CMMS05 Metric and Banach Spaces

Lecturer: Alexander Pushnitski

Autumn 2012

Based on lecture notes by Yuri Safarov
Tuesdays 14:00-16:00, room 436

There are two types of exercises in these notes: the ones in the body of the text and the ones at the ends of sections. Exercises of the first type form an essential part of the exposition; they fill gaps in the text. Attempting (and succeeding!) exercises of both types is a crucial part of the course. Difficult exercises are marked with a star.

1 Metrics, norms and inner products

1.1 Definitions

Definition 1.1. Let X be a non-empty set. A function $\rho : X \times X \to \mathbb{R}$ is called a *metric* on X if it satisfies

(i) $\rho(x, y) > 0$ if $x \neq y$ and $\rho(x, x) = 0$,

(ii) $\rho(x, y) = \rho(y, x)$,

(iii) $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$ (this is called the *triangle inequality*),

where x, y and z are arbitrary elements of X.

The pair (X, ρ) is said to be a metric space. If $X_0 \subset X$ then ρ is also a metric on X_0. The metric space (X_0, ρ) is called a subspace of (X, ρ). The function $\rho(x, y)$ can (and should) be interpreted as the distance between the elements x and y in the space (X, ρ).

If the X is a linear space, it is often possible to express the metric ρ in terms of a function of one variable that can be thought of as the length of each element (i.e., the distance of this element from 0). In this course, we will be dealing with *complex* linear spaces, i.e. linear spaces over the field $\mathbb{C}$. In an entirely similar way, a theory of *real* linear spaces can be developed.

Definition 1.2. Let X be a linear space. A function $\| \cdot \| : X \to [0, \infty)$ is called a *norm* on X if it satisfies

(i) $\|x\| > 0$ if $x \neq 0$ and $\|0\| = 0$,

(ii) $\|\lambda x\| = |\lambda| \, \|x\|$ for all $x \in X$ and $\lambda \in \mathbb{C}$,

(iii) $\|x+y\| \leq \|x\|+\|y\|$ for all $x, y \in X$ (this is a particular case of the triangle inequality).

A linear space provided with a norm is called a *normed* space.

Exercise 1.3. Let X be a linear space with the norm $\|\cdot\|$. Prove that $\rho(x, y) = \|x - y\|$ is a metric on X.

However, not every metric arises in this way (see Example 1.13 below).

Definition 1.4. Let X be a linear space. A function $(\cdot, \cdot) : X \times X \to \mathbb{C}$ is called an *inner product* on X if it satisfies

(i) $(x, x) \geq 0$ for all $x \in X$ and $(x, x) = 0$ if and only if $x = 0$;

(ii) $(x, y) = \overline{(y, x)}$ for all $x, y \in X$, where the bar denotes the complex conjugation;

(iii) $(\alpha x_1 + \beta x_2, y) = \alpha(x_1, y) + \beta(x_2, y)$ for all $x_1, x_2, y \in X$ and all $\alpha, \beta \in \mathbb{C}$.

The inner product generates the norm $\|x\| := \sqrt{(x, x)}$. Therefore an inner product space is a normed space (and is therefore a metric space). The proof is outlined in the following exercises.

Exercise 1.5. Prove that the above defined norm satisfies the axioms (i), (ii) of the norm.

In order to prove the triangle inequality for the norm, one first needs to prove the Cauchy-Schwartz inequality for the inner product:

$$|(x, y)|^2 \leqslant (x, x)(y, y), \qquad \forall x, y \in X.$$

Exercise 1.6. Prove the Cauchy-Schwartz inequality. Proceed as follows. Consider the quadratic polynomial $f(t) = (x + ty, x + ty)$ of $t \in \mathbb{R}$. Observe that $f(t) \geqslant 0$ for all t. State this property in terms of the discriminant of f. Deduce the inequality

$$(\mathrm{Re}\,(x, y))^2 \leqslant (x, x)(y, y), \qquad \forall x, y \in X.$$

Now replace x in the above inequality by αx, where α is a complex number with modulus one. Choose α such that the left hand side of the above inequality becomes $|(x, y)|^2$ and thus complete the proof.

Exercise 1.7. Using the Cauchy-Schwartz inequality, prove that $\|x\| := \sqrt{(x, x)}$ satisfies the triangle inequality.

1.2 Elementary examples

Example 1.8. $\mathbb{R}$ and $\mathbb{C}$ are metric spaces with respect to the usual distance $\rho(x,y) = |x-y|$. One can consider more exotic metrics; for example, $\rho(x,y) = |\tan^{-1}(x) - \tan^{-1}(y)|$ on $\mathbb{R}$.

Example 1.9. Let X be the set of all words in the English language. Elements of X are considered as sequences of letters of the alphabet, appended by an infinite sequence of spaces. For $x, y \in X$, let $\rho(x,y)$ be the number of positions where the words x and y have distinct letters. For example, if $x =$"spin" and $y =$"spans" then $\rho(x,y) = 2$. It is easy to see that ρ is a metric on X. Metrics of this type are considered in information theory.

Example 1.10. $\mathbb{C}^n$ is an inner product space, and therefore is a normed space and a metric space. The standard (Euclidean) inner product, norm and metric are defined by

$$(x,y) = \sum_{i=1}^{n} x_i\,\bar{y}_i, \qquad \|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}, \qquad \rho(x,y) = \left(\sum_{i=1}^{n} |x_i - y_i|^2\right)^{1/2},$$

where x_i and y_i are the coordinates of x and y. The reason for the index 2 in the notation for the norm will soon become clear.

Example 1.11. $\mathbb{C}^n$ can also be equipped with the norms

$$\|x\|_1 = \sum_{i=1}^{n} |x_i| \quad \text{and} \quad \|x\|_\infty = \max_{i=1\ldots n} |x_i|.$$

These norms generate metrics on $\mathbb{C}^n$ but are not associated with inner products.

Exercise 1.12. Check that the axioms of the norm for $\|\cdot\|_1$ and $\|\cdot\|_\infty$ hold true.

Example 1.13. (discrete metric) For any set X, define the metric ρ by

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

The discrete metric is not generated by a norm even if X is a linear space. Indeed, the function $\rho(x,0)$ on X does not satisfy conditions of Definition 1.2.

Example 1.14. Fix $p > 1$. The linear space $\mathbb{C}^n$ can be equipped with the norm

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}.$$

The axioms (i), (ii) of the norm are straightforward to check. The triangle inequality (which in this case is called *Minkowski's inequality*) is less obvious; in the following exercises, we outline its proof. This proof depends on *Hölder's inequality:*

$$\sum_{i=1}^{n} |x_i y_i| \leqslant \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p} \left(\sum_{i=1}^{n} |y_i|^q\right)^{1/q} = \|x\|_p \|y\|_q,$$

where $p > 1$ and $q > 1$ are related by the identity $\frac{1}{p} + \frac{1}{q} = 1$. Note that for $p = 2$ one has $q = 2$ and Hölder's inequality reduces to Cauchy-Schwartz inequality. Also, formally we can take $p = 1$, $q = \infty$, in which case Hölder's inequality is still true and the proof is straightforward.

Exercise 1.15. 1. Let $p \geqslant 1$. Consider the function $f(t) = t - (t^p/p)$, $t \geqslant 0$. Prove that $f(t) \leqslant f(1)$ for all $t \geqslant 0$.

2. Substitute $t = |x_i|/|y_i|^{q-1}$; rearrange to obtain

$$|x_i y_i| \leqslant \frac{1}{p}|x_i|^p + \frac{1}{q}|y_i|^q.$$

(Hint: remember to use $\frac{1}{p} + \frac{1}{q} = 1$).

3. Take a sum over i and use an additional parameter $a > 0$ to obtain

$$\sum_{i=1}^{n} |x_i y_i| \leqslant \frac{1}{p} a^p \|x\|_p^p + \frac{1}{q} a^{-q} \|y\|_q^q.$$

4. Choose an appropriate value of a to obtain the Hölder inequality. (Hint: minimize the r.h.s. in $a > 0$).

Exercise 1.16. 1. Using the elementary triangle inequality for the modulus of a compex number, obtain

$$|x_i + y_i|^p \leqslant |x_i + y_i|^{p-1}|x_i| + |x_i + y_i|^{p-1}|y_i|.$$

2. Summing over i and using the Hölder inequality twice in the r.h.s., obtain

$$\|x + y\|_p^p \leqslant \|x + y\|_p^{p-1}\|x\|_p + \|x + y\|_p^{p-1}\|y\|_p.$$

From here deduce Minkowski's inequality. (Hint: remember to use $\frac{1}{p} + \frac{1}{q} = 1$).

1.3 Balls and bounded sets

Let (X, ρ) be a metric space and r be a strictly positive number.

Definition 1.17. If $\alpha \in X$ then the set $B_r(\alpha) = \{x \in X : \rho(x, \alpha) < r\}$ is called the *open ball*, and the set $B_r[\alpha] = \{x \in X : \rho(x, \alpha) \leq r\}$ is called the *closed ball* with the centre α and radius r.

If there is a need to emphasize the metric, we write $B_r^\rho(\alpha)$ and $B_r^\rho[\alpha]$. Clearly,

$$\alpha \in B_{r-\varepsilon}(\alpha) \subset B_r(\alpha) \subset B_r[\alpha] \subset B_{r+\varepsilon}(\alpha)\,, \qquad \forall r > \varepsilon > 0\,.$$

Definition 1.18. A subset K of a metric space (X, ρ) is called *bounded* if, for some $x \in X$ and $r > 0$, we have $K \subset B_r(x)$. A metric space (X, ρ) is called *bounded* if X is a bounded set.

Definition 1.19. Let $A \subset X$ be a subset of a metric space (X, ρ). The *diameter* of A is

$$\operatorname{diam} A = \sup\{\rho(x, y) \mid x, y \in A\}.$$

Exercise 1.20. Prove that A is bounded if and only if $\operatorname{diam} A < \infty$.

1.4 Isometry and equivalent metrics

Definition 1.21. Let X_1 and X_2 be metric spaces with the metrics ρ_1 and ρ_2 respectively. Let $F : X_1 \to X_2$ be a bijection (i.e. a one-to-one mapping such that the range of F is the whole of X_2). Suppose that $\rho_2(F(a), F(b)) = \rho_1(a, b)$ for all $a, b \in X_1$. Then we say that F is an *isometry* between X_1 and X_2. The metric spaces X_1 and X_2 are said to be *isometric*.

Example 1.22. Let $A_1 = [0, 1]$ and $A_2 = [3, 4]$ with the usual metric. Let $g : A_1 \to A_2$ be the mapping $g(x) = 3 + x$. Clearly, g is an isometry.

Isometric spaces differ only in the explicit nature of their elements; one can say that from the point of view of the metric spaces theory, such spaces are *identical*. If X_1 and X_2 are isometric, then any property that can be stated entirely in terms of the metric will be true or false simultaneously for X_1 and X_2. See Exercise 1.32.

Definition 1.23. We say that the metrics ρ_1 and ρ_2 defined on the same set X are equivalent if

$$c\rho_1(x, y) \leqslant \rho_2(x, y) \leqslant C\rho_1(x, y), \quad \forall x, y \in X$$

with some c, C independent of x, y. Two norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on the same linear space X are said to be equivalent when

$$c\|x\|_1 \leqslant \|x\|_2 \leqslant C\|x\|_1 \quad \forall x$$

with some c, C independent of x.

Example 1.24. The norms $\|\cdot\|_p$ on $\mathbb{C}^n$ are equivalent to each other for all $1 \leqslant p \leqslant \infty$. This follows from the inequality

$$\|x\|_\infty \leqslant \|x\|_p \leqslant n^{1/p}\|x\|_\infty.$$

Exercise 1.25. Prove this inequality.

Exercises:

Exercise 1.26. In $\mathbb{R}^2$, plot the unit ball with respect to the norm $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$.

Exercise 1.27. Let $x \in \mathbb{C}^n$; prove that $\|x\|_\infty = \lim_{p\to\infty}\|x\|_p$. This justifies the notation $\|x\|_\infty$.

Exercise 1.28. Using the triangle inequality, prove that

$$|\rho(x, z) - \rho(y, z)| \leqslant \rho(x, y)$$

for any metric ρ and any elements x, y, z.

Exercise 1.29. Prove that the inner product can be recovered from the norm by the *polarization identity*

$$(x, y) = \frac{1}{4}\{(\|x + y\|^2 - \|x - y\|^2) + i(\|x + iy\|^2 - \|x - iy\|^2)\}.$$

Exercise 1.30. Prove that if normed linear space with the norm $\|\cdot\|$ can be equipped with an inner product $(\cdot,\cdot)$ satisfying $\|x\|^2 = (x,x)$, then the norm satisfies the *parellelogram law:*

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

In fact, one can also prove that if the norm satisfies the parallelogram law, then it is generated by an inner product; but this is somewhat more difficult.

Exercise 1.31. Prove that the discrete metric on $\mathbb{C}^n$ is not equivalent to the metric $\rho_p(x,y) = \|x-y\|_p$ for any p.

Exercise 1.32. Let X_1 and X_2 be isometric metric spaces. Suppose that X_1 is bounded; prove that X_2 is also bounded.

Exercise 1.33. Give an example of a metric space (X,ρ) such that for some $a \in X$ one has $B_1(a) = B_2(a)$.

Exercise 1.34. Prove that for any ball $B_r(a)$ in a metric space, one has $\operatorname{diam} B_r(a) \leqslant 2r$.

Exercise 1.35. Let (X,ρ) and (Y,σ) be metric spaces. Prove that the function

$$d\left((x_1,y_1),(x_2,y_2)\right) = \sqrt{\rho(x_1,x_2)^2 + \sigma(y_1,y_2)^2}$$

defined on the Cartesian product $X \times Y$ is a metric.

2 Sequence and function spaces

2.1 Sequence spaces

Example 2.1. Let ℓ^∞ be the set of all infinite complex sequences $x = (x_1, x_2, \ldots)$ such that $\sup_i |x_i| < \infty$. It is not difficult to see that ℓ^∞ is a linear space. Indeed, let $x, y \in \ell^\infty$ and $\alpha, \beta \in \mathbb{C}$. Then

$$\sup_n |\alpha x_n + \beta y_n| \leqslant |\alpha| \sup_n |x_n| + |\beta| \sup_n |y_n|,$$

and so $\alpha x + \beta y \in \ell^\infty$. It is straightforward to see that the axioms of the linear space hold true. Let $\|x\|_\infty = \sup_i |x_i|$. It is easy to see that this is a norm on ℓ^∞.

