Exercises on metric and Hilbert spaces
An invitation to functional analysis

Alexandru Ghitza*
School of Mathematics and Statistics
University of Melbourne

Version of Fri 18th Oct, 2024 at 11:21

*(aghitza@alum.mit.edu)

Contents

1. Introduction

The next few exercises are about countability/uncountability. See Section 1.2 for clarification on our use of the term “countable”. You may assume without proof that any subset of a countable set is finite or countable.

Exercise 1.1. Let $f\colon X \longrightarrow Y$ be a function, with X a countable set. Then $\operatorname{im}(f)$ is finite or countable.

[*Hint*: Reduce to the case $f\colon \mathbf{N} \longrightarrow Y$ is surjective; construct a right inverse $g\colon Y \longrightarrow \mathbf{N}$, which has to be injective, of f.]

Solution. Without loss of generality, we may assume that f is surjective and we want to show that Y is finite or countable.

Also without loss of generality (by pre-composing f with any bijection $\mathbf{N} \longrightarrow X$), we may assume that $f\colon \mathbf{N} \longrightarrow Y$ is surjective.

As $f\colon \mathbf{N} \longrightarrow Y$ is surjective, there exists a right inverse $g\colon Y \longrightarrow \mathbf{N}$, in other words $f \circ g\colon Y \longrightarrow Y$ is the identity function id_Y: given $y \in Y$, the pre-image $f^{-1}(y) \subseteq \mathbf{N}$ is nonempty, so it has a smallest element n_y; we let $g(y) = n_y$. For any $y \in Y$, we have $f(g(y)) = f(n_y) = y$ as $n_y \in f^{-1}(y)$. So $f \circ g = \mathrm{id}_Y$.

In particular, this forces $g\colon Y \longrightarrow \mathbf{N}$ to be injective, hence realising Y as a subset of the countable set $\mathbf{N}$. We conclude that Y is finite or countable. □

Exercise 1.2. Show that the union S of any countable collection of countable sets is a countable set.

[*Hint*: Construct a surjective function $\mathbf{N} \times \mathbf{N} \longrightarrow S$.]

Solution. Write

$$S = \bigcup_{n \in \mathbf{N}} S_n,$$

with each S_n a countable set. It is clear that S is infinite (as, say, S_1 is, and $S_1 \subseteq S$).

For each $n \in \mathbf{N}$, fix a bijection $\varphi_n\colon \mathbf{N} \longrightarrow S_n$. (As Chengjing rightfully points out to me, this uses the Axiom of Countable Choice.) Define a function $\psi\colon \mathbf{N} \times \mathbf{N} \longrightarrow S$ by:

$$\psi\big((n,m)\big) = \varphi_n(m) \in S_n \subseteq S.$$

This is surjective, and $\mathbf{N} \times \mathbf{N}$ is countable, so S is finite or countable, and we ruled out finite above. □

Exercise 1.3. Let $\mathbf{R}^{\mathbf{N}}$ be the set of arbitrary sequences $(x_1, x_2, \dots)$ of elements of $\mathbf{R}$.

This is a vector space under the naturally-defined addition of sequences and multiplication by a scalar.

Let $e_j \in \mathbf{R}^{\mathbf{N}}$ be the sequence whose j-th entry is 1, and all the others are 0. Describe the subspace $\operatorname{Span}\{e_1, e_2, \dots\}$ of $\mathbf{R}^{\mathbf{N}}$. Is the set $\{e_1, e_2, \dots\}$ a basis of $\mathbf{R}^{\mathbf{N}}$?

Solution. Let $S = \{e_1, e_2, \dots\}$ and $W = \text{Span}(S)$.

For each $n \in \mathbf{N}$, define

$$W_n = \text{Span}\{e_1, e_2, \dots, e_n\} \subseteq W.$$

I claim that

$$W = \bigcup_{n \in \mathbf{N}} W_n.$$

One inclusion is clear, as $W_n \subseteq W$ for all $n \in \mathbf{N}$.

For the other inclusion, let $w \in W$. Then there exist $m \in \mathbf{N}$, $a_1, \dots, a_m \in \mathbf{R}$ and $k_1, \dots, k_m \in \mathbf{N}$ such that

$$w = a_1 e_{k_1} + \dots + a_m e_{k_m}.$$

Set $n = \max\{k_1, \dots, k_m\}$, then $w \in W_n$.

Is $W = \mathbf{R}^{\mathbf{N}}$? No. Any $w \in W$ appears in a W_n for some $n \in \mathbf{N}$, therefore only the first n entries of w can be nonzero. This means, for instance, that $v = (1, 1, 1, \dots) \notin W$. So S does not span $\mathbf{R}^{\mathbf{N}}$. □

Exercise 1.4. (*) Let $V = \mathbf{R}$ viewed as a vector space over $\mathbf{Q}$.

Let $\alpha \in \mathbf{R}$. Show that the set $T = \{\alpha^n : n \in \mathbf{N}\}$ is $\mathbf{Q}$-linearly independent if and only if α is transcendental.

