interior, so there exists a vector $y \in Y$ and a number $r > 0$ such that $B_r(y) \subset \overline{T(B_{n/2}(0))}$. Suppose that $\|z\| < r$. Then $\|(z+y) - y\| < r$ and hence $z + y \in \overline{T(B_{n/2}(0))}$. Since also $B_r(-y) \subset \overline{T(B_{n/2}(0))}$, we similarly have $z - y \in \overline{T(B_{n/2}(0))}$. Using the fact that $\overline{T(B_{n/2}(0))}$ is convex, we see that $z \in \overline{T(B_{n/2}(0))}$, and hence that $B_r(0) \subset \overline{T(B_{n/2}(0))}$. This proves (2) with $\varepsilon = r/n$.

Now, if $y \in B_\varepsilon(0)$, then $\|y - Tx_1\| < 2^{-1}\varepsilon$ for some vector $x_1 \in B_{1/2}(0)$. Thus, $y - Tx_1 \in B_{\varepsilon/2}(0) \subset \overline{T(B_{1/4}(0))}$. Using induction, we can find a sequence $(x_n)_{n=1}^\infty \subset X$ such that $\|x_n\| < 2^{-n}$ and $\|y - T(\sum_{k=1}^n x_k)\| < 2^{-n}\varepsilon$ for $n = 1, 2, \ldots$. Since $\|x_n\| < 2^{-n}$, the series $x = \sum_{k=1}^\infty x_k$ converges absolutely and $x \in B_1(0)$. We also have

$$\|y - Tx\| = \lim_{n\to\infty} \Big\| y - T\Big(\sum_{k=1}^n x_k\Big)\Big\| \le \lim_{n\to\infty} 2^{-n}\varepsilon = 0,$$

so $y = Tx$. ■

The following theorem was proved by J. Schauder in 1930.

Theorem 16.2.2 (The Open Mapping Theorem). *An operator $T \in B(X, Y)$ is surjective if and only if it is open.*

Proof. The sufficiency part follows from Theorem 16.1.3. Now suppose that T is surjective. Let $G \subset X$ be open and let $y = Tx \in T(G)$. Since G is open, there exists a number $r > 0$ such that $B_r(x) \subset G$. Then $B_r(0) \subset G - x$, where we use the notation $G - x = \{g - x : g \in G\}$. According to Lemma 16.2.1, there exists a number $s > 0$ such that $B_s(0) \subset T(B_r(0))$. But since $B_r(0) \subset G - x$, this implies that

$$B_s(0) \subset T(G - x) = T(G) - Tx = T(G) - y,$$

and hence that $B_s(y) \subset T(G)$. This proves that $T(G)$ is open. ■

16.3. The Inverse Mapping Theorem

The inverse mapping theorem was proved by S. Banach in 1929.

Theorem 16.3.1 (The Inverse Mapping Theorem). *If $T \in B(X, Y)$ is bijective, then $T^{-1} \in B(Y, X)$.*

Proof. Accoording to Theorem 16.2.2, T is open, so T^{-1} is bounded. ■

Example 16.3.2. Let us define the operator $T : C[0,1] \to C[0,1]$ by

$$Tx(t) = \int_0^t x(\tau)\,d\tau, \ 0 \le t \le 1, \quad \text{for } x \in C[0,1].$$

According to Example 10.3.1, T is bounded. Moreover, T is injective: If $Tx = 0$, then $Tx(t) = \int_0^t x(\tau)\,d\tau = 0$ for $0 \le t \le 1$, which after differentiation shows that $x = 0$. The range of T is $R(T) = \{y \in C^1[0,1] : y(0) = 0\}$. As a mapping from $C[0,1]$ to $R(T)$, T is invertible with the inverse $T^{-1}y = y'$, $y \in R(T)$. However, the inverse of T is not bounded (see Example 10.3.1), which seems to contradict the inverse mapping theorem. The explanation $R(T)$ is not a closed subset to $C[0,1]$ and therefore not complete. □

Example 16.3.3. For $f \in L^1(-\pi,\pi)$, we define $Tf = (F(n))_{-\infty}^{\infty}$, where $F(n)$ denote the Fourier coefficients of f. According to the Riemann–Lebesgue lemma, $F(n) \to 0$ as $|n| \to \infty$, so $T : L^1(-\pi,\pi) \to \mathbf{c}_0$. It also follows from the uniqueness theorem for Fourier coefficients that T is injective.

Now suppose that T were surjective. Then $T^{-1} : \mathbf{c}_0 \to L^1(-\pi,\pi)$ would be bounded, so there would exist a constant $C > 0$ so that $\|Tf\|_\infty \geq C\|f\|_1$ for every $f \in L^1(-\pi,\pi)$. But for the Dirichlet kernel $D_N \in L^1(-\pi,\pi)$ (see example 14.3.1), $\|TD_N\|_\infty = 1$, while $\|D_N\|_1 \to \infty$ as $N \to \infty$. This shows that T is not surjective, so the range of T is not the whole of $\mathbf{c}_0$. □