Example 2.2. Let ℓ^1 be the set of all infinite complex sequences $x = (x_1, x_2, \ldots)$ such that the series $\sum_{i=1}^\infty |x_i|$ converges. Let us check that ℓ^1 is a linear space. As in the previous example, the key statement to prove is that if $x, y \in \ell^1$ and $\alpha, \beta \in \mathbb{C}$, then $\alpha x + \beta y \in \ell^1$. For any n, we have

$$\sum_{i=1}^n |\alpha x_i + \beta y_i| \leqslant |\alpha| \sum_{i=1}^n |x_i| + |\beta| \sum_{i=1}^n |y_i| \leqslant |\alpha| C_1 + |\beta| C_2,$$

where the constants C_1, C_2 are independent of n. Thus, the series $\sum_{i=1}^n |\alpha x_i + \beta y_i|$ converges and so $\alpha x + \beta y \in \ell^1$.

Let us check that $\|x\|_1 := \sum_{i=1}^\infty |x_i|$ is a norm on ℓ^1. Axioms (i), (ii) of the norm are trivial. Axiom (iii) follows from

$$\sum_{i=1}^n |x_i + y_i| \leqslant \sum_{i=1}^n |x_i| + \sum_{i=1}^n |y_i|, \quad n \geqslant 1$$

by taking $n \to \infty$.

Example 2.3. Fix a positive number $p > 1$. Let ℓ^p be the set of all infinite complex sequences $x = (x_1, x_2, \ldots)$ such that $\sum_{i=1}^\infty |x_i|^p < \infty$. Then ℓ^p is a linear space and $\|x\|_p := (\sum_{i=1}^\infty |x_i|^p)^{1/p}$ is a norm on ℓ^p.

Exercise 2.4. 1. Prove that ℓ^p is a linear space. Hint: the key statement to prove is that $x, y \in \ell^p$ and $\alpha, \beta \in \mathbb{C}$, then $\alpha x + \beta y \in \ell^p$. Use Minkowski's inequality for a finite n and then let $n \to \infty$.

2. Prove the triangle inequality for $\|x\|_p$. Again, use Minkowski's inequality for a finite n and then let $n \to \infty$.

Example 2.5. The norm $\|\cdot\|_2$ on ℓ^2 is generated by the inner product $(x, y) = \sum_{i=1}^\infty x_i \, \bar{y}_i$. The series here is absolutely convergent, which follows from the Cauchy-Schwartz inequality. For $p \neq 2$ the norm $\|\cdot\|_p$ cannot be associated with an inner product.

Exercise 2.6. Check the axioms of the inner product for ℓ^2.

2.2 Function spaces

Example 2.7. Let $[a, b] \subset \mathbb{R}$ be a bounded interval. Let $C[a, b]$ be the set of all continuous complex-valued functions on the interval $[a, b]$. It is clear that $C[a, b]$ is a linear space. The standard norm and the corresponding metric on $C[a, b]$ are defined by

$$\|f\|_\infty = \sup_{x\in[a,b]} |f(x)|, \qquad \rho(f, g) = \sup_{x\in[a,b]} |f(x) - g(x)|.$$

Since any continuous function on a bounded closed interval is bounded, we obtain that $\|f\|_\infty < \infty$ for any $f \in C[a, b]$. It is easy to check that the axioms of the norm for $\|\cdot\|_\infty$ are satisfied. This norm is not generated by an inner product.

Example 2.8. Let $C(\mathbb{R})$ be the set of all bounded continuous functions $f : \mathbb{R} \to \mathbb{C}$. It is easy to see that $C(\mathbb{R})$ is a normed linear space with respect to the norm

$$\|f\|_\infty = \sup_{x\in\mathbb{R}} |f(x)|.$$

In the same way, one can define the spaces of continuous functions corresponding to the semi-infinite intervals $[a, \infty)$ and $(-\infty, a]$.

Exercise 2.9. Check the axioms of the norm for $C[a, b]$.

Example 2.10. The linear space $B(S)$ of all bounded (real or complex-valued) functions on a nonempty set S can also be equipped with the norm $\|f\|_\infty = \sup_{x\in S} |f(x)|$.

If $S = [a, b]$ then $C[a, b] \subset B[a, b]$ and the metric introduced in Example 2.10 is the same as in Example 2.7. However, the space $C[a, b]$ is strictly smaller than $B[a, b]$ (every continuous function on $[a, b]$ is bounded but there are bounded functions which are not continuous). Note that $\ell^\infty = B(\mathbb{N})$.

Example 2.11. Fix $1 \leqslant p < \infty$ and let $(a, b) \subset \mathbb{R}$ be a bounded interval. Let $CL^p(a, b)$ be the linear space of all continuous functions f on the interval $[a, b]$ equipped with the norm

$$\|f\|_p = \left(\int_a^b |f(x)|^p \, dx \right)^{1/p}.$$

Since all continuous functions on $[a, b]$ are bounded, this norm is finite. It is easy to see that the axioms of the norm are satisfied. The triangle inequality in this case is called the (integral) Minkowski inequality. It can be proven in the same way as the usual Minkowski inequality; the proof uses the integral Hölder inequality

$$\int_a^b |f(x)g(x)|dx \leqslant \left(\int_a^b |f(x)|^p dx \right)^{1/p} \left(\int_a^b |g(x)|^q dx \right)^{1/q}$$

where $\frac{1}{p} + \frac{1}{q} = 1$. Of a particular importance is the version of this inequality for $p = q = 2$, which is called the (integral) Cauchy-Schwartz inequality. If $p = 2$, then the space $CL^2(a, b)$ can be equipped with an inner product $(f, g) = \int_a^b f(t)\, \bar{g}(t)\, \mathrm{d}t$.

Later on, we will construct the space $L^p(a,b)$ as a *completion* of the space $CL^p(a,b)$.

Example 2.12. Fix $1 \leqslant p < \infty$. Let $CL^p(\mathbb{R})$ be the linear space of all continuous functions $f : \mathbb{R} \to \mathbb{C}$ such that the integral $\int_{\mathbb{R}} |f(x)|^p dx$ converges. Then this space can be equipped with the norm

$$\|f\|_p = \left(\int_{\mathbb{R}} |f(x)|^p \, dx \right)^{1/p}.$$

An important connection between function and sequence spaces is given by the following example.

Example 2.13. Let $\ell^2(\mathbb{Z})$ be the set of all two-sided sequences $a = (\cdots, a_{-1}, a_0, a_1, a_2, \cdots)$ such that $\sum_{n\in\mathbb{Z}} |a_n|^2 < \infty$. Just as in the case of the usual ℓ^2 space, it is easy to prove that $\ell^2(\mathbb{Z})$ is a normed linear space with the norm

$$\|a\|_{\ell^2(\mathbb{Z})} = \left(\sum_{n\in\mathbb{Z}} |a_n|^2 \right)^{1/2}.$$

Recall that for a function $f \in C[-\pi, \pi]$, its Fourier coefficients are given by

$$f_n = \int_{-\pi}^{\pi} f(x) e^{-inx} dx, \quad n \in \mathbb{Z}.$$

Computing the Fourier coefficients of a function can be considered as a map

$$\Phi : CL^2(-\pi, \pi) \to \ell^2(\mathbb{Z}), \quad f \mapsto \{f_n\}_{n\in\mathbb{Z}}.$$

The Parseval's identity says that $\|\Phi f\|_{\ell^2(\mathbb{Z})} = \|f\|_2$. See a Fourier series course for the details.

2.3 Convergence in metric spaces

Definition 2.14. A sequence x_n of elements of a metric space (X, ρ) is said to *converge* to $x \in X$ if $\rho(x_n, x) \to 0$ as $n \to \infty$. In other words, $x_n \to x$ in (X, ρ) if for any $\varepsilon > 0$ there exists an integer n_ε such that $\rho(x_n, x) < \varepsilon$ for all $n > n_\varepsilon$.

In normed linear spaces it suffices to study convergence to zero. Indeed, in normed linear spaces one has $f_n \to f$ if and only if $(f_n - f) \to 0$ as $n \to \infty$.

Theorem 2.15. *Let ρ_1 and ρ_2 be two equivalent metrics defined on a set X and let $\{x_n\}_{n=1}^\infty$ be a sequence of elements of X. Then $x_n \to x$ in (X, ρ_1) if and only if $x_n \to x$ in (X, ρ_2).*

Proof. Suppose that $x_n \to x$ in (X, ρ_1); then

$$\rho_2(x_n, x) \leqslant C \rho_1(x_n, x) \to 0$$

as $n \to \infty$, and so $x_n \to x$ in (X, ρ_2). The converse statement is proven in the same way. ∎

However, there are many interesting examples where metrics are NOT equivalent and the corresponding notions of convergence are different.

Example 2.16. Consider the set ℓ^1 with respect to two norms: $\|\cdot\|_1$ and $\|\cdot\|_\infty$. Then these two metrics are not equivalent. Indeed, let $g_1 = (1, 0, 0, 0 \dots)$, $g_2 = (1, 1, 0, 0, 0 \dots)$, $g_3 = (1, 1, 1, 0, 0, 0 \dots)$ and so on. Consider the sequence of elements $f_n = \frac{1}{n} g_n$. Then $\|f_n\|_\infty = \frac{1}{n}$ and so $f_n \to 0$ with respect to the ℓ^∞-metric. But $\|f_n\|_1 = 1$ for all n and so f_n does not converge to zero with respect to the ℓ^1-metric.

Example 2.17. Consider the set $C[a, b]$ with respect to two metric: the usual metric of $C[a, b]$ and the metric of $CL^1[a, b]$. These metrics are not equivalent. Indeed, consider the sequence $f_n(x) = \exp(-n(x - a))$. Then $f_n \to 0$ in CL^1 but $\sup_{[a,b]}|f_n| = f_n(a) = 1$ for all n and so f_n does not converge to zero in the metric of $C[a, b]$.

Definition 2.18. Convergence in the metric space $B(S)$ (or any subspace of $B(S)$) is called *uniform convergence* on S. Convergence with respect to the L^2-norm (see Example 2.11) is called *mean square convergence.*

2.4 Embeddings

Definition 2.19. Let X be a linear space with the norm $\|\cdot\|_X$ and Y be a linear space with the norm $\|\cdot\|_Y$. If X is a subspace of Y and there exists a constant C such that for all $f \in X$, one has $\|f\|_Y \leqslant C\|f\|_X$, then we write $X \subset Y$ and say that there is an *embedding* of X into Y.

Example 2.20. We have $\ell^p \subset \ell^\infty$ for any $p \geqslant 1$. Indeed, for any n we have $|x_n|^p \leqslant \sum_{i=1}^\infty |x_i|^p$, and therefore $\|x\|_\infty \leqslant \|x\|_p$.

Example 2.21. We have $C[a, b] \subset CL^p(a, b)$ for any $p \geqslant 1$. This follows from the estimate

$$\int_a^b |f(x)|^p dx \leqslant (b - a) \sup_{x \in [a,b]} |f(x)|^p,$$

which can be rewritten as

$$\|f\|_p \leqslant (b - a)^{1/p} \|f\|_\infty, \quad f \in C[a, b].$$

Note that this embedding does not hold true in function spaces over $\mathbb{R}$. Indeed, a function which is identically equal to 1 belongs to $C(\mathbb{R})$ but not to $CL^1(\mathbb{R})$.

Exercises:

Exercise 2.22. Let $\alpha > 0$. Compute the norm of the sequence $x_n = e^{-\alpha n}$ in ℓ^1; ℓ^2; ℓ^∞.

Exercise 2.23. Consider the sequence $x_n = n^{-1/2}$, $n \geqslant 1$. For which $p \geqslant 1$ is the statement $x \in \ell^p$ true?

Exercise 2.24. Consider the sequence $x_n = \frac{1}{\sqrt{n} \log(1+n)}$, $n \geqslant 1$. For which $p \geqslant 1$ is the statement $x \in \ell^p$ true?

Exercise 2.25. Let $f(t) = \sin(t)$ and $g(t) = \cos(t)$. Consider f and g as elements of $C[0, 2\pi]$ and determine the smallest r such that $g \in B_r[f]$.

Exercise 2.26. Let $n \in \mathbb{N}$. Determine the diameter of the set

$$A_n = \{f \in C[0,1] \mid -\tfrac{1}{n} \leqslant f(x) \leqslant x^n\}$$

in $C[0,1]$ and in $CL^1(0,1)$.

Exercise 2.27. Prove that if $p < r$, then $\ell^p \subset \ell^r$. Hint: first prove that $\ell^p \subset \ell^\infty$. Then write $r = p + s$, $s > 0$ and use the inequality $|x_i|^r \leqslant |x_i|^p \|x\|_\infty^s$.

Exercise 2.28. Let $p < r$; consider ℓ^p with respect to two norms: $\|\cdot\|_r$ and $\|\cdot\|_p$. Prove that these norms are not equivalent. Hint: use the sequence g_n from Example 2.16.

Exercise 2.29. Consider the set $CL^1(\mathbb{R})$ with respect to the metrics $\|\cdot\|_1$ and $\|\cdot\|_\infty$. Construct a sequence of functions $f_n \in CL^1(\mathbb{R})$ such that $\|f_n\|_\infty \to 0$ as but $\|f_n\|_1 \geqslant c > 0$ for all n. Construct a sequence of functions $g_n \in CL^1(\mathbb{R})$ such that $\|g_n\|_1 \to 0$ as but $\|g_n\|_\infty \geqslant c > 0$ for all n.

Repeat the exercise for the $CL^p(\mathbb{R})$, $p > 1$ instead of $CL^1(\mathbb{R})$.

Exercise 2.30. Prove that if $p < r$, then $CL^r(a,b) \subset CL^p(a,b)$. Hint: write $r = p + s$, $s > 0$. Using the Hölder inequality, prove the estimate

$$\|f\|_p^p = \int_a^b |f(x)|^p \cdot 1\, dx \leqslant \|f\|_r^p \, \|1\|_{r/s}.$$

From here deduce the required result.

Exercise 2.31. Let x_n be a sequence in a metric space and let x_∞ be given. Suppose that every subsequence of x_n has a subsubsequence converging to x_∞. Prove that $x_n \to x_\infty$.

Exercise 2.32. Prove that if $x_n \to x_\infty$ in a metric space (X, ρ) then for any $y \in X$ one has $\rho(x_n, y) \to \rho(x_\infty, y)$ as $n \to \infty$.

3 Open set and closed sets. Continuity

3.1 Open sets

Definition 3.1. A set in a metric space is *open* if it contains a ball about each of its points.

Exercise 3.2. Prove that an open ball in a metric space (X, ρ) is open.

Theorem 3.3. *If (X, ρ) is a metric space then*

1. *the whole space X and the empty set $\varnothing$ are both open,*
2. *the union of any collection of open subsets of X is open,*
3. *the intersection of any finite collection of open subsets of X is open.*

Proof. (1) The whole space is open because it contains all open balls and the empty set is open because it does not contain any points.

(2) If x belongs to the union of open sets A_ν then x belongs to at least one of the sets A_ν. Since this set is open, it also contains an open ball about x. This ball lies in the union of A_ν, so the union is an open set.