(Note: An element $\alpha \in \mathbf{R}$ is called *algebraic* if there exists a monic polynomial $f \in \mathbf{Q}[x]$ such that $f(\alpha) = 0$. An element $\alpha \in \mathbf{R}$ is called *transcendental* if it is not algebraic.)

Solution. This is a straightforward rewriting of the definition of algebraic: α is algebraic if and only if it satisfies a polynomial equation with coefficients in $\mathbf{Q}$, which is equivalent to a nontrivial linear relation between the powers of α, which exists if and only if T is linearly dependent. □

Exercise 1.5. (*) Let W be a $\mathbf{Q}$-vector space with a countable basis B. Show that W is a countable set.

[*Hint*: Use Exercise 1.2.]

Conclude that $\mathbf{R}$ does not have a countable basis as a vector space over $\mathbf{Q}$.

Solution. Since B is countable we can enumerate it as $B = \{b_n : n \in \mathbf{N}\}$. For each $n \in \mathbf{N}$, let $W_n = \text{Span}\{b_1, \dots, b_n\}$. Then for each $n \in \mathbf{N}$, W_n is isomorphic (as a $\mathbf{Q}$-vector space) to $\mathbf{Q}^n$, hence W_n is countable. I claim that

$$W = \bigcup_{n \in \mathbf{N}} W_n.$$

One inclusion is obvious, as $W_n \subseteq W$ for all $n \in \mathbf{N}$. For the other direction, let $w \in W = \text{Span}(B)$, so there exist $n \in \mathbf{N}$, $a_1, \dots, a_n \in \mathbf{Q}$ and $k_1, \dots, k_n \in \mathbf{N}$ such that

$$w = a_1 b_{k_1} + \dots + a_n b_{k_n}.$$

Let $k = \max\{k_1, \dots, k_n\}$, then $w \in W_k$.

So W is a countable union of countable sets, hence countable by Exercise 1.2.

The last claim follows directly from the fact that $\mathbf{R}$ is an uncountable set. □

We now turn to posets, Zorn's Lemma, and the existence of bases.

A **p**artially **o**rdered **set** (*poset* for short) is a set X together with a *partial order* $\leqslant$, that is a relation satisfying

- $x \leqslant x$ for all $x \in X$;
- if $x \leqslant y$ and $y \leqslant x$ then $x = y$;
- if $x \leqslant y$ and $y \leqslant z$ then $x \leqslant z$.

A poset X such that for any $x, y \in X$ we have $x \leqslant y$ or $y \leqslant x$ is called a *totally ordered set*, and $\leqslant$ is called a *total order*.

Exercise 1.6. (*) Fix a set Ω and let X be the set of all subsets of Ω. Check that $\subseteq$ is a partial order on X. It is not a total order if Ω has at least two distinct elements.

Solution. The fact that $\subseteq$ is a partial order follows directly from known properties of set inclusion.

If Ω has at least two distinct elements x_1 and x_2, then $\{x_1\}$ and $\{x_2\}$ are not comparable under $\subseteq$, so the latter is not a total order. □

A *chain* in a poset $(X, \leqslant)$ is a subset $C \subseteq X$ that is totally ordered with respect to $\leqslant$.

If $S \subseteq X$ is a subset of a poset, then an *upper bound* for S is an element $u \in X$ such that $s \leqslant u$ for all $s \in S$.

A *maximal element* of a poset X is an element m of X such that there does not exist any $x \in X$ such that $x \neq m$ and $m \leqslant x$. In other words, for any $x \in X$, either $x = m$, or $x \leqslant m$, or x and m are not comparable with respect to the partial order $\leqslant$.

Exercise 1.7. (*) Let $(X, \leqslant)$ be a nonempty finite poset. (This just means that X is a nonempty finite set with a partial order $\leqslant$.) Prove that X has a maximal element.

[*Hint*: You could, for instance, use induction on the number of elements of X.]

Solution. We proceed by induction on n, the cardinality of X.

Base case: if $n = 1$ then $X = \{x\}$ for a single element x. Then trivially x is a maximal element of X.

For the induction step, fix $n \in \mathbf{N}$ and suppose that any poset of cardinality n has a maximal element. Let X be a poset of cardinality $n+1$ and choose an arbitrary element $x \in X$. Let $Y = X \setminus \{x\}$, then Y is a poset of cardinality n so by the induction hypothesis has a maximal element m_Y, and clearly $m_Y \neq x$.

We have two possibilities now:

- If $m_Y \leqslant x$, then x is a maximal element of X. Why? Suppose that x is not maximal in X, so that there exists $z \in X$ such that $z \neq x$ and $x \leqslant z$. Since $z \neq x$, we must have $z \in Y$. If $z = m_Y$, then $z \leqslant x$ and $x \leqslant z$ so $z = x$, contradiction. So $z \neq m_Y$, and $m_Y \leqslant x$ and $x \leqslant z$, so $m_Y \leqslant z$, contradicting the maximality of m_Y in Y.
- Otherwise, (if it is not true that $m_Y \leqslant x$), m_Y is a maximal element of X. Why? Suppose there exists $z \in X$ such that $z \neq m_Y$ and $m_Y \leqslant z$. Since $m_Y \leqslant x$ is not true, we have $z \neq x$, so $z \in Y$, contradicting the maximality of m_Y in Y.