Corollary 16.3.4. *Suppose that X is complete with respect to two norms $\|\cdot\|_1$ and $\|\cdot\|_2$. If there exists a constant $D > 0$ such that*

$$\|x\|_1 \leq D\|x\|_2 \quad \textit{for every } x \in X,$$

then the norms $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent, that is there exists a constant $C > 0$ such that

$$C\|x\|_2 \leq \|x\|_1 \quad \textit{for every } x \in X.$$

Proof. Let $X_j = (X, \|\cdot\|_j)$, $j = 1, 2$. The operator $T : X_2 \to X_1$, which is given by $Tx = x$, $x \in X_2$, is obviously both bijective and bounded. According to Theorem 16.3.1, the operator $T^{-1} : X_1 \to X_2$ is therefore bounded, so there exists a constant $K > 0$ such that $\|x\|_2 = \|T^{-1}x\|_2 \leq K\|x\|_1$. ■

Example 16.3.5. Let $X = \{x \in C[0,1] : x(0) = 0\}$. Then X is complete with respect to the two norms $\|x\|_1 = \|x'\|_\infty$ and $\|x\|_2 = \|x\|_\infty + \|x'\|_\infty$ $(x \in X)$. Moreover, $\|x\|_1 \leq \|x\|_2$ for every $x \in X$. It thus follows from Corollary 16.3.4 that there exists a constant $C > 0$ such that $\|x\|_1 \geq C\|x\|_2$ for every $x \in X$. In particular, $\|x'\|_\infty \geq C\|x\|_\infty$. Of course, this follows directly from the representation

$$x(t) = \int_0^t x'(\tau)\,d\tau, \quad 0 \leq t \leq 1.$$ □

16.4. The Closed Mapping Theorem

Definition 16.4.1. The operator T is said to be **closed** if

$$\left\{ \begin{array}{c} X \ni x_n \to x \in X \\ Tx_n \to y \in Y \end{array} \right\} \quad \text{implies that} \quad y = Tx.$$

Remark 16.4.2. The operator T is continuous if $X \ni x_n \to x \in X$ implies that $Tx_n \to Tx$. Every bounded, i.e., continuous operator T is therefore closed.

Example 16.4.3. Let $D = C^1[0,1]$ considered as a subset to $C[0,1]$. We will show that the operator $T : D \to C[0,1]$, given by $Tx = x'$, $x \in D$, is closed. Suppose therefore that $x_n \to x$ and $Tx_n \to y$ i $C[0,1]$, i.e., x_n and x_n' tend uniformly to x and y, respectively. According to a well-known theorem in analysis, it now follows that $y \in C^1[0,1]$ and $y = x' = Tx$. □

Definition 16.4.4.

(a) A norm on $X \times Y$ is given by $\|(x, y)\| = \|x\| + \|y\|$, $(x, y) \in X \times Y$

(b) The **graph** of T is the set $G_T = \{(x, Tx) : x \in X\} \subset X \times Y$.

Remark 16.4.5.

(a) Notice that $X \times Y$ is a Banach space equipped with the norm $\|(\,\cdot\,,\,\cdot\,)\|$.

(b) The operator T is closed if and only if G_T is closed.

The closed mapping theorem was proved by S. Banach in 1932.

Theorem 16.4.6 (The Closed Mapping Theorem). *The operator T is bounded if and only if it is closed.*

Proof. It remains to prove the sufficiency part, so suppose that T is closed. Then G_T is a Banach space since G_T is closed and $X \times Y$ is complete. The projection $P : G_T \to X$, that maps $(x, Tx) \in G_T$ onto $x \in X$, is obviously linear and bijective. The projection is also bounded:

$$\|P(x, Tx)\| = \|x\| \leq \|x\| + \|Tx\| = \|(x, Tx)\| \quad \text{for every} \quad (x, Tx) \in G_T.$$

It thus follows from the inverse mapping theorem that the inverse $P^{-1} : X \to G_T$ is bounded. Now suppose that $x_n \to x$. Since P^{-1} is continuous, we then have

$$P^{-1}(x_n) = (x_n, Tx_n) \to P^{-1}(x) = (x, Tx),$$

and, in particular, $Tx_n \to Tx$. ■

Example 16.4.7 (Continuation of Example 16.4.3). The operator T despite the fact that it is closed, the reason being that D is not a closed subset to $C[0, 1]$ and therefore not complete. □

Chapter 17

Linear Operators on Hilbert Spaces

In the following chapter, H denotes a Hilbert space over $\mathbf{C}$.

17.1. The Adjoint Operator

Theorem 17.1.1. *For every operator $T \in B(H)$ there exists a uniquely determined operator $T^* \in B(H)$ such that*

$$(Tx, y) = (x, T^*y) \quad \textit{for all } x, y \in H. \tag{1}$$

The operator T^* is called the **adjoint operator** or just the **adjoint** to T.

Proof. For a fixed vector $y \in H$, we define a linear operator f_y on H by

$$f_y(x) = (Tx, y), \quad x \in H.$$

Then $f_y \in H'$ with $\|f_y\| \le \|T\|\|y\|$. Thus, according to Theorem 11.6.1, there exists an element $T^*y \in H$ such that $f_y(x) = (x, T^*y)$ for every $x \in H$, which proves (1). The mapping $y \mapsto T^*y$ is linear since the inner product is linear. We also have

$$\|T^*y\| = \|f_y\| \le \|T\|\|y\|$$

for every $y \in H$, so T^* is bounded with $\|T^*\| \le \|T\|$. Finally, if

$$(Tx, y) = (x, T_1^*y) = (x, T_2^*y)$$

for all $x, y \in H$, then $(x, T_1^*y - T_2^*y) = 0$ for all $x, y \in H$, from which it follows that $T_1^*y = T_2^*y$ for every $y \in H$. ■

Remark 17.1.2. The proof shows that $\|T^*\| \le \|T\|$. In Theorem 17.2.1 (d) below, we will show that actually $\|T^*\| = \|T\|$.