(3) If $A_1, A_2, \ldots, A_k$ are open sets and $x \in \cap_{n=1}^k A_n$ then $x \in A_n$ for every $n = 1, \ldots, k$. Since A_n are open, for each n there exists r_n such that $B_{r_n}(x) \subset A_n$. Let $r = \min\{r_1, r_2, \ldots, r_k\}$. Then $B_r(x) \subset B_{r_n}(x) \subset A_n$ for all $n = 1, \ldots, k$, so $B_r(x) \subset \left(\cap_{n=1}^k A_n\right)$. ∎

Exercise 3.4. Show that the intersection of an infinite collection of open sets is not necessarily open.

Lemma 3.5. *A set is open if and only if it coincides with the union of a collection of open balls.*

Proof. According to Theorem 3.3 the union of any collection of open balls is open. On the other hand, if A is open then for every point $x \in A$ there exists a ball $B(x)$ about x lying in A. We have $A = \cup_{x\in A} B(x)$. Indeed, the union $\cup_{x\in A} B(x)$ is a subset of A because every ball $B(x)$ is a subset of A, and the union contains every point $x \in A$ because $x \in B(x)$. ∎

Definition 3.6. A point $x \in A$ is said to be an *interior* point of the set A if there exists an open ball $B_r(x)$ lying in A. The *interior* of a set A is the union of all open sets contained in A. The interior of A is denoted by $int(A)$.

By Theorem 3.3(ii), the interior of any set is open.

Moreover, the interior of A is easily seen to be the maximal open set contained in A. This means the following: if B is an open set contained in A, then $B \subset int(A)$.

Theorem 3.7. *A set A is open if and only if $A = int(A)$.*

Proof. If $A = int(A)$ then A is open, since $int(A)$ is always open. Conversely, suppose A is open. Then $A \subset A$, A is open and so $A \subset int(A)$. Thus, $A = int(A)$. ∎

Theorem 3.8. *The interior of A coincides with set of all interior points of A.*

Proof. Let $\tilde{A}$ denote the set of all interior points of A.

If x is an interior point then there exists an open ball $B_r(x)$ lying in A. This ball is an open set lying in A and therefore is a subset of the maximal open set $int(A) \subset A$. Thus, $\tilde{A} \subset int(A)$.

Conversely, if $x \in int(A)$ then (since $int(A)$ is open) there exists a ball $B_r(x) \subset int(A) \subset A$, so x is an interior point of A. Thus, $int(A) \subset \tilde{A}$. ∎

Definition 3.9. A set $A \subset X$ is said to be a *neighbourhood* of $\alpha \in X$ if A contains an open ball $B_r(\alpha)$ for some $r > 0$.

Using this terminology, we can rephrase the definition of an open set as follows: *a set is open iff it contains a neighbourhood of each of its points.*

3.2 Closed sets

Definition 3.10. A set A is *closed* if the limit of any convergent sequence of elements of A lies in A. In other words, A is closed if $x_n \in A$, $x = \lim_{n\to\infty} x_n$ implies $x \in A$.

Exercise 3.11. Prove that a closed ball in a metric space (X, ρ) is closed.

Definition 3.12. If $A \subset X$ then A^c denotes the complement of the set A in X, that is, the set of all points $x \in X$ which do not belong to A.

Theorem 3.13. *If A is open then A^c is closed. If A is closed then A^c is open.*

Proof. 1. Suppose A is open and let $x \in A$. Then there exists a ball $B_r(x) \subset A$. This ball does not contain elements of A^c. Thus, if a sequence of elements of A^c converges, it cannot converge to x. It follows that all limits of convergent sequences of elements of A^c lie in A^c, so A^c is closed.

2. Suppose A is closed and let $x \in A^c$. Consider the balls $B_{1/n}(x)$. Suppose that each of these balls contains an element of A. Then we can choose a sequence of elements of A which converges to x. Since A is closed, it follows that $x \in A$; this contradicts to the choice $x \in A^c$. Thus, there exists n such that $B_{1/n}(x)$ does not contain elements of A; then $B_{1/n}(x) \subset A^c$. Since x is arbitrary, it follows that A^c is open. ∎

Exercise 3.14. Prove that, in a metric space (X, ρ),

1. the whole space X and the empty set $\varnothing$ are both closed,

2. the intersection of any collection of closed sets is closed,

3. the union of any finite collection of closed sets is closed.

Hint This follows from Theorems 3.3 and 3.13. It is also a useful exercise to prove this directly, i.e. by using the definition of a closed set.

Definition 3.15. The *closure* of a set A is the intersection of all closed sets containing A. The closure is denoted by $\bar{A}$; another notation is $cl(A)$.

By Exersise 3.14, the closure of any set is closed.

Moreover, the closure of A is easily seen to be the minimal closed set containing A. This means that if $A \subset B$ and B is closed, then $\bar{A} \subset B$.

Theorem 3.16. *A set A is closed if and only if $A = \bar{A}$.*

Proof. If $A = \bar{A}$, then A is closed, since $\bar{A}$ is always closed. Conversely, suppose that A is closed. Then we have $A \subset A$, A is closed and therefore $\bar{A} \subset A$. It follows that $\bar{A} = A$, as required. ∎

Theorem 3.17. *$x \in \bar{A}$ if and only if there exists a sequence $\{x_n\} \subset A$ which converges to x.*

Proof. 1. Suppose that $x = \lim_{n\to\infty} x_n$ where $x_n \in A$ for all n. Let's prove that $x \in \bar{A}$. Suppose $A \subset K$ where K is closed. Then $x_n \in K$ for all n and therefore $x \in K$. Thus, $x \in \bar{A}$.

2. Suppose that $x \in \bar{A}$. If $B_{1/n}(x) \cap A$ is non-empty for any $n \in \mathbb{N}$, then there exists a sequence $x_n \in A$ such that $x = \lim_{n\to\infty} x_n$, and we are done. Suppose that $B_{1/n}(x) \cap A$ is empty for some n. Consider the closed set $K = (B_{1/n}(x))^c$. We have $A \subset K$ and therefore $\bar{A} \subset K$. By assumption, $x \in \bar{A} \subset K$ — contradiction! ∎

Example 3.18. Let (X, ρ) be a metric space with the discrete metric

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

Then

$$B_r[a] = \begin{cases} \{a\}, & \text{if } r < 1, \\ X, & \text{if } r \geq 1. \end{cases} \qquad B_r(a) = \begin{cases} \{a\}, & \text{if } r \leq 1, \\ X, & \text{if } r > 1, \end{cases}$$

Since the open ball is open, this implies that any point is an open set. Since any set coincides with the union of its elements, Theorem 3.3 implies that any subset of X is open. Therefore, by Theorem 3.13, any subset of X is closed.

The closure of the open ball $B_r(a)$ does not necessarily coincide with the closed ball $B_r[a]$. In particular, in the Example 3.18 $\overline{B}_1(a) = B_1(a) = a$ (since $B_1(a)$ is closed) but $B_1[a] = X$.

Definition 3.19. A point $x \in X$ is called a *limit* point of a set A if every ball about x contains a point of A distinct from x. Other terms for "limit point" are point of accumulation or cluster point.

Lemma 3.20. *A point x is a limit point of a set A if and only if there is a sequence x_n of elements of A distinct from x which converges to x.*

Proof. If $x_n \to x$ then every ball about x contains a point x_n. If every ball about x contains a point of A distinct from x then there exists a sequence of points $x_n \in A$ distinct from x and lying in the balls $B_{1/n}(x)$. Obviously, this sequence converges to x. ∎

3.3 Dense sets

Definition 3.21. A subset A of a metric space (X, ρ) is said to be *dense* if its closure $cl(A)$ coincides with X.

Exercise 3.22. Prove that A is dense in (X, ρ) if and only if any open ball in (X, ρ) contains an element of A.

Example 3.23. $\mathbb{Q}$ is dense in $\mathbb{R}$ with the usual metric.

3.4 Continuity

Given a map $f : X \to Y$ and a subset $A \subset Y$, the set $\{x \in X : f(x) \in A\}$ is denoted $f^{-1}(A)$ and called the *inverse image* of A. Note that $f^{-1}(A)$ is a well-defined set irrespective of whether f has an inverse.

Definition 3.24. Let (X, ρ) and (Y, d) be metric spaces. A map $f : X \to Y$ is said to be *continuous at* $\alpha \in X$ if for any open ball $B_\varepsilon(f(\alpha))$ about $f(\alpha)$ there exists a ball $B_\delta(\alpha)$ about α such that $B_\delta(\alpha) \subset f^{-1}\left(B_\varepsilon(f(\alpha))\right)$. A map $f : X \to Y$ is said to be *continuous* if it is continuous at every point $\alpha \in X$.

Theorem 3.25. *Let (X, ρ) and (Y, σ) be metric space and $f : X \to Y$ be a map from X to Y. Then the following statements are equivalent:*

1. *f is continuous;*
2. *the inverse image of every open subset of Y is an open subset of X.*

Proof. Assume first that f is continuous. Let A be an open subset of Y and let $x \in f^{-1}(A) \subset X$. Since A is open, there exists a ball $B_\varepsilon(f(x))$ about the point $f(x)$ such that $B_\varepsilon(f(x)) \subset A$. Since f is continuous, there exists a ball $B_\delta(x)$ about x such that $B_\delta(x) \subset f^{-1}(B_\varepsilon(f(x))) \subset f^{-1}(A)$. Therefore for every point $x \in f^{-1}(A)$ there exists a ball $B_\delta(x)$ lying in $f^{-1}(A)$, which means that $f^{-1}(A)$ is open.

Assume now that the inverse image of any open set is open. Let $x \in X$ and $B_\varepsilon(f(x))$ is a ball about $f(x) \in Y$. The inverse image $f^{-1}(B_\varepsilon(f(x)))$ is an open set which contains the point x. Therefore there exists a ball $B_\delta(x)$ about x such that $B_\delta(x) \subset f^{-1}(B_\varepsilon(f(x)))$. This implies that f is continuous. ∎

Exercise 3.26. Let (X_k, ρ_k) be metric spaces and $T_1 : X_1 \to X_2$, $T_2 : X_2 \to X_3$ be continuous maps. Prove that the composition $T_2 \circ T_1$ is a continuous map from X_1 to X_3.

Exercises:

Exercise 3.27. Prove that the closure satisfies the following *Kuratowski axioms* (here $cl(A) = \bar{A}$):

(i) $cl(cl(A)) = cl(A)$;

(ii) $cl(A \cup B) = cl(A) \cup cl(B)$;

(iii) $A \subset cl(A)$;
(iv) $cl(\varnothing) = \varnothing$.

Of course, (iv) is a triviality and (iii) follows directly from Theorem 3.17; your task is to prove (i) and (ii).

Exercise 3.28. Give an example which shows that $\overline{A \cap B} = \overline{A} \cap \overline{B}$ is FALSE.

Exercise 3.29. Show that in any metric space, $\overline{B_r(\alpha)} \subset B_r[\alpha]$. Give an example which shows that in general in may happen that $\overline{B_r(\alpha)} \neq B_r[\alpha]$.

Exercise 3.30*. Prove that in a normed linear space $\overline{B_r(\alpha)} = B_r[\alpha]$.

Exercise 3.31. Prove that the following sets are open in the given metric spaces:

1. $\{f \in C[0,1] \mid -1 < f(0) < 1\}$ in $C[0,1]$;
2. $\{f \in C[-1,1] \mid \int_0^1 |f(x)| dx < 1\}$ in $C[-1,1]$;
3. $\{x \in \ell^1 \mid x_1 < \sum_{n=2}^{\infty} x_n 2^{-n}\}$ in ℓ^1.

Exercise 3.32. Prove that the following sets are closed in the given metric spaces:

1. $\{x \in \ell^2 \mid |x_{2n}| = 1/n \quad \forall n \in \mathbb{N}\}$ in ℓ^2;
2. $\{f \in C[-1,1] \mid \int_{-1}^0 f(x) dx \leqslant \int_0^1 f(x) dx\}$ in $C[-1,1]$;
3. $\{f \in C[0,\pi] \mid |f(x)| \leqslant |\sin(x)| \quad \forall x \in [0,\pi]\}$ in $C[0,\pi]$.

Exercise 3.33. Show that the map $f : (X,\rho) \to (Y,d)$ is continuous at $x_0 \in X$ if and only if for every sequence x_n converging to x_0 in (X,ρ), the sequence $f(x_n)$ converges to $f(x_0)$ in (Y,d).

Exercise 3.34. Show by constructing an example that the image of an open set under a continuous map need not be open.

Exercise 3.35. Let X be a set with two equivalent metrics, ρ_1 and ρ_2 defined on it. Prove that a set $A \subset X$ is open in the metric space (X,ρ_1) if and only if it is open in (X,ρ_2).

Exercise 3.36. Let X and Y be two normed linear spaces such that the embedding $X \subset Y$ holds true. Let A be a subset of X. Suppose that A is open as a subset of Y; prove that A is open as a subset of X. Construct an example which shows that the converse is not true (i.e. A may be open in X yet not open in Y).

Exercise 3.37. Let $\{x_n\}_{n=1}^{\infty}$ be a convergent sequence and $K = \cup_{n=1}^{\infty}\{x_n\}$. Prove that K has exactly one limit point, which is $\lim_{n\to\infty} x_n$.

Exercise 3.38*. Prove that the subset of $C[a,b]$, which consists of all polynomials, is not closed. (In fact, one can prove that its closure is the whole of $C[a,b]$; this is the Weierstrass' approximation theorem). Hint: consider the function $\exp(x)$ and prove that this function can be approximated by polynomials uniformly on $[a,b]$. Use Taylor's series for $\exp(x)$ to construct this series of polynomials. You will need to use some estimate for the remainder in the Taylor's series in order to prove convergence of the polynomials to $\exp(x)$.

4 Completeness. Separable spaces

4.1 Complete and incomplete spaces; completion

Definition 4.1. A sequence x_n of elements of a metric space (X, ρ) is called a *Cauchy sequence* if, given any $\varepsilon > 0$, there exists n_ε such that $\rho(x_n, x_m) < \varepsilon$ for all $n, m > n_\varepsilon$.

Exercise 4.2. Show that every convergent sequence is a Cauchy sequence.

Definition 4.3. A metric space (X, ρ) is said to be *complete* if any Cauchy sequence $\{x_n\} \subset X$ converges to a limit $x \in X$.

There exist incomplete metric spaces. If a metric space (X, ρ) is not complete then it has Cauchy sequences which do not converge. This means, in a sense, that there are gaps (or missing elements) in X. Every incomplete metric space can be made complete by adding new elements, which can be thought of as the missing limits of non-convergent Cauchy sequences. More precisely, we have the following theorem.