In either case we found a maximal element for X.

An alternative approach is to proceed by contradiction: suppose $(X, \leqslant)$ is a nonempty finite poset that does not have a maximal element. Use this to construct an unbounded chain of elements of X, contradicting finiteness. □

Zorn's Lemma (Lemma 1.3) is used to deduce the existence of maximal elements in infinte posets.

Exercise 1.8. (*) Prove Theorem 1.2: any vector space V has a basis.

[*Hint*: Let X be the set of all linearly independent subsets of V, partially ordered by inclusion. Prove that X has a maximal element B, and prove that this must also span V.]

Solution. If $V = \{0\}$, then $\varnothing$ is vacuously a (in fact, the only) basis of V.

Suppose $V \neq \{0\}$. If $v \in V \setminus \{0\}$, then $\{v\}$ is a linearly independent subset of V. Let X be the set of all linearly independent subsets of V, then X is nonempty. We consider the partial order $\subseteq$ on X given by inclusion of subsets.

Let C be a nonempty chain in X and define

$$U = \bigcup_{S \in C} S,$$

then clearly $S \subseteq U$ for all $S \in C$, so we'll know that U is an upper bound for C as soon as we can show that it is linearly independent (so that $U \in X$).

Suppose there exist $n \in \mathbf{N}$, $a_1, \ldots, a_n \in \mathbf{F}$, and $u_1, \ldots, u_n \in U$ such that

$$a_1 u_1 + \cdots + a_n u_n = 0. \tag{1.1}$$

Let $J = \{1, \ldots, n\}$. For each $j \in J$, there exists $S_j \in C$ such that $u_j \in S_j$. As C is totally ordered, there exists $i \in J$ such that $S_j \subseteq S_i$ for all $j \in J$. But this means that $u_1, \ldots, u_n \in S_i$, so that the linear relation of Equation (1.1) takes place in the linearly independent set S_i. Therefore $a_1 = \cdots = a_n = 0$.

We conclude that X satisfies the conditions of Zorn's Lemma, hence it has a maximal element B. I claim that B spans V, so that it is a basis of V.

We prove this last claim by contradiction: if $v \in V \setminus \operatorname{Span}(B)$, then $B' := B \cup \{v\}$ is linearly independent, hence an element of X. But $B \subseteq B'$ and $B \neq B'$, contradicting the maximality of B. □

2. Metric and topological spaces

Exercise 2.1. Let (X, d) be a metric space. Show that

$$|d(x,y) - d(t,y)| \leqslant d(x,t)$$

for all $x, y, t \in X$.

Solution. We need to show that

$$-d(x,t) \leqslant d(x,y) - d(t,y) \leqslant d(x,t).$$

One application of the triangle inequality gives

$$d(x,y) \leqslant d(x,t) + d(t,y) \qquad \Rightarrow \qquad d(x,y) - d(t,y) \leqslant d(x,t).$$

Another application gives

$$d(t,y) \leqslant d(t,x) + d(x,y) \qquad \Rightarrow \qquad -d(x,t) \leqslant d(x,y) - d(t,y).$$ □

Exercise 2.2. Let (X, d) be a metric space. Show that

$$|d(x,y) - d(s,t)| \leqslant d(x,s) + d(y,t)$$

for all $x, s, y, t \in X$.

Solution. We have

$$\begin{aligned} |d(x,y) - d(s,t)| &= |d(x,y) - d(y,s) + d(y,s) - d(s,t)| \\ &\leqslant |d(x,y) - d(y,s)| + |d(y,s) - d(s,t)| \\ &\leqslant d(x,s) + d(y,t) \end{aligned}$$

after one application of the triangle inequality and two applications of Exercise 2.1. □

Exercise 2.3. (*) Fix a prime p and consider the metric space $(\mathbf{Q}, d_p)$ where d_p is the p-adic metric from Example 2.1.

(a) Let $p = 3$ and write down 4 elements of $\mathbf{B}_1(2)$ and 4 elements of $\mathbf{B}_{1/9}(3)$.

(b) Back to general prime p now: show that every triangle is isosceles. In other words, given three points in $\mathbf{Q}$, at least two of the three resulting (p-adic) distances are equal.

(c) Show that every point of an open ball is a centre. In other words, take an open ball $\mathbf{B}_r(c)$ with $r \in \mathbf{R}_{\geqslant 0}$ and $c \in \mathbf{Q}$ and suppose $x \in \mathbf{B}_r(c)$; prove that $\mathbf{B}_r(c) = \mathbf{B}_r(x)$.

(d) Show that given any two open balls, either one of them is contained in the other, or they are completely disjoint.