Example 17.1.3. Suppose that $T : \mathbf{C}^d \to \mathbf{C}^d$ is given by the matrix A in the standard basis for $\mathbf{C}^d$. Let $A^* = \overline{A}^t$. Then

$$(Tx, y) = (Ax)^t\overline{y} = x^tA^t\overline{y} = x^t\overline{A^*y} = (x, T^*y)$$

for all $x, y \in H$ if T^* is the operator that is given by the matrix A^* in the standard basis for $\mathbf{C}^d$. Notice that we here use the uniqueness part of Theorem 11.6.1. □

Example 17.1.4. Let $K \in L^2((a, b) \times (a, b))$. In Example 10.4.4, we saw that the operator T, defined by

$$Tx(t) = \int_a^b K(t, \tau)x(\tau)\, d\tau, \quad a \le t \le b,$$

for $x \in L^2(a, b)$, maps $L^2(a, b)$ into $L^2(a, b)$. Now, according to Fubini's theorem,

$$\begin{aligned}(Tx, y) &= \int_a^b \left(\int_a^b K(t, \tau)x(\tau)\, d\tau\right)\overline{y(t)}\, dt = \int_a^b x(\tau)\overline{\left(\int_a^b \overline{K(t, \tau)}y(t)\, dt\right)}\, d\tau \\ &= (x, T^*y)\end{aligned}$$

for all $x, y \in L^2(a,b)$ if T^* is the operator defined by

$$T^*y(\tau) = \int_a^b \overline{K(t,\tau)}y(t)\,dt, \quad a \le \tau \le b. \qquad \square$$

Example 17.1.5. Let $R : \ell^2 \to \ell^2$ and $L : \ell^2 \to \ell^2$ be the right and left shift operators, respectively, defined in Example 10.3.4. Then

$$(Rx, y) = ((0, x_1, ...), (y_1, y_2, ...)) = ((x_1, x_2, ...), (y_2, y_3, ...)) = (x, Ly)$$

for all $x, y \in \ell^2$, which shows that $R^* = L$. In the same manner, one shows that $L^* = R$. $\square$

Example 17.1.6. The operator $T_m : L^2(0,1) \to L^2(0,1)$ is defined by

$$T_m x(t) = m(t)x(t), \quad 0 \le t \le 1,$$

for $x \in L^2(0,1)$, where the **multiplicator** $m \in L^\infty(0,1)$. In this case,

$$(T_m x, y) = \int_0^1 m(t)x(t)\overline{y(t)}\,dt = \int_0^1 x(t)\overline{\overline{m(t)}y(t)}\,dt = (x, T_{\overline{m}}y)$$

for all $x, y \in L^2(0,1)$, which shows that $(T_m)^* = T_{\overline{m}}$. $\square$

17.2. Properties of the Adjoint

Theorem 17.2.1. *Suppose that $S, T \in B(H)$. Then the following properties hold*:

(a) $(\alpha S + \beta T)^* = \overline{\alpha}S^* + \overline{\beta}T^*$ *for all* $\alpha, \beta \in \mathbf{C}$;

(b) $(ST)^* = T^*S^*$;

(c) $(T^*)^* = T$;

(d) $\|T^*\| = \|T\|$.

Proof.

(a) For $x, y \in H$, $((\alpha S + \beta T)x, y) = \alpha(Sx, y) + \beta(Tx, y) = (x, (\overline{\alpha}S^* + \overline{\beta}T^*)(y))$.

(b) For $x, y \in H$, $((ST)(x), y) = (S(Tx), y) = (Tx, S^*y) = (x, (T^*S^*)(y))$.

(c) Since $(Tx, y) = (x, T^*y)$ for all $x, y \in H$, $\overline{(Tx, y)} = \overline{(x, T^*y)}$, from which it follows that $(T^*y, x) = (y, Tx)$ for all $x, y \in H$. Hence, $(T^*)^* = T$.

(d) According to Remark 17.1.2, $\|T^*\| \le \|T\|$. If we apply this inequality to the operator $T = (T^*)^*$, we obtain $\|T\| = \|(T^*)^*\| \le \|T^*\|$ ∎

17.3. Self-adjoint, Normal, and Unitary Operators

Definition 17.3.1. An operator $T \in B(H)$ is called

(i) **self-adjoint** if $T^* = T$;

(ii) **normal** if $TT^* = T^*T$;

(iii) **unitary** if T is invertible and $T^{-1} = T^*$.

Notice that all self-adjoint and all unitary operators are normal.