Theorem 4.4. *Let X be an arbitrary metric space. Then there exists a complete metric space Y such that X is isometric to a dense subset Y_0 of Y.*

The metric space Y is said to be the *completion* of X. Note that the condition that Y_0 is dense simply ensures that Y is "not too large". For the proof of the this theorem, see e.g. exercises to chapter 3 of the book W. Rudin, *Principles of mathematical analysis.* One can also prove that the completion is unique up to an isometry, i.e. if Y_1 and Y_2 are completions of X, then Y_1 is isometric to Y_2.

If the space X is bounded then one can give a simple proof of Theorem 4.4. This proof is outlined in the exercises.

In practice, completions are often obtained by means of the following construction.

Theorem 4.5. *Let (A, ρ) be a subspace of a complete metric space (X, ρ) and $\bar{A}$ be the closure of A in (X, ρ). Then $(\bar{A}, \rho)$ is a completion of (A, ρ).*

Proof. Let us prove that $(\bar{A}, \rho)$ is complete. Let $\{x_n\}$ be a Cauchy sequence in $\bar{A}$. Since (X, ρ) is complete and $\bar{A} \subset X$, this sequence converges to some element $x \in X$. Since $\bar{A}$ is closed, we have $x \in \bar{A}$. Therefore the space $(\bar{A}, \rho)$ is complete.

The fact that A is dense in $\bar{A}$ follows from the definition of a dense set. ∎

Example 4.6. Let $\mathbb{Q}$ be the set of rational numbers with the standard metric $\rho(x, y) = |x - y|$. This metric space is not complete because any sequence x_n which converges to an irrational number is a Cauchy sequence but does not have a limit in $\mathbb{Q}$. The completion of this space is the set of all real numbers $\mathbb{R}$ with the same metric $\rho(x, y) = |x - y|$.

Any irrational number in the interval $(0, 1)$ can be written as an infinite decimal fraction $0.a_1a_2\ldots$ or, in other words, can be identified with the Cauchy sequence 0, $0.a_1$, $0.a_1a_2$, ... of rational numbers which does not converge to a rational limit.

The space of real numbers $\mathbb{R}$ is *defined* as the completion of the space of rational numbers and therefore, by definition, is complete.

Example 4.7. Since $\mathbb{R}$ is complete, the space of complex numbers $\mathbb{C}$ with the standard metric $\rho(x, y) = |x - y|$ is also complete. Indeed, if $\{c_n\}$ is a sequence of complex numbers and $c_n = a_n + ib_n$, where $a_n = \operatorname{Re} c_n$ and $b_n = \operatorname{Im} c_n$, then

1. $\{c_n\}$ is a Cauchy sequence if and only if $\{a_n\}$ and $\{b_n\}$ are Cauchy sequences of real numbers;
2. the sequence $\{c_n\}$ converges if and only if the sequences $\{a_n\}$ and $\{b_n\}$ converge.

4.2 Banach spaces

Definition 4.8. A complete normed linear space is called a *Banach space.*

This term is in honour of the great Polish mathematician Stefan Banach (1892–1945), the founder of functional analysis.

Example 4.9. The space $\mathbb{C}^n$ with the norm $\|\cdot\|_p$, $1 \leqslant p \leqslant \infty$ is a Banach space. The proof of this follows the same argument as in Example 4.7. Indeed, if $x^{(k)}$ is a Cauchy sequence, then all of its coordinates $x_j^{(k)}$, $j = 1, \dots, n$ form Cauchy sequences and therefore each of the sequences of the coordinates converges to a limit x_j. Then $x^{(k)}$ converges to $x = (x_1, \dots, x_n)$.

Theorem 4.10. *ℓ^p is a Banach space for any p, $1 \leqslant p \leqslant \infty$.*

Proof. 1. Let $x^{(k)}$ be a Cauchy sequence in ℓ^p. Fix $i \in \mathbb{N}$. Since $|x_i| \leqslant \|x\|_p$, we see that $x_i^{(k)}$ is a Cauchy sequence of complex numbers. Since $\mathbb{C}$ is complete, we see that $x_i^{(k)} \to x_i$ as $k \to \infty$.

2. Since $x^{(k)}$ is Cauchy, for a given ε there exists $N_\varepsilon \in \mathbb{N}$ such that

$$\sum_{i=1}^{n} |x_i^{(k)} - x_i^{(m)}|^p \leqslant \varepsilon, \qquad \forall k, m \geqslant N_\varepsilon, \quad \forall n \in \mathbb{N}.$$

Let $m \to \infty$ here; then we get

$$\sum_{i=1}^{n} |x_i^{(k)} - x_i|^p \leqslant \varepsilon, \qquad \forall k \geqslant N_\varepsilon, \quad \forall n \in \mathbb{N}.$$

Since n is arbitrary, we get the convergence of the series and the inequality

$$\sum_{i=1}^{\infty} |x_i^{(k)} - x_i|^p \leqslant \varepsilon, \qquad \forall k \geqslant N_\varepsilon. \tag{4.1}$$

For any given k this means that the sequence $x_i^{(k)} - x_i$ is an element of ℓ^p. Since $x_i^{(k)}$ is also an element of ℓ^p and ℓ^p is a linear space, we get that $x = (x_1, x_2, \dots) \in \ell^p$.

Now (4.1) can be rewritten as $\|x^{(k)} - x\|_p^p \leqslant \varepsilon$, $k \geqslant N_\varepsilon$. Since ε can be taken arbitrary small, we get that $\|x^{(k)} - x\|_p \to 0$ as $k \to \infty$. ∎

Theorem 4.11. *Let S be any nonempty set. Then $B(S)$ is a Banach space.*

Proof. 1. Let f_k be a Cauchy sequence in $B(S)$. Fix $x \in S$. Since $|f(x)| \leqslant \|f\|_\infty$, we see that $f_k(x)$ is a Cauchy sequence of complex numbers. Since $\mathbb{C}$ is complete, we see that the limit $\lim_{k\to\infty} f_k(x)$ exists. Denote this limit by $f(x)$. We need to prove that $f \in B(S)$ and $\|f_k - f\|_\infty \to 0$ as $k \to \infty$.

2. Since f_k is Cauchy, for a given ε there exists N_ε such that

$$\|f_k - f_m\|_\infty \leqslant \varepsilon \qquad \forall k, m \geqslant N_\varepsilon,$$

or

$$|f_k(x) - f_m(x)| \leqslant \varepsilon \qquad \forall k, m \geqslant N_\varepsilon, \quad \forall x \in S.$$

Let $m \to \infty$ here; we get

$$|f_k(x) - f(x)| \leqslant \varepsilon \qquad \forall k \geqslant N_\varepsilon, \quad \forall x \in S.$$

Since x is arbitrary, we get

$$\sup_{x\in S} |f_k(x) - f(x)| \leqslant \varepsilon \qquad \forall k \geqslant N_\varepsilon.$$

Thus, we get that f is bounded and

$$\|f_k - f\|_\infty \leqslant \varepsilon \qquad \forall k \geqslant N_\varepsilon.$$

This means that $\|f_k - f\|_\infty \to 0$ as $k \to \infty$. ∎

Corollary 4.12. *$C[a, b]$ is a Banach space.*

Proof. Since continuous functions on $[a, b]$ are bounded, Theorem 4.11 implies that any Cauchy sequence of continuous functions f_k uniformly converges to a bounded function f on $[a, b]$, and we only need to prove that the function f is continuous.

In order to prove that we have to show that for any convergent sequence $x_n \to x$ in $[a, b]$ and any $\varepsilon > 0$ there exists N such that for all $n \geqslant N$ one has $|f(x_n) - f(x)| \leqslant \varepsilon$. First choose m such that $\|f_m - f\| \leqslant \varepsilon/3$. Next, since f_m is continuous, we can choose N such that for all $n \geqslant N$, one has $|f_m(x_n) - f_m(x)| \leqslant \varepsilon/3$. Then for all $n \geqslant N$ we have

$$\begin{aligned} |f(x_n) - f(x)| &= |f(x_n) - f_m(x_n) + f_m(x_n) - f_m(x) + f_m(x) - f(x)| \\ &\leq |f(x_n) - f_m(x_n)| + |f_m(x_n) - f_m(x)| + |f_m(x) - f(x)| \leqslant \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon, \end{aligned}$$

as required. ∎

Example 4.13. The space $CL^p(a, b)$ is incomplete for any $1 \leqslant p < \infty$. For simplicity of notation, consider the case $(a, b) = (-1, 1)$. Let $f_n \in CL^p(-1, 1)$ be defined by

$$f_n(x) = \begin{cases} -1 & \text{if } -1 \leqslant x \leqslant -1/n; \\ nx & \text{if } -1/n \leqslant x \leqslant 1/n; \\ 1 & \text{if } 1/n \leqslant x \leqslant 1. \end{cases}$$

Then f_n is continuous and a simple calculation shows that

$$\|f_n - f_m\|_p^p \leqslant \frac{2}{n} + \frac{2}{m}.$$

Thus, $\{f_n\}$ is a Cauchy sequence. However, we claim that f_n does not converge in the space $CL^p(-1,1)$. Let us prove this claim. Let g be the discontinuous function,

$$g(x) = \begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0. \end{cases}$$

Since g is discontinuous, it is not an element of $CL^p(-1,1)$. However, one can still compute the norm of g and $g - f$ according to the same formula as in the definition of the norm of CL^p. By the same argument as above one shows that $\|f_n - g\|_p \to 0$ as $n \to \infty$.

Now suppose that there exists $f \in CL^p(-1,1)$ such that $\|f_n - f\|_p \to 0$. It is not difficult to see that $\|f - g\|_p > 0$. By the triangle inequality for the norm $\|\cdot\|_p$, we get that $\|f - g\|_p \leqslant \|f - f_n\|_p + \|f_n - g\|_p \to 0$ as $n \to \infty$ — contradiction.

Definition 4.14. Let $1 \leqslant p < \infty$. Then $L^p(a,b)$ is defined as the completion of $CL^p[a,b]$. That is, $L^p(a,b)$ is the completion of $C[a,b]$ with respect to the norm $\|\cdot\|_p$.

Let x_n be a sequence of elements of a normed linear space X.

Definition 4.15. The series $\sum_{n=1}^{\infty} x_n$ is said to be *convergent* if the sequence σ_k defined by $\sigma_k = \sum_{n=1}^{k} x_n$ is convergent in X. If $\sigma_k \to x \in X$ as $k \to \infty$ then we write $\sum_{n=1}^{\infty} x_n = x$. The series $\sum_{n=1}^{\infty} x_n$ is said to be *absolutely convergent* if $\sum_{n=1}^{\infty} \|x_n\| < \infty$.

Theorem 4.16. *In a Banach space every absolutely convergent series is convergent.*

Proof. Let $s_k = \sum_{n=1}^{k} \|x_n\|$. Since the series $\sum_{n=1}^{\infty} \|x_n\|$ is convergent, the sequence of positive numbers $\{s_k\}$ converges and therefore it is a Cauchy sequence. If $m > k$, $\sigma_m = \sum_{n=1}^{m} x_n$ and $\sigma_k = \sum_{n=1}^{k} x_n$ then

$$\rho(\sigma_m, \sigma_k) = \|\sigma_m - \sigma_k\| = \| \sum_{n=k+1}^{m} x_n \| \leq \sum_{n=k+1}^{m} \|x_n\| = |s_m - s_k| \,.$$

This implies that $\{\sigma_n\}_{n=1,2,\ldots}$ is a Cauchy sequence in our Banach space, and therefore it converges. ∎

4.3 Separable spaces

Definition 4.17. Two sets A and B are said to have the same *cardinality* if there exists a bijection $f : A \to B$. A set A is said to be *finite* if there exists $n \in \mathbb{N}$ such that A has the same cardinality as $\{1, 2, \ldots, n\}$. A set A is said to be *countable* if it is either finite or has the same cardinality as $\mathbb{N}$.

For finite sets, cardinality is the same as the number of elements.

Theorem 4.18. *A subset of a countable set is countable.*

Proof. It suffices to prove that any subset A of $\mathbb{N}$ is countable. Let

$$\begin{aligned} a_1 &= \min A, \quad A_1 = A \setminus \{a_1\}, \\ a_2 &= \min A_1, \quad A_2 = A_1 \setminus \{a_2\}, \\ a_3 &= \min A_2, \quad A_3 = A_2 \setminus \{a_3\}, \text{etc.} \end{aligned}$$

If at some stage A_n is empty, then A is finite. If not, then the map $n \mapsto a_n$ establishes a bijection between $\mathbb{N}$ and A. ∎

Theorem 4.19. *Let A be a set such that there exists a surjection $f : \mathbb{N} \to A$. Then A is countable.*

Proof. It is easy to construct a set $B \subset \mathbb{N}$ such that the restriction of f onto B is a bijection between B and A (for example, for any $a \in A$ let $g(a) = \min f^{-1}(a)$ and set $B = \{g(a) \mid a \in A\}$). Now it remains to use the previous theorem. ∎

Theorem 4.20. *A countable union of countable sets is countable.*

Proof. Let A_n be countable sets, where $n \in N$ and N is countable. First assume that all of the sets N, A_1, A_2, ... are infinite. Then we may assume that $N = \mathbb{N}$. Let $A_n = \{a_1^{(n)}, a_2^{(n)}, a_3^{(n)}, \dots\}$. The elements of $A = \cup_{n=1}^{\infty} A_n$ can be listed as

$$a_1^{(1)}, a_2^{(1)}, a_1^{(2)}, a_3^{(1)}, a_2^{(2)}, a_1^{(3)}, a_4^{(1)}, \dots$$

(plot a diagram). This establishes a surjection $f : \mathbb{N} \to A$. Note that if the intersection $\cap_n A_n$ is non-empty, then this will not be an injection, because some of the elements of A will be listed more than once. In any case, applying the previous theorem, we obtain that A is countable.

If some of the sets N, A_1, A_2, ... are finite, then we can add extra elements to these sets to make them infinite and then use Theorem 4.18. ∎

Corollary 4.21. *If A and B are countable sets, then the Cartesian product $A \times B$ is countable.*

Example 4.22. The set $\mathbb{Z}$ is countable. The set $\mathbb{Q}$ is countable. If $A_1, A_2, \dots A_n$ are countable, then the Cartesian product $A_1 \times A_2 \times \cdots \times A_n$ is countable.

Theorem 4.23. *For any non-empty set X, the set of all subsets of X (usually denoted by 2^X) cannot have the same cardinality as X.*

Proof. Assume $f : X \to 2^X$ is a bijection. Consider $S = \{x \in X \mid x \notin f(x)\} \subset X$. Let $s \in X$ be such that $f(s) = S$. Is $s \in S$? Both answers (yes and no) lead to a contradiction. ∎

Corollary 4.24. *$\mathbb{R}$ and $\mathbb{N}$ have different cardinalities.*

Proof. It is easy to see that the interval $[0,1]$ has the same cardinality as $2^{\mathbb{N}}$. Indeed, any real number from this interval can be written down uniquely in the form of a binary expansion $0.a_1a_2a_3\dots$ where $a_j \in \{0,1\}$. Such a binary expansion uniquely defines a subset of $\mathbb{N}$. This establishes a bijection between $[0,1]$ and $2^{\mathbb{N}}$. ∎

Definition 4.25. A Banach space is called *separable*, if it contains a countable dense subset.