Solution. (a) We have

$$\left\{2, 5, -7, \frac{4}{5}\right\} \subseteq \mathbf{B}_1(2)$$
$$\left\{3, 30, -24, \frac{39}{4}\right\} \subseteq \mathbf{B}_{1/9}(3).$$

(b) Recall that in the proof of the triangle inequality for the p-adic metric in Example 2.1, the following stronger result was shown:

$$d_p(x, y) \leqslant \max\{d_p(x, t), d_p(t, y)\}.$$

with equality holding if $d_p(x,t) \neq d_p(t,y)$. But this precisely says that if $d_p(x,t) \neq d_p(t,y)$, then $d_p(x,y)$ has to be equal to the largest of $d_p(x,t)$ and $d_p(t,y)$.

(c) First $x \in \mathbf{B}_r(c)$ iff $c \in \mathbf{B}_r(x)$ (this is true for any metric space). So it suffices to show that $x \in \mathbf{B}_r(c)$ implies $\mathbf{B}_r(x) \subseteq \mathbf{B}_r(c)$. Let $y \in \mathbf{B}_r(x)$, then $d_p(y,x) < r$, so that

$$d_p(y, c) \leqslant \max\left\{d_p(y, x), d_p(x, c)\right\} < r,$$

in other words $y \in \mathbf{B}_r(c)$.

(d) Consider two open balls $\mathbf{B}_r(x)$ and $\mathbf{B}_t(y)$. Without loss of generality $r \leqslant t$. Suppose that the balls are not disjoint and let $z \in \mathbf{B}_r(x) \cap \mathbf{B}_t(y)$. By part (c) this implies that $\mathbf{B}_r(z) = \mathbf{B}_r(x)$ and $\mathbf{B}_t(z) = \mathbf{B}_t(y)$, so that

$$\mathbf{B}_r(x) = \mathbf{B}_r(z) \subseteq \mathbf{B}_t(z) = \mathbf{B}_t(y).$$ □

Exercise 2.4. Let $n \in \mathbf{N}$, $X = \mathbf{R}^n$ with the dot product $\cdot$, $\|x\| = \sqrt{x \cdot x}$ for $x \in X$, and $d(x,y) = \|x - y\|$ for $x, y \in X$. Then (X, d) is a metric space. (The function d is called the *Euclidean metric* or *ℓ^2 metric* on $\mathbf{R}^n$.)

[*Hint*: The Cauchy–Schwarz inequality can be useful for checking the triangle inequality.]

Solution. We have

(a) $d(x,y) = \|x - y\| = \sqrt{(x-y)\cdot(x-y)} = \sqrt{(-1)^2\,(y-x)\cdot(y-x)} = \|y - x\| = d(y,x)$;

(b) Let $u = x - t$ and $v = t - y$, then we are looking to show that $\|u + v\| \leqslant \|u\| + \|v\|$. But:

$$\begin{aligned}\|u+v\|^2 &= (u+v)\cdot(u+v) = \|u\|^2 + 2\,u\cdot v + \|v\|^2 \leqslant \|u\|^2 + 2\,|u\cdot v| + \|v\|^2 \\ &\leqslant \|u\|^2 + 2\,\|u\|\,\|v\| + \|v\|^2 = \big(\|u\| + \|v\|\big)^2,\end{aligned}$$

where the last inequality sign comes from the Cauchy–Schwarz inequality.

(c) $d(x,y) = 0$ iff $(x-y)\cdot(x-y) = 0$ iff $x - y = 0$ iff $x = y$. □

Exercise 2.5. Draw the unit open balls in the metric spaces $(\mathbf{R}^2, d_1)$ (Example 2.4), $(\mathbf{R}^2, d_2)$ (Exercise 2.4), and $(\mathbf{R}^2, d_\infty)$ (Example 2.5).

Solution. The Manhattan unit open ball is the interior of the square with vertices $(1,0)$, $(0,-1)$, $(-1,0)$, and $(0,1)$.

The Euclidean unit open ball is the interior of the unit circle centred at $(0,0)$.

The sup metric unit open ball is the interior of the square with vertices $(1,1)$, $(1,-1)$, $(-1,-1)$, and $(-1,1)$. □

Exercise 2.6. Let X be a nonempty set and define

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

Prove that (X,d) is a metric space. (The function d is called the *discrete metric* on X.)

Solution. It is clear from the definition that $d(y,x) = d(x,y)$ and that $d(x,y) = 0$ iff $x = y$.

For the triangle inequality, take $x, y, t \in X$ and consider the different cases:

$x = y$	$x = t$	$t = y$	$d(x,y)$	$d(x,t) + d(t,y)$
True	True	True	0	$0 + 0 = 0$
True	False	False	0	$1 + 1 = 2$
False	True	False	1	$1 + 0 = 1$
False	False	True	1	$0 + 1 = 1$
False	False	False	1	$1 + 1 = 2$

In all cases we see that $d(x,y) \leqslant d(x,t) + d(t,y)$. □

Exercise 2.7. Let $n \in \mathbf{N}$, $X = \mathbf{F}_2^n$, and let $d(x,y)$ be the number of indices $i \in \{1,\dots,n\}$ such that $x_i \neq y_i$. Prove that (X,d) is a metric space. (The function d is called the *Hamming metric*.)