Example 17.3.2. The operator in Example 17.1.3 is self-adjoint if $\overline{A}^t = A$. □

Example 17.3.3. The operator in Example 17.1.4 is self-adjoint if

$$\overline{K(\tau, t)} = K(t, \tau) \quad \text{for } 0 \le t, \tau \le 1.$$

For instance, if $K(t, \tau) = k(t - \tau)$, where k is an even, real-valued function, then the operator is self-adjoint. □

Example 17.3.4. The operator R in Example 17.1.5 is not normal:

$$RR^*x = (0, x_2, ...) \quad \text{but} \quad R^*Rx = (x_1, x_2, ...) \quad \text{for } x \in \ell^2.$$ □

Example 17.3.5. For the operator T_m in Exempel 17.1.6, $T_m^* = T_{\overline{m}}$, so T is self-adjoint if m is real-valued. We also see that T_m is normal since

$$T_m T_m^* = T_{|m|^2} \quad \text{and} \quad T_m^* T_m = T_{|m|^2}.$$

From these identities, it also follows that T_m is unitary if and only if $|m| = 1$, in which case $T_m^{-1} = T_{m^{-1}}$. □

Lemma 17.3.6. *If $T \in B(H)$ is self-adjoint, then (Tx, x) is real for every $x \in H$.*

Proof. If $x \in H$, then

$$(Tx, x) = \overline{(x, Tx)} = \overline{(T^*x, x)} = \overline{(Tx, x)}.$$ ■

Theorem 17.3.7 (Rayleigh's Principle). *If $T \in B(H)$ is self-adjoint, then*

$$\|T\| = \sup_{\|x\|=1} |(Tx, x)|. \tag{2}$$

Proof. Denote the right-hand side (2) by M. Since

$$|(Tx, x)| \le \|Tx\|\|x\| \le \|T\|\|x\|^2 = \|T\|$$

for every $x \in H$ that satisfies $\|x\| = 1$, we have $M \le \|T\|$. To prove the reverse inequality, suppose that $\|x\| = \|y\| = 1$. Then, using the fact that $T = T^*$, we see that

$$(T(x + y), x + y) - (T(x - y), (x - y)) = 2((Tx, y) + (Ty, x)) = 4\,\mathrm{Re}(Tx, y).$$

As above, $|(Tu, u)| \le M\|u\|^2$ for every vector $u \in H$. This then implies that

$$4\,\mathrm{Re}(Tx, y) \le M(\|x + y\|^2 + \|x - y\|^2) = 2M(\|x\|^2 + \|y\|^2) = 4M.$$

Choose $\theta \in \mathbf{R}$ so that $(Tx, y) = e^{i\theta}|(Tx, y)|$. If we replace x with $e^{-i\theta}x$ in the last inequality, we then obtain that $|(Tx, y)| \le M$ for every unit vector y. Corollary 12.4.1 and Theorem 11.6.1 now shows that $\|Tx\| \le M$ for every unit vector x, and hence that $\|T\| \le M$. ■

Chapter 18

Compact Operators on Hilbert Spaces

In the following chapter, H denotes a Hilbert space over $\mathbf{C}$.

18.1. Compact Operators

Definition 18.1.1. An operator $T \in L(H)$ is said to be **compact** if $T(B)$ is relatively compact for every bounded subset B to K. The class of compact operators on H is denoted $K(H)$.

According to Definition 5.3.1, $T(B)$ is relatively compact if $\overline{T(B)}$ is compact.

Theorem 18.1.2. *Every compact operator is bounded.*

Proof. Suppose that $T \in K(H)$. Since B_H is bounded and T is compact, $\overline{T(B_H)}$ is compact. It then follows from Theorem 5.1.6 that $\overline{T(B_H)}$ is bounded and, in particular, that $T(B_H)$ is bounded. This fact now implies that

$$\|T\| = \sup_{\|x\| \leq 1} \|Tx\| < \infty. \qquad \blacksquare$$

Example 18.1.3. Suppose that $\dim(H) < \infty$ and that $T \in B(H)$. If B is a bounded subset to H, then $\overline{T(B)}$ is closed and bounded. Since H is finite-dimensional, this implies that $\overline{T(B)}$ is compact. Thus, T is compact. □

Example 18.1.4. Suppose that $T \in B(H)$ has **finite rank**: $\dim R(T) < \infty$. If B is a bounded subset to H, then $T(B)$ is bounded since T is bounded. It now follows that $\overline{T(B)}$ is compact since $R(T)$ is finite-dimensional. □

Example 18.1.5. If $\dim(H) = \infty$, then the identity operator $I : H \to H$ is never compact because $\overline{I(B_H)} = B_H$ and B_H is not compact according to Riesz' lemma (cf. Theorem 8.6.1 and Corollary 8.6.2). □

The following alternative characterization of compact operators follows from Theorem 5.3.3.

Theorem 18.1.6. *An operator $T \in B(H)$ is compact if and only if $(x_n)_{n=1}^{\infty} \subset H$ is bounded implies that $(Tx_n)_{n=1}^{\infty} \subset H$ has a convergent subsequence.*

18.2. Properties of Compact Operators

The following theorem summarizes the algebraic properties of $K(H)$; it shows that $K(H)$ is a subspace to $B(H)$ and an algebra $\mathbf{C}$. The theorem is a direct consequence of Theorem 18.1.6.

Theorem 18.2.1. *Suppose that $S, T \in K(H)$. Then the following properties hold:*

(a) $\alpha S + \beta T \in K(H)$ *for all* $\alpha, \beta \in \mathbf{C}$;

(b) $ST \in K(H)$.

The next theorem shows that $K(H)$ is a closed subspace to $B(H)$.

Theorem 18.2.2. *If* $T_n \in K(H)$ *and* $T_n \to T \in B(H)$*, then* $T \in K(H)$.