Example 4.26. Since $\mathbb{Q}$ is countable and dense in $\mathbb{R}$, the space $\mathbb{R}$ is separable. In the same way, $\mathbb{R}^n$ is separable.

Theorem 4.27. *For any $p < \infty$, the Banach space ℓ^p is separable. The space ℓ^∞ is not separable.*

Proof. 1. Let $p < \infty$. Let $A_n \subset \ell^p$ be the set of all sequences $x \in \ell^p$ such that all coordinates of x are rational and $x_k = 0$ for all $k > n$. Let $A = \cup_{n=1}^{\infty} A_n$. It is clear that each of the sets A_n is countable and therefore A is countable.

Let us prove that A is dense in ℓ^p. For a given n, denote by $P_n : \ell^p \to \ell^p$ the map

$$P_n(y) = (y_1, y_2, \dots, y_n, 0, 0, \dots).$$

For a given $y \in \ell^p$ and a given $\varepsilon > 0$ one can always find $n > 0$ such that $\|y - P_n(y)\|_p \leqslant \varepsilon/2$. Now one can find $x \in A_n$ such that $\|P_n(y) - x\|_p \leqslant \varepsilon/2$. Thus, $\|y - x\|_p \leqslant \varepsilon$.

2. Let us prove that ℓ^∞ is not separable. For any $X \subset \mathbb{N}$, let $e^X \in \ell^\infty$ be defined by the following rule: $e_n^X = 1$ if $n \in X$ and $e_n^X = 0$ if $n \notin X$. Then the number of such elements in ℓ^∞ is uncountable and for any $X \neq Y$ we have $\|e^X - e^Y\|_\infty = 1$.

Suppose that ℓ^∞ has a dense countable subset D. Consider the collection of balls $B_{1/2}(e^X)$ over all $X \subset \mathbb{N}$. Since $2^{\mathbb{N}}$ is uncountable, this collection of balls is uncountable. On the other hand, these balls are disjoint and there is an element of D in each of these balls. Thus, we have a bijection between a subset of D (which by Theorem 4.18 is countable) and the set of all such balls. This is a contradiction. ∎

Theorem 4.28. *The spaces $C[a,b]$ and $L^p(a,b)$ are separable.*

Proof. The fact that $C[a,b]$ is separable follows from the Weierstrass approximation theorem: *the set of all polynomials is dense in $C[a,b]$*. Every polynomial can be approximated by a polynomial with rational coefficients. The set of all polynomials with rational coefficients is countable. Thus, $C[a,b]$ is separable.

Since $C[a,b]$ is dense in $L^p(a,b)$, the countable set of all polynomials with rational coefficients is also dense in $L^p(a,b)$. ∎

Exercise 4.29. Prove that the set of all polynomials with rational coefficients is countable.

Exercises

Exercise 4.30. Prove the following statement: if a Cauchy sequence has a convergent subsequence then it is convergent with the same limit.

Exercise 4.31. Prove that the following metric spaces are incomplete and construct their completions: (1) $\mathbb{R}$ with the metric $\rho(x,y) = |\tan^{-1}(x) - \tan^{-1}(y)|$; (2) $\mathbb{R}$ with the metric $\rho(x,y) = |e^x - e^y|$.

Exercise 4.32. Is the space $\mathbb{N}$ with the metric $\rho(x,y) = \frac{|n-m|}{nm}$ complete?

Exercise 4.33. Give an example of two metrics ρ_1 and ρ_2 on the set $X = (0,1)$ such that (X,ρ_1) is complete and (X,ρ_2) is incomplete.

Exercise 4.34. Prove that if the series $\sum_{n=1}^{\infty} \rho(x_n, x_{n+1})$ converges, then $x_1, x_2, \ldots$ is a Cauchy sequence.

Exercise 4.35. The aim of this exercise is to provide a proof of Theorem 4.4 in the case of a bounded metric space. Let (X,ρ) be our metric space. Consider the map $f : X \to B(X)$ defined as follows:

$$f : X \ni x \mapsto \rho(x,\cdot) \in B(X).$$

Prove that f is an isometry between X and the image of X in $B(X)$. Using this fact, Theorem 4.5 and Theorem 4.11, conclude that X has a completion.

Exercise 4.36.* Prove that if every absolutely convergent series in a normed linear space is convergent, then this space is complete. Hint: Given a Cauchy sequence x_n, find a subsequence x_{n_k} such that $\sum_{k=1}^{\infty} \|x_{n_{k+1}} - x_{n_k}\| < \infty$. Then use the result of Exercise 4.30.

Exercise 4.37. Let f be the set of all sequences of complex numbers, $x = (x_1, x_2, \ldots)$ such that only finitely many terms of these sequences are non-zero. Prove that f is a linear space. Prove that f is dense in ℓ^p for any $p < \infty$. Prove that f is NOT dense in ℓ^∞.

Exercise 4.38.* Prove that completeness of a metric space is equivalent to the following *nested balls property:*

Let $B_1 \supset B_2 \supset \cdots$ be a sequence of closed balls in a metric space such that the radius of B_n tends to zero as $n \to \infty$. Then the intersection $\cap_{n=1}^{\infty} B_n$ consists of one (and only one) point.

Remarks: (i) the fact that the intersection consists of at most one point is a triviality. (ii) The fact that a complete metric space satisfies the nested balls property is easy: consider the sequence of centers of the balls,... (iii) The fact that a nested balls property implies completeness is slightly more difficult. You have to use the trick of Exercise 4.36 and given a Cauchy sequence, use an appropriate subsequence and consider balls centered at the elements of this subsequence.

5 Compactness in metric spaces

5.1 Three definitions of compactness

Intervals which are bounded and closed figure prominently in analysis on the real line. The appropriate generalization of their essential properties that are relevant to analysis in more general spaces is compactness. There are three definitions of compactness in metric spaces which can be shown to be equivalent.

Definition 5.1. A subset K of a metric space (X, ρ) is said to be *(sequentially) compact* if any sequence of elements of K has a subsequence which converges to a limit in K.

It is clear from the definition that K is compact in (X, ρ) if and only if it is compact in (X, σ) for any metric σ equivalent to ρ.

The second definition needs a little terminology. If $\hat{S}$ is a family of subsets of X and $K \subset \cup_{S \in \hat{S}} S$ then $\hat{S}$ is called a *cover* of K. If each member of $\hat{S}$ is open, it is called an *open cover* of K. If $\hat{S}$ is a cover of K and a subset $\hat{S}_0$ of $\hat{S}$ also covers K then $\hat{S}_0$ is called a *subcover* of $\hat{S}$. A cover (or subcover) is said to be finite if it has a finite number of members.

Definition 5.2. A subset K of a metric space (X, ρ) is said to be *compact* if any open cover of K has a finite subcover.

Definition 5.3. A metric space (X, ρ) is said to be (sequentially) compact if the set X is (sequentially) compact.

Definition 5.4. A set K in a metric space is said to be *totally bounded*, if for any $\varepsilon > 0$ this set is contained in a union of finitely many balls of the radius ε.

Obviously, every totally bounded set is bounded. As we will see shortly, the converse is not true.

Theorem 5.5. *In a complete metric space, the following statements are equivalent for a set K:*

(i) K is compact;
(ii) K is sequentially compact;
(iii) K is closed and totally bounded.

In fact, completeness is not needed in the proof of the equivalence of (i) and (ii). The formulation (iii) appears to be especially useful in applications. The most common way of proving compactness in concrete function spaces is by constructing the set of balls as in Definition 5.4.

Proof. We prove (i)⇒(ii)⇒(iii)⇒(i).

1. Assume (i) and let $\{x_n\}_{n=1}^{\infty}$ be a sequence in K. Let E be the set of values of this sequence, i.e. $E = \cup_{n=1}^{\infty}\{x_n\}$. Let us prove that E is finite; then it straightforward to see that the sequence $\{x_n\}_{n=1}^{\infty}$ must have a convergent subsequence.

Take any $x \in K$. Since x is not a point of accumulation of E, there exists $r > 0$ such that $B_r(x)$ contains finitely many elements of the set E. Choosing r in this way for every element $x \in K$, we obtain an open cover of K such that each set in this cover contains finitely many elements of E. Choosing a finite subcover of this cover, we obtain that E is finite, as required.

2. Assume (ii). Let us prove that K is closed. Let $\{x_n\}_{n=1}^{\infty}$ be a convergent sequence of elements of K. By (ii), this sequence has a limit point in K. Since the sequence is convergent, it has only one limit point, which is its limit. Thus, $x = \lim_{n\to\infty} x_n \in K$, as required.

Let us prove that K is totally bounded. Let $\varepsilon > 0$ be given. We need to choose finitely many elements x_1, x_2, ... such that the balls $B_\varepsilon > 0$ cover the set K. Pick $x_1 \in K$. We have two possibilities: either $K \subset B_\varepsilon(x_1)$ or not. In the first case, we are done: the ball $B_\varepsilon(x_1)$ covers K. Suppose not; then we can choose $x_2 \in K$ such that $\rho(x_1, x_2) \geqslant \varepsilon$. We have two possibilities: either $K \subset B_\varepsilon(x_1) \cup B_\varepsilon(x_2)$ or not. In the first case, we are done. Suppose not; then choose $x_3 \in K$ such that $\rho(x_1, x_3) \geqslant \varepsilon$ and $\rho(x_2, x_3) \geqslant \varepsilon$.

Continuing this way, at the nth step we obtain the elements x_1, x_2, ..., x_n such that $\rho(x_j, x_i) \geqslant \varepsilon$ for any $i \neq j$. We claim that this process must eventually stop, i.e. at some stage the constructed balls of radius ε will cover K (this will prove the statement). If not, then we obtain an infinite sequence $\{x_n\}_{n=1}^{\infty}$ which has no convergent subsequences; this contradicts our assumption.

3. Assume (iii) and let Γ be an open cover of K. Suppose (to reach a contradiction) that no finite subcollection of Γ covers K.

By (iii), $K \subset \cup_{i=1}^{n} B_1(x_i) \subset \cup_{i=1}^{n} B_1[x_i]$, i.e. K is covered by finitely many closed balls of radius 1. For at least one index i, the set $K \cap B_1[x_i]$ cannot be covered by finitely many members of Γ. Denote $K_1 = K \cap B_1[x_i]$.

Using (iii) again in the same way, we obtain that K_1 can be covered by closed balls $\cup_{i=1}^{m} B_{1/2}[y_i]$. Denote $K_2 = K_1 \cap B_{1/2}[y_i]$, where i is such that K_2 cannot be covered by finitely many members of Γ.

Continuing this process, we obtain a sequence of closed sets $K \supset K_1 \supset K_2 \supset \dots$ such that $\operatorname{diam} K_n \leqslant 2/n$ and no K_n can be covered by finitely many members of Γ.

For each n, choose $x_n \in K_n$. Then x_n is a Cauchy sequence. Since our metric space is complete, the limit $x = \lim_{n\to\infty} x_n$ exists. Since K is closed, we have $x \in K$. Hence $x \in V$ for some $V \in \Gamma$.

Since V is open, $B_r(x) \subset V$ for some $r > 0$. Choose n such that $2/n < r$. Then $K_n \subset B_r(x) \subset V$. This contradicts the assumption. ∎

5.2 Properties and examples of compact sets

By Theorem 5.5, a compact set is bounded and closed.

However, the converse is not true. In fact, we have:

Theorem 5.6. *A closed ball in a Banach space is compact if and only if the space is finite dimensional.*

This theorem will be proven in the next section. Here we illustrate the "only if" part by giving two examples.

Example 5.7. Let us show that a closed unit ball centered at zero in ℓ_p is not compact for any $1 \leqslant p \leqslant \infty$. Indeed, the elements $x_n = (0, \dots, 0, 1, 0, \dots)$ (where 0 is on the n'th position) belong to this ball and the distance between any pair of these elements is $2^{1/p}$ if $p < \infty$ and is 1 if $p = \infty$. Thus, this sequence is not Cauchy and so cannot have a convergent subsequence.

Example 5.8. Let us show that the ball $B_1[0]$ is not compact in $C[a, b]$. Choose a sequence of disjoint open intervals $I_n \subset [a, b]$, $n \in \mathbb{N}$. Next, for each n, let f_n be a continuous function which vanishes on $[a, b] \setminus I_n$ such that $0 \leqslant f_n(x) \leqslant 1$ for all $x \in I_n$ and $f(x_n) = 1$ for some point $x_n \in I_n$. Clearly, such a sequence can be constructed (for example, one can take f_n to be piecewise linear and equal to 1 at the midpoint of I_n). Then $f_n \in B_1[0]$ and $\|f_n - f_m\| = 1$ for any $n \neq m$ and so the sequence $\{f_n\}_{n=1}^{\infty}$ does not have a convergent subsequence.

In particular, these examples show that not all bounded sets are totally bounded.

Lemma 5.9. *A closed subset of a compact set is compact.*

Proof. Let K be compact, K_0 be a closed subset of K and $\{x_n\}_{n=1}^{\infty}$ be a sequence of elements of K_0. By the compactness of K, this sequence has a convergent subsequence in K. Since K_0 is closed, the limit of this subsequence lies in K_0. Therefore any sequence of elements of K_0 has a subsequence which converges to a limit in K_0 which means that K_0 is compact. ∎

Theorem 5.10. *Let $\mathbb{R}^n_2$ be the space $\mathbb{R}^n$ with the usual Euclidean metric generated by the norm $\|\cdot\|_2$. Then any bounded and closed subset in $\mathbb{R}^n_2$ is compact.*

In order to prove this theorem, we need an argument which allows to reduce the situation to the case $n = 1$. This argument is presented in a more general form than we need it for the proof of the Theorem.

Lemma 5.11. *If K and L are compact subsets of metric spaces (X, ρ) and (Y, σ) respectively then $K \times L$ as a subset of $X \times Y$ with the metric*

$$d\left((x_1, y_1), (x_2, y_2)\right) = \sqrt{\rho(x_1, x_2)^2 + \sigma(y_1, y_2)^2}$$

is compact.