Solution. Do this from scratch if you want to, but I prefer to deduce it from other examples we have seen.

First look at the case $n = 1$, $X = \mathbf{F}_2$. Then $d(x,y)$ is precisely the discrete metric on $\mathbf{F}_2$ (see Exercise 2.6), in particular it is a metric. I'll denote it $d_{\mathbf{F}_2}$ for a moment to minimise confusion.

Back in the arbitrary $n \in \mathbf{N}$ case, note that $d(x,y)$ defined above can be expressed as

$$d(x,y) = d_{\mathbf{F}_2}(x_1,y_1) + \cdots + d_{\mathbf{F}_2}(x_n,y_n),$$

which is a special case of Example 2.4, therefore also a metric. □

Exercise 2.8. Let (X,d) be a metric space and let $A \subseteq X$.

(a) Prove that the set A is open if and only if it is the union of a collection of open balls.

(b) Conclude that the set of all open balls in X generates the metric topology of X.

Solution. (a) In one direction, if A is a union of a collection of open balls, then A is open by Example 2.10 and Proposition 2.11.

In the other direction, suppose A is open. Let $a \in A$, then there exists an open ball $\mathbf{B}_{r(a)}(a) \subseteq A$. Then

$$A = \bigcup_{a \in A} \mathbf{B}_{r(a)}(a).$$

(b) Follows immediately from the definition of the topology generated by a set. □

Exercise 2.9. Let Y be a subset of a metric space (X, d) and consider the induced metric on Y.

(a) Prove that for any $y \in Y$ and any $r \in \mathbf{R}_{\geqslant 0}$ we have

$$\mathbf{B}_r^Y(y) = \mathbf{B}_r^X(y) \cap Y,$$

where $\mathbf{B}_r^X(y)$ is the open ball of radius r centred at y in X, and $\mathbf{B}_r^Y(y)$ is the open ball of radius r centred at y in Y.

(b) Let $A \subseteq Y$. Prove that A is an open set in Y if and only if there exists an open set U in X such that $A = U \cap Y$.

Solution. (a) We have

$$\mathbf{B}_r^X(y) = \{x \in X : d(x, y) < r\}$$
$$\mathbf{B}_r^Y(y) = \{x \in Y : d(x, y) < r\},$$

so that

$$\mathbf{B}_r^X(y) \cap Y = \{x \in X : d(x, y) < r\} \cap Y = \{x \in Y : d(x, y) < r\} = \mathbf{B}_r^Y(y).$$

(b) In one direction, suppose A is open in Y; by Exercise 2.8 we have some indexing set I such that

$$A = \bigcup_{i \in I} \mathbf{B}_{r_i}^Y(a_i),$$

with $r_i > 0$ and $a_i \in A$ for all $i \in I$. We can then let

$$U = \bigcup_{i \in I} \mathbf{B}_{r_i}^X(a_i),$$

which by Exercise 2.8 is an open in X. It is clear that $A = U \cap Y$ by part (a).

Conversely, suppose $A = U \cap Y$ with U open in X. Let $a \in A$, then $a \in U$ so there exists an open (in X) ball $\mathbf{B}_r^X(a)$ such that $\mathbf{B}_r^X(a) \subseteq U$. Consider $\mathbf{B}_r^Y(a) = \mathbf{B}_r^X(a) \cap Y \subseteq U \cap Y = A$. So every point $a \in A$ is contained in an open (in Y) ball, hence A is open in Y. □

Exercise 2.10. Prove that any closed ball is a closed set.

Solution. This is a variation on Example 2.10 and a generalisation of Example 2.9 (which is the case $r = 0$).

Consider $C = \mathbf{D}_r(x)$ with $x \in X$, $r \in \mathbf{R}_{\geqslant 0}$. Let $y \in X \smallsetminus C$, then $d(x,y) > r$. Set $t = d(x,y) - r$ and consider the open ball $\mathbf{B}_t(y)$.

I claim that $\mathbf{B}_t(y) \subseteq \big(X \smallsetminus C\big)$: if $w \in \mathbf{B}_t(y)$ then $d(w,y) < t$ so

$$d(x,y) \leqslant d(x,w) + d(w,y) \leqslant d(x,w) + t \qquad \Rightarrow \qquad d(x,w) \geqslant d(x,y) - t = r,$$

hence $w \notin C$. □

Exercise 2.11. **(*)** Show that any p-adic open ball in $\mathbf{Q}$ is both an open set and a closed set.

Solution. Any open ball in any metric space is an open set (Example 2.10). Let's show that an arbitrary p-adic open ball $\mathbf{B}_r(c)$ is closed.

Let $U = \mathbf{Q} \smallsetminus \mathbf{B}_r(c)$. Given $u \in U$, we have $|u - c|_p \geqslant r$.

I claim that $\mathbf{B}_r(u) \subseteq U$, which would imply that U is open, so that $\mathbf{B}_r(c)$ is closed.