Proof. Let $(x_n^{(0)})_{n=1}^\infty \subset H$ be a bounded sequence; we will construct a convergent subsequence to $(Tx_n^{(0)})_{n=1}^\infty \subset H$ using **Cantor's diagonal procedure**. First choose a subsequence $(x_n^{(1)})_{n=1}^\infty \subset (x_n^{(0)})_{n=1}^\infty$ such that $(T_1 x_n^{(1)})_{n=1}^\infty$ is convergent. Then choose a subsequence $(x_n^{(2)})_{n=1}^\infty \subset (x_n^{(1)})_{n=1}^\infty$ such that $(T_2 x_n^{(2)})_1^\infty$ is convergent. Continue in the same way and let $x_n = x_n^{(n)}$, $n = 1, 2, \dots$, the diagonal sequence.

Notice that

$$\begin{aligned}\|T(x_m - x_n)\| &= \|(T - T_k)(x_m - x_n) + T_k(x_m - x_n)\| \\ &\le \|T - T_k\|\|x_m - x_n\| + \|T_k(x_m - x_n)\|.\end{aligned}$$

The right-hand side tends to 0 as $k \to \infty$ since $T_k \to T$ and $\|x_m - x_n\|$ is bounded. Also, for a fixed k, $\|T_k(x_m - x_n)\| \to 0$ as $m, n \to \infty$. This shows that $(Tx_n)_{n=1}^\infty$ is a Caucy sequence and therefore convergent. ■

Corollary 18.2.3. *The space* $K(H)$ *is complete.*

18.3. Approximation with Finite Rank Operators

Theorem 18.2.2 in combination with 18.1.4 gives a method for proving that a operator is compact. Here, the **rank** of an operator is the dimension of its range.

Example 18.3.1. Consider the operators T and T_n, $n = 1, 2, \dots$, from ℓ^2 to ℓ^2, defined by

$$Tx = \left(\frac{x^1}{1}, \frac{x^2}{2}, \dots\right) \quad \text{and} \quad T_n x = \left(\frac{x^1}{1}, \frac{x^2}{2}, \dots, \frac{x^n}{n}, 0, \dots\right) \quad \text{for } x \in \ell^2.$$

Then every operator T_n has finite rank. Also, for $x \in \ell^2$,

$$\|(T - T_n)x\|_2^2 = \sum_{j=n+1}^\infty \frac{|x^j|^2}{j^2} \le \frac{1}{(n+1)^2} \sum_{j=1}^\infty |x_j|^2 \le \frac{1}{(n+1)^2} \|x\|_2,$$

which shows that $\|T - T_n\| \le 1/(n+1)$, and hence that $T_n \to T$. It now follows from Theorem 18.2.2 that T is compact. □

Exactly as in this example, one can show that every **diagonal operator**

$$Tx = (a^1 x^1, a^2 x^2, \dots), \quad x \in \ell^2,$$

where $(a^j)_{j=1}^\infty \in \mathbf{c}_0$, is compact. In general, if $T \in B(H)$ is the limit of a sequence of operators T_n with finite rank, then T is compact. Theorem 18.3.5 shows that the converse is also true.

Example 18.3.2. The operator $T : L^2(0,1) \to L^2(0,1)$ is defined by

$$Tx(t) = \int_0^1 K(t,\tau)x(\tau)\,d\tau,\ 0 \le \tau \le 1, \quad \text{for } x \in L^2(0,1),$$

where the kernel K is assumed to belong to $L^2((0,1)\times(0,1))$. According to Example 10.4.4, $\|T\| \le \|K\|_2$. An orthonomal basis for $L^2(0,1)$ is given by the functions $e_n(t) = e^{2\pi int}$, $0 \le t \le 1$, $n = 0, \pm1, \dots$. If we for a fixed τ expand $K(t,\tau)$ with respect to this basis, we obtain

$$K(t,\tau) = \sum_{-\infty}^{\infty} K_n(\tau)e_n(t), \quad 0 \le t \le 1,$$

where the series converges at least in $L^2(0,1)$. Parseval's identity also shows that

$$\|K\|_2^2 = \int_0^1 \left(\int_0^1 |K(t,\tau)|^2\,dt\right) d\tau = \int_0^1 \left(\sum_{-\infty}^{\infty} |K_n(\tau)|^2\right) d\tau = \sum_{-\infty}^{\infty} \int_0^1 |K_n(\tau)|^2\,d\tau.$$

For $N = 0, 1, \dots$, we now define $K_N(t,\tau) = \sum_{-N}^{N} K_n(\tau)e_n(t)$ and let T_N be the operator, which is given by the kernel K_N. Since $\dim R(T_N) = 2N+1$, the operator T_N is compact. We now obtain

$$\|T - T_N\|^2 \le \|K - K_N\|_2^2 \le \sum_{|n|>N} \int_0^1 |K_n(\tau)|^2\,d\tau \longrightarrow 0 \quad \text{as } N \to \infty,$$

which shows that T is compact. □

To prove that every compact operator is the limit of a sequence of finite rank operators, we first show that the range of a compact operator is separable.

Theorem 18.3.3. *If $T \in K(H)$, then the range $R(T)$ is separable.*

In the proof, we use the following lemma (see Definition 5.2.1).