Proof. Let (x_n, y_n) be an arbitrary sequence in $K \times L$. Since K is compact, there is a subsequence x_{n_k} which converges to a limit $x \in K$ as $k \to \infty$. Since L is compact, the sequence y_{n_k} has a subsequence $y_{n_{k_i}}$ which converges to a limit $y \in L$ as $i \to \infty$. Since $x_{n_k} \to x$ as $k \to \infty$, we also have $x_{n_{k_i}} \to x$ as $i \to \infty$. By definition of convergence, $\rho(x_{n_{k_i}}, x) \to 0$ and $\sigma(y_{n_{k_i}}, y) \to 0$ as $i \to \infty$. This implies that $d\big((x_{n_{k_i}}, y_{n_{k_i}}), (x, y)\big) \to 0$ as $i \to \infty$, that is, $(x_{n_{k_i}}, y_{n_{k_i}}) \to (x, y) \in K \times L$. Therefore any sequence (x_n, y_n) of elements of $K \times L$ has a subsequence which converges to a limit in $K \times L$. ∎

Proof of Theorem 5.10. Since any bounded subset lies in a closed cube Q^n, in view of Lemma 5.9 it is sufficient to prove that the closed cube is compact. The closed cube Q^n is a direct product of a one dimensional closed cube Q^1 (a closed interval) and a closed cube $Q^{n-1} \subset \mathbb{R}^{n-1}$. If Q^1 and Q^{n-1} are compact then, by Lemma 5.11, Q^n is also compact. Therefore it is sufficient to prove that a closed interval is compact (after that the required result is obtained by induction in n).

Let x_n be an arbitrary sequence of numbers lying in a closed interval $[a, b]$. Let us split $[a, b]$ into the union of two intervals of length $\delta/2$, where $\delta = b - a$. At least one of these intervals contains infinitely many elements x_n of our sequence. Let us choose one of these elements and denote it by y_1. Now we split the interval of length $\delta/2$ which contains infinitely many elements x_n into the union of two intervals of length $\delta/4$. Again, at least one of these intervals contains infinitely many elements x_n. We choose one of these elements (distinct from y_1) and denote it by y_2. Repeating this procedure, we obtain a subsequence $\{y_k\}$ of the sequence $\{x_n\}$ such that y_k lie in an interval of length 2^{-k_0} for all $k \geq k_0$. Clearly, $\{y_k\}$ is a Cauchy sequence. Since $\mathbb{R}$ is a complete metric space, $\{y_k\}$ converges to a limit. Since a closed interval is a closed set, this limit belongs to $[a, b]$. Thus, any sequence of elements of $[a, b]$ has a subsequence which converges to a limit in $[a, b]$, which means that the closed interval is compact. ∎

Example 5.12. In ℓ^2, consider the set K of all sequences x satisfying

$$|x_1| \leqslant 2^{-1}, \quad |x_2| \leqslant 2^{-2}, \quad |x_2| \leqslant 2^{-3}, \cdots$$

The set K is sometimes called "Hilbert's brick". Let us prove that K is compact. By Theorem 5.5, it suffices to prove that K is closed and totally bounded. The set K is closed because it can be represented as an intersection of closed sets

$$K = \cap_{n=1}^{\infty} K_n, \quad K_n = \{x \in \ell^2 \mid |x_n| \leqslant 2^{-n}\}.$$

Next, given $\varepsilon > 0$, we need to choose a finite set S such that

$$K \subset \cup_{x \in S} B_\varepsilon(x). \tag{5.1}$$

Choose n so that $2^{-n-1} \leqslant \varepsilon$. Consider the map

$$x = (x_1, x_2, \dots, x_n, \dots) \mapsto x^* = (x_1, x_2, \dots, x_n, 0, 0 \dots).$$

It is easy to see that $\|x - x^*\| \leqslant \varepsilon/2$. The set K^* of points x^* in K is totally bounded (as a closed bounded set in a finite dimensional space) and therefore there exists a finite set S such that $K^* \subset \cup_{x \in S} B_{\varepsilon/2}(x)$. Then it is clear that (5.1) is also satisfied.

Exercises

Exercise 5.13. The family of intervals $(\frac{1}{n+2}, \frac{1}{n})$, $n \in \mathbb{N}$ is an open cover of the set $(0, 1)$. Is it possible to choose a finite subcover of this cover?

Exercise 5.14. Give an example of a non-compact set $S \subset \mathbb{R}$ and an open cover of S such that there exists a finite subcover of this cover.

Exercise 5.15. Prove that any finite set is compact.

Exercise 5.16. Prove that the intervals (a, b) and $[0, \infty)$ are not compact. Provide three proofs: (i) by exhibiting an open cover with no finite subcover; (ii) by exhibiting a sequence with no convergent subsequence; (iii) by showing that the set fails to be closed and totally bounded.

Exercise 5.17. Construct a sequence in the unit ball $B_1[0]$ in $L^1(0, 1)$ such that this sequence has no convergent subsequences. Repeat this for $L^2(0, 1)$.

Exercise 5.18. For the following subsets of $C[0, 1]$, decide whether these subsets are compact:

1. The whole space $C[0, 1]$;
2. The unit ball in $C[0, 1]$;
3. The unit sphere: $\{f \in C[0, 1] \mid \|f\|_C = 1\}$;
4. The set of all polynomials;
5. The set of all polynomials whose coefficients are bounded above by 1.
6. The set of all polynomials of degree $\leqslant n$ whose coefficients a_i satisfy $|a_i| \leqslant 1$.

Exercise 5.19. Prove that a union and an intersection of two compact sets are compact.

Exercise 5.20. Let X be an infinite set and ρ be a discrete metric on X. Prove that the metric space (X, ρ) is not compact.

Exercise 5.21. Prove that if a set A in a metric space is totally bounded, then the closure of A is also totally bounded.

6 Compactness: Further properties

6.1 Compactness and continuity

Theorem 6.1. *The image of a compact set by a continuous map is compact.*

Proof. Let K be a compact set and f be a continuous map. Let y_n be an arbitrary sequence of elements of $f(K)$. Then $y_n = f(x_n)$ where $x_n \in K$. Since K is compact, the sequence $\{x_n\}$ has a subsequence $\{x_{n_k}\}$ which converges to a limit $x \in K$. Then by Exercise 3.33 the subsequence $y_{n_k} = f(x_{n_k})$ converges to the limit $f(x) \in f(K)$. This proves that $f(K)$ is compact. ∎

Corollary 6.2. *Let (X, ρ) be a compact metric space and let $f : X \to \mathbb{R}$ be a continuous function. Then the maximum and minimum of f on X are finite and are attained on X.*

Proof. The image $f(X)$ is a closed bounded set, hence $\max f(X)$ and $\min f(X)$ exist and are finite. ∎

Definition 6.3. We say that a (real or complex-valued) function f defined on a metric space $f : (X, \rho)$ is *uniformly continuous* if for any $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(x) - f(y)| \leq \varepsilon$ whenever $\rho(x, y) \leq \delta$.

Obviously, a uniformly continuous function is continuous.

Theorem 6.4. *If (X, ρ) is a compact metric space then any continuous function f on (X, ρ) is uniformly continuous.*

Proof. Let $\varepsilon > 0$. Since f is continuous, for every point $x \in X$ there exists $\delta_x > 0$ such that

$$|f(y) - f(x)| \leq \varepsilon/2 \quad \text{whenever} \quad \rho(y, x) \leq \delta_x. \tag{6.1}$$

Let $J_x = B_{\delta_x/2}(x)$. Since $x \in J_x$, the collection of open balls $\{J_x\}_{x \in X}$ is an open cover of X. Since X is compact, it has a finite subcover, that is, there exists a finite collection of points $x_1, x_2, \ldots, x_n$ such that $X = \cup_{k=1}^n J_{x_k}$. Denote $\delta = \frac{1}{2} \min\{\delta_{x_1}, \ldots \delta_{x_n}\}$. Since the number of points x_k is finite, we have $\delta > 0$.

Let $x, y \in X$ and $\rho(x, y) \leq \delta$. Since $X = \cup_{k=1}^n J_{x_k}$, there exists k such that $x \in J_{x_k}$, that is, $\rho(x, x_k) \leq \delta_{x_k}/2$. By the triangle inequality

$$\rho(y, x_k) \leq \rho(x, x_k) + \rho(x, y) \leq \delta_{x_k}/2 + \delta \leq \delta_{x_k}$$

and, in view of (6.1), $|f(y) - f(x)| \leq |f(y) - f(x_k)| + |f(x_k) - f(x)| \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$. Thus we have proved that for any $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(x) - f(y)| \leq \varepsilon$ whenever $\rho(x, y) \leq \delta$. ∎

6.2 Finite dimensional subspaces

Corollary 6.5. *In a finite dimensional space all norms are equivalent.*

Proof. It suffices to prove that any norm $\|\cdot\|$ on $\mathbb{R}^n$ is equivalent to the standard Euclidean norm $\|\cdot\|_2$. Let $e_1 \ldots, e_n$ be the standard basis in $\mathbb{R}^n$. Then we have

$$\|x\| = \|\sum_{i=1}^{n} x_i e_i\| \leqslant \sum_{i=1}^{n} |x_i| \|e_i\| \leqslant C\|x\|_2, \tag{6.2}$$

where $C^2 = \|e_1\|^2 + \cdots + \|e_n\|^2$. In order to prove the opposite inequality, consider the function $f(x) = \|x\|$ on the unit sphere $K = \{x \in \mathbb{R}^n \mid \|x\|_2 = 1\}$. The function f is continuous by (6.2). The sphere K is compact in $(\mathbb{R}^n, \|\cdot\|_2)$ since it is a bounded and closed set. Thus, f attains a minumum m on K. By the axioms of the norm, $m > 0$. Thus, $\|x\| \geqslant m$ for all $x \in K$. By rescaling, we get $\|x\| \geqslant m\|x\|_2$ for all $x \in \mathbb{R}^n$, which completes the proof. ∎

Corollary 6.6. *In a finite dimensional Banach space, any closed ball is compact.*

Proof. Let our Banach space be $\mathbb{R}^n$ with the norm $\|\cdot\|$. By Corollary 6.5, any closed ball $B_r[x]$ in this Banach space is bounded and closed in $\mathbb{R}^n_2$. Thus, $B_r[x]$ is compact with respect to the Euclidean norm. Thus, it is compact with respect to the norm $\|\cdot\|$. ∎

This proves the "if" part of Theorem 5.6. Let us prove the "only if" part.

Exercise 6.7. Prove that any finite-dimensional subspace of a Banach space is closed.

Lemma 6.8. *Let X_0 be a closed linear subspace of a normed space X, $X_0 \neq X$. Then for any $\varepsilon > 0$ there exists $x_0 \in X \setminus X_0$ such that $\|x_0\| = 1$ and $\|x_0 - x\| \geq 1 - \varepsilon$, $\forall x \in X_0$.*

Proof. Take an arbitrary $\varepsilon \in (0, 1)$ (there is nothing to prove if $\varepsilon \geq 1$). Since $X_0 \neq X$, there exists $x_1 \in X \setminus X_0$. Since X_0 is closed, $d := \inf_{x \in X_0} \|x_1 - x\| > 0$ (why?). There exists $y \in X_0$ such that $d \leq \|x_1 - y\| \leq d/(1 - \varepsilon)$ (why?). Let

$$x_0 := \frac{1}{\|x_1 - y\|}(x_1 - y).$$

Then $\|x_0\| = 1$. Since X_0 is a linear subspace, $y + \|x_1 - y\|x \in X_0$, $\forall x \in X_0$, and

$$\begin{aligned} \|x_0 - x\| &= \frac{1}{\|x_1 - y\|} \|x_1 - (y + \|x_1 - y\|x)\| \\ &\geq \frac{1 - \varepsilon}{d} \|x_1 - (y + \|x_1 - y\|x)\| \geq \frac{1 - \varepsilon}{d} d = 1 - \varepsilon. \end{aligned}$$

∎

Theorem 6.9. *If every bounded subset of a Banach space X is compact, then X is finite-dimensional.*

Proof. Suppose, to get a contradictions, that X is infinite dimensional. Take an arbitrary $x_1 \in X$ with $\|x_1\| = 1$ and let X_1 be the one-dimensional linear subspace spanned by x_1, i.e. $X_1 := \{\lambda x_1 | \lambda \in \mathbb{C}\} \subset X$. Since X_1 is closed (see Exercise 6.7) and $X_1 \neq X$, Lemma 6.8 (with $\varepsilon = 1/2$) implies that there exists $x_2 \in X$ such that $\|x_2\| = 1$ and $\|x_2 - x\| \geq 1/2$, $\forall x \in X_1$. In particular, $\|x_2 - x_1\| \geq 1/2$.

Suppose we have already constructed $x_1, \ldots, x_n \in X$, such that

$$\|x_k\| = 1, \quad \|x_k - x_j\| \geq 1/2, \quad k \neq j \tag{6.3}$$

for all $k, j = 1, \ldots, n$. Let X_n be the n-dimensional linear subspace spanned by $x_1, \ldots, x_n$. Since X_n is closed (see Exercise 6.7) and $X_n \neq X$ (because X is infinite dimensional), Lemma 6.8 implies that there exists $x_{n+1} \in X$ such that $\|x_{n+1}\| = 1$ and $\|x_{n+1} - x\| \geq 1/2$, $\forall x \in X_n$. This inductive procedure produces an infinite sequence $x_k \in X$, which satisfies (6.3) for all $k, j \in \mathbb{N}$. It is clear that this sequence does not have Cauchy subsequences. Hence the bounded set $\{x_k\}_{k \in \mathbb{N}} \subset X$ is not compact. ∎

6.3 Arzela-Ascoli theorem

The question of compactness of subsets of concrete metric spaces is a common problem is analysis. One usually proceeds by using Theorem 5.5 and checking that the set in question is closed and totally bounded. Closedness is usually not difficult and the real issue is total boundedness. Because of this, it is convenient to have concrete criteria of total boundedness in concrete metric spaces.

Here we consider such a criterion for $C[a, b]$. We need to introduce some definitions. A family $\Phi \subset C[a, b]$ of functions is called *uniformly bounded*, if there exists a number M such that $|\phi(x)| \leqslant M$ for all $\phi \in \Phi$ and all $x \in [a, b]$. A family $\Phi \subset C[a, b]$ of functions is called *equicontinuous*, if for any $\varepsilon > 0$ there exists $\delta > 0$ such that $|\phi(x_1) - \phi(x_2)| \leqslant \varepsilon$ for all $\phi \in \Phi$ and all $x_1, x_2 \in [a, b]$ such that $|x_1 - x_2| \leqslant \delta$.

Theorem 6.10. *A subset $\Phi \subset C[a, b]$ is totally bounded if and only if it is uniformly bounded and equicontinuous.*

Proof. 1. Suppose that Φ is totally bounded. Then Φ is bounded, i.e. $\Phi \subset B_r(0)$ for some $r > 0$. Thus, Φ is uniformly bounded.