Suppose, on the contrary, that there exists $t \in \mathbf{B}_r(u) \cap \mathbf{B}_r(c)$. Then $|u - t|_p < r$ and $|t - c|_p < r$, so that

$$|u - c|_p = |(u-t) + (t-c)|_p \leqslant \max\{|u-t|_p, |t-c|_p\} < r,$$

contradicting the fact that $|u - c|_p \geqslant r$. □

Exercise 2.12. Let (X, d) be a metric space and define

$$d'(x,y) = \frac{d(x,y)}{1 + d(x,y)}.$$

Prove that (X, d') is a metric space.

[*Hint*: Before tackling the triangle inequality, show that if $a, b, c \in \mathbf{R}_{\geqslant 0}$ satisfy $c \leqslant a + b$, then $\frac{c}{1+c} \leqslant \frac{a}{1+a} + \frac{b}{1+b}$.]

Solution. It is clear from the definition that $d'(x,y) = d'(y,x)$ and that $d'(x,y) = 0$ iff $d(x,y) = 0$ iff $x = y$.

For the triangle inequality, apply the inequality in the hint with $c = d(x,y)$, $a = d(x,t)$, $b = d(t,y)$. □

Exercise 2.13.

(a) Let $f\colon X \longrightarrow Y$ be a function between two sets X and Y, and let $S \subseteq Y$. Prove that

$$f^{-1}(S) = X \smallsetminus f^{-1}(Y \smallsetminus S).$$

(b) Let $f\colon X \longrightarrow Y$ be a function between topological spaces. Prove that f is continuous if and only if: for any closed subset $C \subseteq Y$, the inverse image $f^{-1}(C) \subseteq X$ is a closed subset.

Solution.

(a) We have $x \in f^{-1}(S)$ iff $f(x) \in S$ iff $f(x) \notin \big(Y \smallsetminus S\big)$ iff $x \notin f^{-1}(Y \smallsetminus S)$.

(b) Suppose f is continuous and $C \subseteq Y$ is closed. By part (a) we have

$$f^{-1}(C) = X \setminus f^{-1}(Y \setminus C).$$

Then $(Y \setminus C) \subseteq Y$ is open and f is continuous, so $f^{-1}(Y \setminus C) \subseteq X$ is open, therefore $f^{-1}(C)$ is closed.

Conversely, suppose the inverse image of any closed subset is closed. Let $V \subseteq Y$ be open, then by part (a) we have

$$f^{-1}(V) = X \setminus f^{-1}(Y \setminus V).$$

So $(Y \setminus V) \subseteq Y$ is closed, so $f^{-1}(Y \setminus V) \subseteq X$ is closed, hence $f^{-1}(V)$ is open. We conclude that f is continuous. □

Exercise 2.14. This is a variation on Tutorial Question 2.7.

Let $f\colon X \longrightarrow Y$ be a function and $\mathcal{T}_X$ a topology on X. Define

$$\mathcal{T}_Y = \{U \in \mathcal{P}(Y) : f^{-1}(U) \in \mathcal{T}_X\}.$$

(a) Prove that $\mathcal{T}_Y$ is the finest topology on Y such that f is continuous. (This topology is called the *final topology* induced by f.)

(b) Let $\mathcal{T}$ be another topology on Y. Prove that $f\colon (X, \mathcal{T}_X) \longrightarrow (Y, \mathcal{T})$ is continuous if and only if $\mathcal{T}$ is coarser than $\mathcal{T}_Y$.

(c) Use an example to prove that $\mathcal{T}_Y$ need not be metrisable even when $\mathcal{T}_X$ is a metric topology.

(d) Give an example in which $\mathcal{T}_Y$ is metrisable but $\mathcal{T}_X$ is not.

[*Hint*: For (c) and (d), consider using Tutorial Question 2.3.]

Solution.

(a) We start with proving that $\mathcal{T}_Y$ is a topology:

- Since $\emptyset = f^{-1}(\emptyset)$ and $X = f^{-1}(Y)$, it follows that $\mathcal{T}_Y$ contains $\emptyset$ and Y.
- If $\{U_i : i \in I\}$ is a collection of members of $\mathcal{T}_Y$, then

$$\bigcup_{i \in I} f^{-1}(U_i) = f^{-1}\Big(\bigcup_{i \in I} U_i\Big) \in \mathcal{T}_X.$$

- If $U_1, \ldots, U_n$ are members of $\mathcal{T}_Y$, then

$$\bigcap_{i=1}^{n} f^{-1}(U_i) = f^{-1}\Big(\bigcap_{i=1}^{n} U_i\Big) \in \mathcal{T}_X.$$

If $\mathcal{T}$ is a topology on Y such that f is continuous, then $f^{-1}(U) \in \mathcal{T}_X$ for every member U of $\mathcal{T}$, so $\mathcal{T} \subseteq \mathcal{T}_Y$. Therefore, $\mathcal{T}_Y$ is the finest topology such that f is continuous.