Lemma 18.3.4. *Every totally bounded metric space X is separable.*

Proof. There exist elements $x_1^{(n)}, \dots, x_{k_n}^{(n)} \in X$ such that $X \subset \bigcup_{j=1}^{k_n} B_{1/n}(x_j^{(n)})$ for $n = 1, 2, \dots$. This shows that the set $\{x_j^{(n)} : j = 1, \dots, k_n,\ n = 1, 2, \dots\}$ is dense in X. ■

Proof (Theorem 18.3.3). Notice that $R(T) = \bigcup_{n=1}^{\infty} T(B_n(0))$. It also follows from Lemma 18.3.4 that every set $T(B_n(0))$ contains a countable, dense subset D_n. But then $\bigcup_{n=1}^{\infty} D_n$ is countable and dense in $R(T)$. ■

Theorem 18.3.5. *If $T \in K(H)$, then there exits a sequence T_n, $n = 1, 2, \dots$, with finite rank such that $T_n \to T$.*

Proof. According to Theorem 18.3.3, $H_1 = \overline{R(T)}$ is a separable Hilbert space. Let $(e_n)_{n=1}^{\infty}$ be an orthonormal basis for H_1 (see Theorem 9.5.11), let P_n denote the orthogonal projection on $\operatorname{span}\{e_1, \dots, e_n\}$, and define $T_n = P_n T$, $n = 1, 2, \dots$. For $x \in H$, Parseval's identity (Theorem 9.5.9) shows that

$$\|Tx - T_n x\|^2 = \left\| \sum_{j=1}^{\infty} (Tx, e_j) e_j - \sum_{j=1}^{n} (Tx, e_j) e_j \right\|^2 = \sum_{j=n+1}^{\infty} |(Tx, e_j)|^2 \longrightarrow 0$$

as $n \to \infty$. Let $\varepsilon > 0$. Since $\overline{T(B_H)}$ is compact, there exists elements $x_1, \dots, x_k \in H$ such that $T(B_H) \subset \bigcup_{j=1}^{k} B_\varepsilon(Tx_j)$. Then choose N so large that $\|Tx_j - T_n x_j\| < \varepsilon$ for every j if $n \geq N$. If $x \in B_H$ and $\|Tx - Tx_j\| < \varepsilon$, we now obtain

$$\|Tx - T_n x\| \leq \|Tx - Tx_j\| + \|Tx_j - T_n x_j\| + \|P_n(Tx_j - Tx)\| < 3\varepsilon.$$

This holds for every $x \in B_H$, so it follows that $\|T - T_n\| \leq 3\varepsilon$ if $n \geq N$. ■

Corollary 18.3.6. *If $T \in K(H)$, then also $T^* \in K(H)$.*

Proof. Since T is compact, there exist finite rank operators T_n, $n = 1, 2, \dots$, such that $T_n \to T$. Obviously, every operator T_n^* has finite rank. It also follows from Theorem 17.2.1 that $\|T^* - T_n^*\| = \|T - T_n\|$, from which it follows that $T_n^* \to T^*$, and hence that T^* is compact since every operator T_n is compact. ■

Chapter 19

Spectral Theory for Selfadjoint, Compact Operators on Hilbert Spaces

In this chapter, H denotes a complex Hilbert space.

19.1. Eigenvalues and Eigenvectors

Definition 19.1.1. Let $T \in B(H)$.

(a) A number $\lambda \in \mathbf{C}$ is said to be an **eigenvalue** of T if $\ker(T - \lambda I) \neq \{0\}$.

(b) The set of eigenvalues of T is called the **point spectrum** of T and is denoted $\sigma_p(T)$.

(c) If $\lambda \in \sigma_p(T)$, the space $\ker(T - \lambda I)$ is called the **eigenspace** of T and the non-zero elements in $\ker(T - \lambda I)$ are the **eigenvectors** corresponding to λ.

Remark 19.1.2.

(a) If $\lambda \in \sigma_p(T)$ and $x \in \ker(T - \lambda I)$, then $Tx = \lambda x$.

(b) Suppose that $\lambda \in \sigma_p(T)$ and that $x \in \ker(T - \lambda I)$ satisfies $\|x\| = 1$. Since T is bounded, it follows that $|\lambda| = \|\lambda x\| = \|Tx\| \leq \|T\|$. This shows that all the eigenvalues of T belong to the disk $\{z \in \mathbf{C} : |z| \leq \|T\|\}$.

(c) As we will see in Example 19.1.3 (a) and (c), the point spectrum $\sigma_p(T)$ of a operator T may be empty.

Example 19.1.3.

(a) The operator $T : \ell^2 \to \ell^2$, defined by $Tx = (0, x^1, x^2, ...)$ for $x \in \ell^2$, has no eigenvalues. Indeed, suppose that $Tx = \lambda x$, i.e.,

$$(0, x^1, x^2, ...) = \lambda(x^1, x^2, x^3, ...).$$

If $\lambda = 0$, then it follows that $x = 0$, so $\ker(T) = \{0\}$. If $\lambda \neq 0$, it follows that $x^1 = 0$ and $x^{j+1} = x^j/\lambda$ for $j = 1, 2, \dots$, which again shows that $x = 0$, and hence that $\ker(T - \lambda I) = \{0\}$.

(b) Consider the operator $T : \ell^2 \to \ell^2$, defined by $Tx = (\frac{x_1}{1}, \frac{x_2}{2}, ...)$ for $x \in \ell^2$. Suppose that $Tx = \lambda x$, i.e.,

$$(\tfrac{x_1}{1}, \tfrac{x_2}{2}, ...) = \lambda(x^1, x^2, ...).$$

If $\lambda = 0$, it follows that $x = 0$. If $\lambda \neq 0$, then either $x = 0$ or $\lambda = \frac{1}{j}$ and x^j is arbitrary for some value of j and $x^j = 0$ for all other values of j. Thus, the eigenvalues of T are $\lambda = \frac{1}{j},\ j = 1, 2, \dots$, with the corresponding eigenvectors $te_j,\ t \neq 0$, where $e_1, e_2, \dots$, is the standard basis of ℓ^2.