Next, let us prove that Φ is equicontinuous. Given $\varepsilon > 0$, there exist finitely many functions $f_1, \ldots, f_n$ such that $\Phi \subset \cup_{k=1}^n B_{\varepsilon/3}(f_k)$. Each of the functions f_k is continuous on $[a, b]$ and therefore (by Theorem 6.4) there exists $\delta > 0$ such that for all k,

$$|f_k(x) - f_k(y)| \leqslant \varepsilon/3 \quad \text{whenever} \quad |x - y| \leqslant \delta.$$

Now let $f \in \Phi$; then $f \in B_{\varepsilon/3}(f_k)$ for some k. It follows that for any x, y such that $|x - y| < \delta$:

$$|f(x) - f(y)| \leqslant |f(x) - f_k(x)| + |f_k(x) - f_k(y)| + |f_k(y) - f(y)| \leqslant \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon.$$

Thus, Φ is equicontinuous.

2. Let Φ be uniformly bounded and equicontinuous, and let $\varepsilon > 0$ be given. Equicontinuity of Φ guarantees that for each $x \in [a, b]$ there exists $\delta = \delta(x) > 0$ such that if $|x - y| < \delta$ then $|f(x) - f(y)| < \varepsilon$. By compactness of $[a, b]$, the open cover $\{B_{\delta(x)}(x)\}_{x \in [a,b]}$ has a finite subcover $\{B_{\delta_i}(x_i)\}_{i=1}^n$.

Let M be the constant from the definition of uniform boundedness of Φ, and let $D^n \subset \mathbb{C}^n$ be the set

$$D^n = \{(z_1, \dots, z_n) \in \mathbb{C}^n \mid |z_1| \leqslant M, \quad , |z_n| \leqslant M\}.$$

For each $f \in \Phi$, let $p(f) = (f(x_1), \dots, f(x_n))$ where $\{x_i\}_{i=1}^n$ are as above. Since D^n is a closed bounded set in $\mathbb{C}^n$, it is compact and therefore totally bounded. Thus, there exist functions $f_1, \dots, f_m \in \Phi$ such that every $p(f)$ lies within ε of some $p(f_k)$.

Let us prove that the collection of balls $B_{3\varepsilon}(f_k)$, $k = 1, \dots, m$, covers Φ. Choose any $f \in \Phi$; then there exists k, $1 \leqslant k \leqslant m$, such that

$$|f(x_i) - f_k(x_i)| < \varepsilon, \quad \forall i = 1, \dots, n.$$

Next, any $x \in [a, b]$ lies in some $B_{\delta_i}(x_i)$, $i = 1, \dots, n$, and therefore

$$|f(x) - f(x_i)| < \varepsilon \quad \text{and} \quad |f_k(x) - f_k(x_i)| < \varepsilon.$$

Thus,

$$|f(x) - f_k(x)| \leqslant |f(x) - f(x_i)| + |f(x_i) - f_k(x_i)| + |f_k(x_i) - f_k(x)| \leqslant 3\varepsilon,$$

as required. Since $\varepsilon > 0$ is arbitrary, we get that Φ is totally bounded. ∎

Example 6.11. Let $K \subset C[a, b]$ be the set of all continuously differentiable functions which satisfy

$$|\phi(x)| \leqslant 1, \quad |\phi'(x)| \leqslant 1, \quad \forall x \in [a, b].$$

Then K is totally bounded. Indeed,

$$|\phi(x_2) - \phi(x_1)| \leqslant \int_{x_1}^{x_2} |\phi'(x)| dx \leqslant |x_2 - x_1|.$$

It follows that K is equicontinuous.

6.4 Compact embeddings

Definition 6.12. Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed linear spaces such that there is an embedding $Y \subset X$. This embedding is called *compact* if the unit ball $\{\phi \in Y \mid \|\phi\|_Y \leqslant 1\}$ is totally bounded in X.

Example 6.13. Let $C^1[a, b]$ be the linear space of all continuously differentiable functions $f : [a, b] \to \mathbb{C}$ with the norm

$$\|f\|_{C^1} = \sup_{x \in [a,b]} |f(x)| + \sup_{x \in [a,b]} |f'(x)|.$$

Then, by Example 6.11, the embedding $C^1[a, b] \subset C[a, b]$ is compact.

Exercise 6.14. Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed linear spaces such that there is an embedding $Y \subset X$. Let $D \subset Y$ be a dense set in $(Y, \|\cdot\|_Y)$ and suppose that the set $\{\phi \in D \mid \|\phi\|_Y \leqslant 1\}$ is totally bounded in X. Prove that the embedding $Y \subset X$ is compact. Hint: use Exercise 5.21.

Example 6.15. Let $C_0^1[a, b]$ be the linear space of all continuously differentiable functions $f : [a, b] \to \mathbb{C}$ such that $f(a) = f(b) = 0$. Let $H_0^1(a, b)$ be the completion of $C_0^1[a, b]$ with respect to the norm

$$\|f\|_{H^1}^2 = \int_a^b (|f'(x)|^2 + |f(x)|^2)dx.$$

The space $H_0^1(a, b)$ is one of the large and extremely useful family of Sobolev spaces.

Let us prove that the embedding $H_0^1(a, b) \subset L^2(a, b)$ is compact. Since we have a bounded embedding $C[a, b] \subset L^2(a, b)$, it suffices to prove (see Exercise 6.22) that the embedding $H_0^1(a, b) \subset C(a, b)$ is compact. By the previous exercise, it suffices to prove that the set

$$B = \{f \in C_0^1[a, b] \mid \|f\|_{H^1} \leqslant 1\}$$

is totally bounded. By the Arzela-Ascoli theorem, we need to prove that B is uniformly bounded and equicontinuous. For $x_1, x_2 \in [a, b]$, $x_1 < x_2$, and any $f \in C_0^1[a, b]$ we have:

$$\begin{aligned} |f(x_2) - f(x_1)| &= \left| \int_{x_1}^{x_2} f'(t)dt \right| \leqslant \int_{x_1}^{x_2} |f'(t)|dt \\ &\leqslant \left(\int_{x_1}^{x_2} |f'(t)|^2 dt \right)^{1/2} \sqrt{x_2 - x_1} \leqslant \|f'\|_2 \sqrt{x_2 - x_1} \leqslant \|f\|_{H^1} \sqrt{x_2 - x_1}. \end{aligned}$$

Thus, B is equicontinuous. Taking $x_1 = a$, we get

$$|f(x)| \leqslant \|f'\|_2 \sqrt{b - a}$$

and so B is uniformly bounded.

Example 6.16. One can look at the above example from a different point of view. In $L^2[a, b]$, consider the set K of f such that the derivative f' is well defined and

$$\int_a^b (|f(x)|^2 + |f'(x)|^2)dx \leqslant 1.$$

This set is compact in $L^2[a, b]$. This can be proved as follows. Each $f \in L^2[a, b]$ can be expanded in trigonometric series. The map from f to the set of coefficients of the trigonometric series of f is an isometry from $L^2[a, b]$ to ℓ^2. In ℓ^2, the set isometric to K is the one defined by

$$\sum_{n=1}^{\infty} n^2 |x_n|^2 \leqslant 1.$$

We need to prove that this set is compact. This can be done exactly as in Example 5.12.

Exercises

Exercise 6.17. Prove (by constructing a counterexample) that the assumption of compactness cannot be dropped from Theorem 6.1.

Exercise 6.18. In ℓ^2, consider the set K of all sequences x satisfying $|x_n| \leqslant \frac{1}{n}$, $n \in \mathbb{N}$. Prove that K is compact in ℓ^2.

Exercise 6.19. Consider the set ℓ^2_1 of all sequences x satisfying $\sum_{n=1}^{\infty} n^2|x_n|^2 < \infty$. Prove that ℓ^2_1 is a normed linear space with the norm

$$\|x\| = \left(\sum_{n=1}^{\infty} n^2|x_n|^2 \right)^{1/2}.$$

Prove that the embedding $\ell^2_1 \subset \ell^2$ is compact.

Exercise 6.20. For each of the following functions f, decide whether f is uniformly continuous on the given set and prove your answer.

1. $f(x) = \sin(1/x)$, $x \in (0, 1)$.
2. $f(x) = x\sin(1/x)$, $x \in (0, 1)$.
3. $f(x) = \sqrt{x}\sin(1/x)$, $x \in (0, 1)$.
4. $f(x) = 1/(1 + x^2(\sin x)^2)$, $x \in \mathbb{R}$.
5. $f(x) = (\sin(x^2))/(1 + x^2)$, $x \in \mathbb{R}$.

Exercise 6.21. For each of the following sets of functions, decide whether this set is totally bounded in the given space and prove your answer. You may use the Arzela-Ascoli theorem.

1. $f_a(x) = \sin(ax)$, $a \in [0, 1]$, in $C[0, \pi]$.
2. $f_a(x) = \sin(ax)$, $a \in [0, \infty)$, in $C[0, \pi]$.
3. $f_a(x) = x^a$, $a \in [0, 2]$, in $C[0, 1]$.
4. $f_a(x) = x^a$, $a \in [1, \infty)$, in $C[0, 1]$.
5. $f_a(x) = x^a$, $a \in [1, \infty)$, in $C[0, \frac{1}{2}]$.

Exercise 6.22. Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be two Banach spaces such that we have a bounded embedding $X \subset Y$. Suppose that $A \subset X$ is a set which is totally bounded in $(X, \|\cdot\|_X)$. Prove that A is totally bounded in $(Y, \|\cdot\|_Y)$.

Exercise 6.23. Construct a non-compact metric space such that any continuous function on this space is uniformly continuous.

7 Linear continuous functionals. The Hahn-Banach Theorem

7.1 Definition and examples

Let X be a Banach space with the norm $\|\cdot\|_X$.

Definition 7.1. A linear continuous functional on X is a map $f : X \to \mathbb{C}$ such that

(i) f is linear, i.e. $f(x+y) = f(x) + f(y)$ and $f(\lambda x) = \lambda f(x)$ for all $x, y \in X$ and $\lambda \in \mathbb{C}$;

(ii) f is bounded, i.e. there exists $C > 0$ such that

$$|f(x)| \leqslant C\|x\| \quad \forall x \in X.$$

The set of all linear continuous functionals on X is called a dual space to X and is denoted by X^*.

The norm on X^* is defined by of the norm on $\mathcal{B}(X, \mathbb{C})$:

$$\|f\|_{X^*} = \sup_{x \in X \setminus \{0\}} \frac{|f(x)|}{\|x\|_X}.$$

Theorem 7.2. *X^* is a Banach space.*

Proof. The proof bears a close resemblance to the proof of completeness of $C[a, b]$. We need to show that X^* is complete.

Let $f_n \in X^*$ be a Cauchy sequence. Fix $x \in X$. Since $|f(x)| \leqslant \|f\|_{X^*}\|x\|_X$, we see that $f_n(x)$ is a Cauchy sequence of complex numbers. Since $\mathbb{C}$ is complete, we see that $f_n(x) \to f(x)$ for some functional f on X. It is easy to see that f is linear (see Exercise below). Our task is to show that f is bounded and $\|f_n - f\|_{X^*} \to 0$ as $n \to \infty$.

Since f_n is Cauchy, for a given ε there exists $N_\varepsilon \in \mathbb{N}$ such that

$$|f_n(x) - f_m(x)| \leqslant \varepsilon\|x\|_X, \quad \forall n, m \geqslant N_\varepsilon, \quad \forall x \in X.$$

Let $m \to \infty$ here; then we get

$$|f_n(x) - f(x)| \leqslant \varepsilon\|x\|_X, \quad \forall n \geqslant N_\varepsilon, \quad \forall x \in X.$$

Since x is arbitrary, we get that f is a bounded functional and

$$\|f_n - f\|_{X^*} \leqslant \varepsilon \quad \forall n \geqslant N_\varepsilon.$$

This proves that $\|f_n - f\|_{X^*} \to 0$ as $n \to \infty$. ∎

Exercise 7.3. Complete the above proof by showing that f is linear.

Example 7.4. Linear functionals on $C[a,b]$: $f \mapsto \int_a^b f(x)dx$, $f \mapsto \int_a^b g(x)f(x)dx$, $f \mapsto f(x_0)$ for some $x_0 \in [a,b]$.

In a Hilbert space H: $f \mapsto (f,g)$ for some element $g \in H$ is a continuous functional. We will see later that in a Hilbert space, all functionals have this form and so one has $H^* = H$.

Example 7.5. Let us prove that ℓ_1^* is isometric to ℓ_∞. First, it is easy to see that for any $y \in \ell_\infty$, the functional

$$\lambda_y(x) = \sum_{n=1}^{\infty} y_n x_n, \quad x \in \ell_1,$$

is bounded in ℓ_1 and $\|\lambda_y\|_{\ell_1^*} \leqslant \|y\|_\infty$. Thus, we have a bounded linear map $J : \ell_\infty \to \ell_1^*$ with $\|J\| \leqslant 1$.

Let $y \in \ell_\infty$. For any $\epsilon > 0$, let $n \in \mathbb{N}$ be such that $|y_n| \geqslant \|y\|_\infty - \epsilon$. Then

$$|\lambda_y(e_n)| = |y_n| \geqslant \|y\|_\infty - \epsilon \geqslant (\|y\|_\infty - \epsilon)\|e_n\|_1,$$

and therefore $\|\lambda_y\|_{\ell_1^*} \geqslant \|y\|_\infty - \epsilon$. Since $\epsilon > 0$ is arbitrary, we get $\|\lambda_y\|_{\ell_1^*} = \|y\|_\infty$.

It remains to prove that J is onto. Let $\lambda \in \ell_1^*$. Set $y_n = \lambda(e_n)$; then, clearly, we have $y = (y_1, y_2, \dots) \in \ell_\infty$. Consider the functional $\mu = \lambda - \lambda_y \in \ell_1^*$. We have $\mu(e_n) = 0$ for all n. Thus, μ vanishes on f (here f is the set of all sequences x such that $\{n \in \mathbb{N} \mid x_n \neq 0\}$ is finite). Since f is dense in ℓ_1, we get that $\mu = 0$ and therefore $\lambda = \lambda_y$.

Example 7.6. Let us prove that ℓ_p^* is isometric to ℓ_q, where $1 < p < \infty$ and $1/p + 1/q = 1$. First, by using Hölder inequality, it is easy to see that for any $y \in \ell_q$, the functional

$$\lambda_y(x) = \sum_{n=1}^{\infty} y_n x_n, \quad x \in \ell_p,$$

is bounded in ℓ_p and $\|\lambda_y\|_{\ell_p^*} \leqslant \|y\|_q$. Thus, we have a bounded linear map $J : \ell_q \to \ell_p^*$ with $\|J\| \leqslant 1$.

Let $y \in \ell_p$. Define $x_n = \overline{y_n}|y_n|^{q-2}$ if $y_n \neq 0$ and $x_n = 0$ if $y_n = 0$. Then $\|x\|_p^p = \|y\|_q^q$ and $\lambda_y(x) = \|y\|_q^q$. It follows that $\lambda_y(x) = \|x\|_p \|y\|_q$, and so $\|\lambda_y\|_{\ell_p^*} = \|y\|_q$.