(b) The 'only if' part has been proven in part (a), so it suffices to prove the 'if' part.

Suppose $\mathcal{T}$ is coarser than $\mathcal{T}_Y$. If U is a member of $\mathcal{T}$, then $U \in \mathcal{T}_Y$, which implies that $f^{-1}(U)$ is open in X. It follows that f is continuous when the topology on Y is $\mathcal{T}$.

(c) Let $(X, \mathcal{T}_X)$ be the set of real numbers equipped with the Euclidean topology. Put $Y = \{0, 1\}$. If $f\colon X \longrightarrow Y$ is defined by

$$f(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{otherwise,} \end{cases}$$

then $\mathcal{T}_Y = \big\{\varnothing, \{1\}, \{0,1\}\big\}$. The topology $\mathcal{T}_X$ is defined by the Euclidean metric, but $\mathcal{T}_Y$ is not metrisable (see Tutorial Question 2.3).

(d) Put $X = \{0, 1\}$, $Y = \{1\}$, $\mathcal{T}_X = \big\{\varnothing, \{1\}, \{0,1\}\big\}$. Let $f\colon X \longrightarrow Y$ be the function sending both 0 and 1 to 1. It follows that $\mathcal{T}_Y = \{\varnothing, \{0,1\}\}$. The topology $\mathcal{T}_Y$ is defined by the discrete metric (see Tutorial Question 2.1), but $\mathcal{T}_X$ is not metrisable (see Tutorial Question 2.3). □

Exercise 2.15. Let X be a topological space and $U \subseteq X$ a subset. Prove that U is open in X if and only if: for all $u \in U$, there exists an open neighbourhood V_u of u such that $V_u \subseteq U$.

Solution. One direction is obvious: if U is open in X, then given any $u \in U$ we can take $V_u = U$ as an open neighbourhood contained in U.

In the other direction, suppose U has the given property at every $u \in U$. Then

$$U = \bigcup_{u \in U} V_u,$$

therefore U is open, since it is the union of the collection $\{V_u : u \in U\}$ of open sets. □

Exercise 2.16. Prove Proposition 2.21:

Let X be a set and $\mathcal{T}_1, \mathcal{T}_2$ two topologies on X. The following statements are equivalent:

(a) $\mathcal{T}_2$ is coarser than $\mathcal{T}_1$ (that is, $\mathcal{T}_2 \subseteq \mathcal{T}_1$);

(b) for any $x \in X$ and any $\mathcal{T}_2$-open neighbourhood U_x^2 of x, there exists a $\mathcal{T}_1$-open neighbourhood U_x^1 of x such that $U_x^1 \subseteq U_x^2$;

(c) the function $f : (X, \mathcal{T}_1) \longrightarrow (X, \mathcal{T}_2)$ given by $f(x) = x$ is continuous.

Solution. **(a)⇔(c):** Since $f^{-1}(S) = S$ for any subset S of X, we have:

($\mathcal{T}_2$ is coarser then $\mathcal{T}_1$) if and only if (if $U \in \mathcal{T}_2$ then $U \in \mathcal{T}_1$) if and only if (if $U \in \mathcal{T}_2$ then $f^{-1}(U) \in \mathcal{T}_1$) if and only if ($f$ is continuous).

(a)⇒(b): trivial, since if $x \in U_x^2$ and $U_x^2 \in \mathcal{T}_2 \subseteq \mathcal{T}_1$, we can take $U_x^1 = U_x^2$ and we are done.

(b)⇒(a): Let $U \in \mathcal{T}_2$. We use Exercise 2.15 to prove that $U \in \mathcal{T}_1$. Let $x \in U$, then setting $U_x^2 = U$ we have that U_x^2 is a $\mathcal{T}_2$-open neighbourhood of x, so by (b) there exists a

cT_1-open neighbourhood U_x^1 of x such that $U_x^1 \subseteq U$. By Exercise 2.15 we conclude that U is open in the topology $\mathcal{T}_1$. □

Exercise 2.17. Generalise Example 2.9 to the setting of Hausdorff topological spaces; in other words, prove that if X is a Hausdorff topological space, then any singleton $\{x\} \subseteq X$ is a closed set.

Solution. Let $U = X \setminus \{x\}$ and let $u \in U$. Then $u \neq x$, so by the Hausdorff property of X, there exist open neighbourhoods V_1 of u and V_2 of x such that $V_1 \cap V_2 = \varnothing$. In particular, $x \notin V_1$, so $V_1 \subseteq U$. As we have exhibited an open neighbourhood contained in U around every element of U, we conclude by Exercise 2.15 that U is open, so its complement $\{x\}$ is closed. □

Exercise 2.18. Show that the union of any finite collection of closed sets is closed. Show that the intersection of any arbitrary collection of closed sets is closed.