(c) Let $T : C[0,1] \to C[0,1]$ be defined by $Tx(t) = \int_0^t x(\tau)\,d\tau,\ 0 \le t \le 1$, for $x \in C[0,1]$; here, we consider $C[0,1]$ as a subspace to $L^2(0,1)$. Suppose that $Tx = \lambda x$, i.e.,

$$\int_0^t x(\tau)\,d\tau = \lambda x(t) \quad \text{for } 0 \le t \le 1. \tag{1}$$

After differentiating, we obtain the differential equation $x = \lambda x',\ 0 \le t \le 1$. If $\lambda = 0$, this shows that $x = 0$. If $\lambda \ne 0$, it follows from (1) that every solution satisfies the initial condition $x(0) = 0$. It is easy to see that the only solution to this problem is $x = 0$. Thus, T has no eigenvalues. □

19.2. Spectrum for Normal, Self-adjoint, and Unitary Operators

Theorem 19.2.1. *An operator $T \in B(H)$ is normal if and only if $\|Tx\| = \|T^*x\|$ for every $x \in H$.*

Proof. The necessity part follows directly from the definition of an normal operator: For $x \in H$,

$$\|Tx\|^2 = (Tx, Tx) = (x, T^*Tx) = (x, TT^*x) = (T^*x, T^*x) = \|T^*x\|^2.$$

We leave the proof of the sufficiency part as an exercise to the reader. ■

Theorem 19.2.2. *Suppose that $T \in B(H)$. Then the following properties hold:*

(a) *If T is normal, then* $\ker(T - \lambda I) = \ker(T^* - \overline{\lambda} I)$. *In particular, there holds $\sigma_p(T^*) = \overline{\sigma_p(T)}$.*

(b) *If T is self-adjoint, then $\sigma_p(T) \subset \mathbf{R}$.*

(c) *If T is unitary, then $\sigma_p(T) \subset \{z \in \mathbf{C} : |z| = 1\}$.*

Proof.

(a) It follows from Theorem 19.2.1 and Theorem 17.2.1 (a) that

$$\|(T - \lambda I)x\| = \|(T - \lambda I)^*x\| = \|(T^* - \overline{\lambda} I)x\|$$

for every $\lambda \in \mathbf{C}$ and every $x \in H$, which proves the assertion.

(b) If $x \ne 0$ and $x \in \ker(T - \lambda I)$, then $\lambda x = Tx = T^*x = \overline{\lambda} x$, which shows that $\lambda = \overline{\lambda}$, i.e., $\lambda \in \mathbf{R}$.

(c) If $x \ne 0$ and $x \in \ker(T - \lambda I)$, then $\|x\|^2 = (T^*Tx, x) = \|Tx\|^2 = |\lambda|^2\|x\|^2$, which shows that $|\lambda| = 1$. ■

Corollary 19.2.3. *If $T \in B(H)$ is normal and $\lambda, \mu \in \sigma_p(T)$, where $\lambda \ne \mu$, then* $\ker(T - \lambda I) \perp \ker(T - \mu I)$.

Proof. Suppose that $x \in \ker(T - \lambda I)$ and $y \in \ker(T - \mu I)$. Then

$$\lambda(x, y) = (Tx, y) = (x, T^*y) = (x, \overline{\mu} y) = \mu(x, y).$$

This implies that $(x, y) = 0$ since $\lambda \ne \mu$. ■

19.3. The Spectrum of Compact Operators

Theorem 19.3.1. *Suppose that* $T \in K(H)$ *and* $\lambda \in \sigma_p(T) \setminus \{0\}$*. Then the eigenspace* $\ker(T - \lambda I)$ *is finite-dimensional.*

Proof. To obtain a contradiction, we assume that $\dim(\ker(T - \lambda I)) = \infty$ and choose an orthonormal sequence $(e_n)_{n=1}^{\infty}$ in $\ker(T - \lambda I)$. If $m \neq n$, then

$$\|Te_m - Te_n\|^2 = \|\lambda(e_m - e_n)\|^2 = |\lambda|^2\|e_m - e_n\|^2 = 2|\lambda|^2 > 0,$$

which shows that $(Te_n)_{n=1}^{\infty}$ does not contain a convergent despite the fact that T is compact and $(e_n)_{n=1}^{\infty}$ is bounded. It follows that $\ker(T - \lambda I)$ has to be finite-dimensional. ∎

Theorem 19.3.2. *Suppose that* $T \in K(H)$*. Then* $\sigma_p(T)$ *is countable with* 0 *as the only possible accumulation point.*