It remains to prove that J is onto. Let $\lambda \in \ell_p^*$. Set $y_n = \lambda(e_n)$ and let x_n be as defined above. For $N \in \mathbb{N}$, let $x = (x_1, \dots, x_N, 0, \dots)$. Then, by the same calculation as above,

$$\sum_{n=1}^{N} |y_n|^q = \lambda(x) \leqslant \|\lambda\|_{\ell_p^*} \|x\|_p = \|\lambda\|_{\ell_p^*} \big(\sum_1^N |y_n|^q\big)^{q/p},$$

and so $\sum_1^N |y_n|^q \leqslant \|\lambda\|_{\ell_p^*}^q$. Since N is arbitrary, we get that $y \in \ell_q$. Now as in the previous example, one checks that $\lambda = \lambda_y$.

Example 7.7. In the same way, $(L^p[a,b])^* = L^q[a,b]$ with $1 < p < \infty$ and $1/p + 1/q = 1$ (without proof).

7.2 Hahn–Banach Theorem

As above, let X be a Banach space with the norm $\|\cdot\|_X$.

Theorem 7.8. *Let Y be a linear subset of X. Let $\lambda : Y \to \mathbb{C}$ be a linear functional such that $|\lambda(x)| \leqslant a\|x\|_X$ for all $x \in Y$. Then there exists $\Lambda \in X^*$ such that $\|\Lambda\|_{X^*} \leqslant a$ and $\Lambda(x) = \lambda(x)$ for all $x \in Y$.*

In other words, any bounded linear functional on a subspace of a Banach space can be extended onto the whole Banach space without increasing the norm.

Proof of the Hahn-Banach theorem. For simplicity assume $a = 1$. We will only give the proof for a separable space X.

1. *Reduction of the complex case to the real case.* Set $\ell(x) = \operatorname{Re}\lambda(x)$; this is a bounded real-linear functional. We have $\lambda(x) = \ell(x) - i\ell(ix)$. If we manage to extend ℓ to L without increasing the norm, then $\Lambda(x) = L(x) - iL(ix)$ extends $\lambda(x)$. Moreover, Λ is bounded with norm 1. Indeed, for $x \in X$, let $\alpha = \Lambda(x)/|\Lambda(x)|$; then

$$|\Lambda(x)| = \overline{\alpha}\Lambda(x) = \Lambda(\overline{\alpha}x) = L(\overline{\alpha}x) \leqslant \|\overline{\alpha}x\| = \|x\|.$$

2. *Extension by one real dimension.* Let $z \in X \setminus Y$. Let us extend ℓ onto the real-linear set $Z = \{y + tz \mid y \in Y, t \in \mathbb{R}\}$ without increasing the norm. Since $\ell(y + tz) = \ell(y) + t\ell(z)$, it suffices to define the value of $\ell(z)$. Thus, we denote $p = \ell(z)$ and seek p such that

$$|\ell(y) + tp| \leqslant \|y + tz\|, \quad \forall y \in Y, \quad \forall t \in \mathbb{R}. \tag{7.1}$$

First let us prove an auxiliary inequality. Let $y_1, y_2 \in Y$ and $\alpha > 0$, $\beta > 0$. Using the boundedness of ℓ and the triangle inequality, we get

$$\begin{aligned}\ell(\beta y_1 - \alpha y_2) \leqslant \|\beta y_1 - \alpha y_2\| \leqslant \|\beta y_1 - \beta\alpha z\| + \|\beta\alpha z - \alpha y_2\| \\ = \beta\|y_1 - \alpha z\| + \alpha\|y_2 - \beta z\|.\end{aligned}$$

It follows that

$$\frac{1}{\alpha}(\ell(y_1) - \|y_1 - \alpha z\|) \leqslant \frac{1}{\beta}(\ell(y_2) + \|y_2 - \beta z\|).$$

Since y_1, y_2 and α, β are arbitrary, we get

$$\sup_{\alpha>0, y_1\in Y} \frac{1}{\alpha}(\ell(y_1) - \|y_1 - \alpha z\|) \leqslant \inf_{\beta>0, y_2\in Y} \frac{1}{\beta}(\ell(y_2) + \|y_2 - \beta z\|).$$

Thus, we can find $p \in \mathbb{R}$ such that

$$\frac{1}{\alpha}(\ell(y) - \|y - \alpha z\|) \leqslant p \leqslant \frac{1}{\alpha}(\ell(y) + \|y - \alpha z\|), \quad \forall y \in Y, \quad \forall \alpha > 0. \tag{7.2}$$

Replacing y by $-y$, we see that p also satisfies

$$-\frac{1}{\alpha}(\ell(y) + \|y + \alpha z\|) \leqslant p \leqslant -\frac{1}{\alpha}(\ell(y) - \|y + \alpha z\|), \quad \forall y \in Y, \quad \forall \alpha > 0. \tag{7.3}$$

The inequalities (7.2), (7.3) can be written as

$$\ell(y) - \|y - \alpha z\| \leqslant \alpha p \leqslant \ell(y) + \|y - \alpha z\|, \tag{7.4}$$

$$-\ell(y) - \|y + \alpha z\| \leqslant \alpha p \leqslant -\ell(y) + \|y + \alpha z\|. \tag{7.5}$$

Finally, (7.4) can be rewritten as

$$-\ell(y) - \|y - \alpha z\| \leqslant -\alpha p \leqslant -\ell(y) + \|y - \alpha z\|, \tag{7.6}$$

and (7.5) and (7.6) can be combined together as

$$-\ell(y) - \|y + tz\| \leqslant tp \leqslant -\ell(y) + \|y + tz\|, \quad \forall y \in Y, \quad \forall t \in \mathbb{R}.$$

This is equivalent to (7.1).

3. *Extension by one complex dimension.* Now we can in the same way extend ℓ onto the set $\{y + itz \mid y \in Y, t \in \mathbb{R}\}$ without increasing the norm. Indeed, it suffices to repeat the argument of the previous step of the proof with iz instead of z. Thus, we obtain an extension of ℓ onto $\{y + tz \mid y \in Y, t \in \mathbb{C}\}$ without increasing the norm of ℓ.

4. *Induction argument.* Let x_1, x_2, ..., be a dense set in X. Using this set, we define a sequence of linear sets $Y \subseteq Y_1 \subseteq Y_2 \subseteq \cdots \subseteq X$ such that $x_n \subset X_n$ for all n. The previous step shows how to extend ℓ onto $\tilde{Y} = \cup_n Y_n$ without increasing the norm. By construction, Y contains all points x_1, x_2, ..., and so $\tilde{Y}$ is dense in X. Now we can extend ℓ onto X by continuity and we are done. ∎

7.3 Second Dual space

Corollary 7.9. *For any $x \in X$, there exists $\ell \in X^*$ such that $\|\ell\|_{X^*} = 1$ and $\ell(x) = \|x\|_X$.*

Proof. Let Y be the one dimensional subspace spanned by x and let $\ell(tx) = t\|x\|_X$ for $t \in \mathbb{C}$. Now extend ℓ onto the rest of X, using the Hahn-Banach theorem. ∎

Consider the dual of the dual space X^*, denoted by X^{**}. Every element x generates an element $\lambda_x \in X^{**}$ according to $\lambda_x(\ell) = \ell(x)$, $\ell \in X^*$. Clearly, $|\lambda_x(\ell)| = |\ell(x)| \leqslant \|\ell\|_{X^*}\|x\|_X$, and so $\|\lambda_x\|_{X^{**}} \leqslant \|x\|$. Thus, we have a bounded linear map $J : X \to X^{**}$.

Next, Corollary 7.9 ensures that $\|\lambda_x\|_{X^{**}} = \|x\|$. Indeed, choosing ℓ for x as in Corollary 7.9, we get $\|\ell\| = 1$ and $|\lambda_x(\ell)| = |\ell(x)| = \|x\|_X$, so $\|\lambda_x\|_{X^{**}} \geqslant \|x\|$. Thus, the map $J : X \to X^{**}$ is isometric (i.e. $\|Jx\|_{X^{**}} = \|x\|_X$ but the range of J is not necessarily the whole space X^{**}).

By a slight abuse of terminology, one usually identifies x with λ_x and thus one says that we have a natural isometric embedding $X \subset X^{**}$.

Definition 7.10. A Banach space X is called *reflexive*, if $X = X^{**}$.

In other words, X is reflexive if the range of J is the whole space X^{**}.

Example 7.11. ℓ_p, $1 < p < \infty$ are reflexive. c_0 is not reflexive; in fact, $c_0^* = \ell_1$ (see [Reed Simon, Example 3, Section III.2) and $\ell_1^* = \ell_\infty$ (see Example 7.6). All Hilbert spaces are reflexive (see next section).

7.4 Weak convergence

Let X be a Banach space with the norm $\|\cdot\|_X$.

Definition 7.12. One says that a sequence of elements $x_n \in X$ weakly converges to $x \in X$ if $\ell(x_n) \to \ell(x)$ for any $\ell \in X^*$.

Notation: $x = w - \lim x_n$.

Clearly, if $\|x_n - x\| \to 0$, then $x_n \to x$ weakly. The converse *is false:* $e_n \to 0$ in ℓ_2, but $\|e_n\| = 1$. Here e_n are the standard basis vectors, $e_n = (0, \dots, 0, 1, 0, \dots)$, with 1 on the n'th position.

Theorem 7.13. *Any weakly convergent sequence in a Banach space is bounded.*

The proof uses the following very useful result, which is called the Banach-Steinhaus theorem or the Principle of Uniform Boundedness (for obvious reasons).

Theorem 7.14. *Let X be a Banach space and let $\mathcal{F}$ be a family of bounded linear operators from X into a normed space Y such that the set $\{Tx : T \in \mathcal{F}\}$ is bounded for each $x \in X$. Then the set of norms $\{\|T\| : T \in \mathcal{F}\}$ is bounded.*

The proof is fairly complicated and uses the *Baire's category theorem.* We do not discuss the proof in this course.

Proof of Theorem 7.13. Let $x_n \to x$ weakly. Consider x_n, x as elements of X^{**}. Then for any $\ell \in X^*$, $\ell(x_n) = x_n(\ell)$ converges, and so is bounded. By the Banach-Steinhaus theorem, the norms $\|x_n\|_{X^{**}} = \|x_n\|_X$ are bounded. ∎

Example 7.15. Consider the weak convergence in ℓ_2. We claim that $x^{(k)} \to x$ weakly in ℓ_2 iff

(i) $x_n^{(k)} \to x_n$ for all n;

(ii) the norms $\|x^{(k)}\|$ are bounded.

Indeed, suppose $x^{(k)} \to x$ weakly. Then (i) is true by definition of weak convergence, taking the linear functional $\ell_n(x) = x_n$. (ii) is true by Theorem 7.13.

Next, assume (i), (ii). Let $\ell \in f$ ($f \subset \ell_2$ consists of all elements such that only finitely many coordinates are non-zero). Then $\ell(x^{(k)}) \to \ell(x)$. Now let $y \in \ell_2$ be arbitrary; approximating y by elements of f and using (ii), it is easy to prove that $\ell(x^{(k)}) \to \ell(x)$.

Example 7.16. Consider the weak convergence in $C[a,b]$. Suppose that $f_n \to f$ weakly. Then:

(i) $f_n(x) \to f(x)$ for all $x \in [a,b]$;

(ii) the norms $\|f\|_C$ are bounded.

Indeed, (i) follows from the definition of weak convergence by taking the functional $\ell_x(f) = f(x)$. (ii) follows from Theorem 7.13. In fact, one can prove that (i), (ii) imply weak convergence, but we will not prove this, as this requires description of the space dual to $C[a,b]$, which involves some measure theory.

Example 7.17. Consider some examples of weak convergence in L^2:

1. $f_n(x) = \sin(nx)$ converges weakly to zero in $L^2(-1,1)$.
2. Let $\varphi \in L^2(\mathbb{R})$. Then $f_n(x) = \sqrt{n}\varphi(nx)$ converges weakly to zero in $L^2(-1,1)$.
3. Let $\varphi \in L^2(\mathbb{R})$. Then $f_n(x) = \varphi(x-n)$ converges weakly to zero in $L^2(\mathbb{R})$.

7.5 Exercises

Exercise 7.18. Let X be a normed space and Y be a Banach space. Prove that $\mathcal{B}(X,Y)$ is a Banach space (i.e. you need to prove completeness of $\mathcal{B}(X,Y)$ with respect to the operator norm).

Exercise 7.19. Prove that any linear functional on $\mathbb{R}^n$ (with the usual Euclidean norm) is bounded. Describe all such functionals.

Exercise 7.20. Let X be a reflexive Banach space. Using Corollary 7.9, prove that for any $\ell \in X^*$ there is an element $x \in X$ such that $\|x\|_X = 1$ and $\ell(x) = \|\ell\|_{X^*}$.

Exercise 7.21. For $x \in \ell^\infty$, let $f(x) = \sum_{n=1}^\infty x_n 2^{-n}$. Determine the norm of f in $(\ell^\infty)^*$; in $(\ell^1)^*$; in any $(\ell^p)^*$, $1 < p < \infty$.

Exercise 7.22. Let $\ell : C[-1,1] \to \mathbb{C}$ be defined by

$$\ell(f) = \int_{-1}^{0} f(x)dx - \int_{0}^{1} f(x)dx.$$

Prove that ℓ is a bounded linear functional on $C[-1,1]$ and determine its norm $\|\ell\|_{C^*}$. Prove that there is no element $f \in C[-1,1]$ such that $\|f\|_C = 1$ and $\ell(f) = \|\ell\|_{C^*}$.

Compare this with the result of the previous problem; what can be said about the space $C[-1,1]$?

Exercise 7.23. Using the Hahn-Banach Theorem, prove that there exists a linear functional λ on ℓ^∞ with the following property. If $x \in \ell^\infty$ is such that the limit $x_\infty = \lim_{n\to\infty} x_n$ exists, then $\lambda(x) = x_\infty$. That is, λ extends the definition of the limit to all bounded sequences.

Exercise 7.24. Suppose that $\ell_n \to \ell$ in X^* weakly. Prove that $\ell_n \to \ell$ $*$-weakly.

Exercise 7.25. Let X be a reflexive Banach space. Prove that $\ell_n \to \ell$ in X^* weakly if and only if $\ell_n \to \ell$ $*$-weakly.

Exercise 7.26. Let X be a Banach space and let $T_n : X \to X$ be a sequence of bounded linear operators such that for all $x \in X$, the limit $\lim_{n\to\infty} T_n x$ exists in the norm of X. Using the principle of uniform boundedness, prove that there exists a bounded operator $T : X \to X$ such that $T_n \to T$ strongly as $n \to \infty$.