Proof. Suppose that $\sigma_p(T)$ is infinite and let $\lambda_1, \lambda_2, \ldots \in \sigma_p(T) \setminus \{0\}$ with corresponding eigenvectors $e_1, e_2, \ldots$. Put $H_n = \operatorname{span}\{e_1, e_2, \ldots e_n\}$, $n = 1, 2, \ldots$. According to Riesz' lemma (Theorem 8.6.1), there exist vectors $x_n \in S_{H_n}$ such that $\operatorname{dist}(x_n, H_{n-1}) \geq \frac{1}{2}$ for $n = 2, 3, \ldots$. Since $x_n = \alpha_n e_n + y_n$, where $y_n \in X_{n-1}$, it follows that

$$(T - \lambda_n I)x_n = (T - \lambda_n I)y_n \in X_{n-1}.$$

Hence, if $m > n$, then

$$\left\|T\Big(\frac{x_m}{\lambda_m}\Big) - T\Big(\frac{x_n}{\lambda_n}\Big)\right\| = \left\|x_m + (T - \lambda_m)\Big(\frac{x_m}{\lambda_m}\Big) - T\Big(\frac{x_n}{\lambda_n}\Big)\right\| \geq \frac{1}{2}.$$

It follows that the sequence $(T(\lambda_n^{-1}x_n))_{n=2}^{\infty}$ has no convergent subsequences, so the sequence $(\lambda_n^{-1}x_n)_{n=2}^{\infty}$ does not contain any bounded subsequences. This in turn implies that $|\lambda_n|^{-1} = |\lambda_n|^{-1}\|x_n\| \to \infty$, i.e., $\lambda_n \to 0$. The only accumulation point of $\sigma_p(T)$ is thus 0. This then implies that the set $\{\lambda \in \sigma_p(T) : \frac{1}{m} \leq |\lambda| \leq \|T\|\}$ is finite for $m = 1, 2, \ldots$, so $\sigma_p(T)$ is countable. ∎

Theorem 19.3.3. *Let* $T \in K(H)$ *be a self-adjoint operator such that* $T \neq 0$*. Then either* $\|T\| \in \sigma_p(T)$ *or* $-\|T\| \in \sigma_p(T)$*. In particular,* $\sigma_p(T) \neq \varnothing$*.*

Proof. We will use Rayleigh's principle (Theorem 17.3.7). First choose a sequence $(x_n)_{n=1}^{\infty} \in B_H$ such that $|(Tx_n, x_n)| \to \|T\|$. By passing to a subsequence, we may assume that $(Tx_n, x_n) \to \lambda$, where $|\lambda| = \|T\|$. Since $T^* = T$ and (Tx_n, x_n) is real for every n (Lemma 17.3.6), we have

$$\begin{aligned}\|(T - \lambda I)x_n\|^2 &= \|Tx_n\|^2 + \lambda^2 - 2\lambda(Tx_n, x_n) \leq \|T\|^2 + \lambda^2 - 2\lambda(Tx_n, x_n)\\ &= 2\lambda^2 - 2\lambda(Tx_n, x_n),\end{aligned}$$

that is $\|(T - \lambda I)x_n\| \to 0$. Again passing to a subsequence, we may assume that $Tx_n \to y \in H$. It now follows that

$$\lambda x_n = (\lambda I - T)x_n + Tx_n \to y.$$

Since $|\lambda| = \|T\| \neq 0$, we have $y \neq 0$. Finally,

$$\lambda y = \lambda \lim_{n\to\infty} Tx_n = \lim_{n\to\infty} T(\lambda x_n) = Ty,$$

which shows that $y \in \ker(T - \lambda I)$ and that $\lambda \in \sigma_p(T)$. ∎

19.4. A Spectral Theorem for Compact, Self-adjoint Operators

Suppose that $T \in K(H)$ is self-adjoint with eigenvalues $\lambda_1, \lambda_2, \dots$, ordered so that $|\lambda_1| \geq |\lambda_2| \geq \dots$. Let P_j be the orthogonal projection on $E_j = \ker(T - \lambda_j I)$ for $j = 1, 2, \dots$.

Theorem 19.4.1. *With the assumptions and notation above,* $T = \sum_j \lambda_j P_j$.

Proof. We will prove the theorem in the case when $\sigma_p(T)$ is infinite and leave the remaining case as an exercise to the reader. For $n = 1, 2, \dots$, let $Y_n = E_1 \oplus \dots \oplus E_n$. If $T_n = T|_{Y_n^\perp}$, then $T_n : Y_n^\perp \to Y_n^\perp$; indeed, if $x \in Y_n^\perp$ and $y \in E_j$, where $1 \leq j \leq n$, then

$$(T_n x, y) = (Tx, y) = (x, T^* y) = (x, \overline{\lambda_j} y) = \lambda_j (x, y) = 0.$$

Since T is self-adjoint and compact, the same properties hold for T_n. For $x \in H$, we let $x_n = \sum_{j=1}^n P_j x$ denote the orthogonal projection of x on Y_n. Then

$$\left\| Tx - \sum_{j=1}^n \lambda_j P_j x \right\| = \|T(x - x_n)\| = \|T_n(x - x_n)\| \leq \|T_n\| \|x - x_n\| \leq |\lambda_{n+1}| \|x\|.$$

Hence, $\|T - \sum_{j=1}^n \lambda_j P_j\| \leq |\lambda_{n+1}| \to 0$, so $T = \sum_{j=1}^\infty \lambda_j P_j$. ∎

Corollary 19.4.2. *If* $x \in H$, *then* $x = y + \sum_j P_j x$, *where* $y \in \ker(T)$.

Proof. If y is defined by $y = x - \sum_j P_j x$, then

$$Ty = Tx - \sum_j P_j x = Tx - \sum_j \lambda_j P_j x = 0,$$

which shows that $y \in \ker(T)$. ∎

It follows from the corollary that if $\ker(T)$ is separable, then there exists an orthonormal basis for H consisting of eigenvectors of T.