Solution. Let $n \in \mathbf{N}$ and let $C_1, \dots, C_n$ be closed subsets of X. Let

$$C = \bigcup_{i=1}^{n} C_i,$$

then the complement of C is

$$X \setminus C = X \setminus \left(\bigcup_{i=1}^{n} C_i \right) = \bigcap_{i=1}^{n} (X \setminus C_i) .$$

For each $i = 1, \dots, n$, C_i is closed so $X \setminus C_i$ is open, therefore $X \setminus C$ is the intersection of finitely many open sets, hence is itself open by the topology axioms. We conclude that C is closed.

For the second statement, let $\{C_i : i \in I\}$ be a collection of closed subsets of X, indexed by a set I. Let

$$C = \bigcap_{i \in I} C_i,$$

then the complement of C is

$$X \setminus C = X \setminus \left(\bigcap_{i \in I} C_i \right) = \bigcup_{i \in I} (X \setminus C_i) .$$

For each $i \in I$, C_i is closed so $X \setminus C_i$ is open, hence $X \setminus C$ is the union of a collection of open sets, so is itself open by the topology axioms. We conclude that C is closed. □

Exercise 2.19. Prove Proposition 2.27: A subset D of a topological space X is disconnected if and only if there exist open subsets $U, V \subseteq X$ such that

$$D \subseteq U \cup V, \qquad D \cap U \cap V = \varnothing, \qquad D \cap U \neq \varnothing, \qquad D \cap V \neq \varnothing.$$

Solution. By definition D is a disconnected subset of X if and only if it is a disconnected topological space in the induced topology. The latter is by definition: there exist U', V'

open subsets of D such that

$$D = U' \cup V', \qquad U' \cap V' = \varnothing, \qquad U' \neq \varnothing, \qquad V' \neq \varnothing.$$

But U', V' are open in D if and only if there exist open subsets U, V of X such that $U' = U \cap D$, $V' = V \cap D$, from which the claim follows. □

Exercise 2.20. Let X be a topological space and let $\{y\}$ be a one-point topological space. Prove that $X \times \{y\}$ (with the product topology) is homeomorphic to X.

Solution. Let $f\colon X \times \{y\} \longrightarrow X$ be the map $f(x,y) = x$ and let $g\colon X \longrightarrow X \times \{y\}$ be the map $g(x) = (x,y)$. It is clear that g is the inverse of f. Since f is simply the projection onto the first factor of the product, it is continuous by Proposition 2.19. To show that g is continuous, consider a rectangle in $X \times \{y\}$: this is either $\varnothing$ or $U \times \{y\}$ for some open set $U \subseteq X$. Then $g^{-1}(U \times \{y\}) = U$ is open in X. □

Exercise 2.21. Let X and Y be topological spaces, where the topology on Y is the trivial topology. Prove that every function from X to Y is continuous.

Solution. Let $f\colon X \longrightarrow Y$ be a function. The only open subsets of Y are $\varnothing$ and Y. Since $f^{-1}(\varnothing) = \varnothing$ and $f^{-1}(Y) = X$, it follows that f is continuous. □

Exercise 2.22. Prove that every constant function between topological spaces is continuous.

Solution. Let X and Y be topological spaces. Pick a point y in Y and define $f\colon X \longrightarrow Y$ to be the constant function sending every element of X to y. If U is an open subset of Y, then

$$f^{-1}(U) = \begin{cases} X & \text{if } y \in U, \\ \varnothing & \text{otherwise.} \end{cases}$$

Hence $f^{-1}(U)$ is open. □

Exercise 2.23. Let X be a topological space and let S be a subset of X. Prove that the inclusion $\iota\colon S \longrightarrow X$ defined by $\iota(x) = x$ is continuous when S is given the subspace topology induced from X.

Conclude that the identity function $\mathrm{id}_X\colon X \longrightarrow X$ is continuous.

Solution. If U is an open subset of X, then $\iota^{-1}(U) = U \cap S$, which is open in S by the definition of the subspace topology. Hence ι is continuous.

The identity function is the special case $S = X$. □

Exercise 2.24. A subset $D \subseteq X$ of a topological space X is dense in X if and only if $D \cap U \neq \varnothing$ for all nonempty open sets U in X.

Solution. Suppose D is dense, so $\overline{D} = X$, and let U be nonempty open. If $D \cap U = \varnothing$ then $D \subseteq X \setminus U$. But $X \setminus U$ is a closed subset of X containing D, so by the minimality property of $\overline{D}$ we have $\overline{D} \subseteq X \setminus U$. As $U \neq \varnothing$, this means $\overline{D} \neq X$, contradiction.

Conversely, suppose $D \cap U$ is nonempty for any nonempty open U. If $\overline{D} \neq X$ then $U := X \setminus \overline{D}$ is a nonempty open subset of X, so $D \cap (X \setminus \overline{D}) \neq \varnothing$. But this is absurd since $D \subseteq \overline{D}$. □

Exercise 2.25. Let X be a topological space. The intersection of two dense open sets U_1 and U_2 is dense and open.

Solution. Let $U_{12} = U_1 \cap U_2$. We know already that U_{12} is open.

To show that U_{12} is dense, we use Exercise 2.24 and show that $U_{12} \cap U \neq \varnothing$ for all nonempty open U:

$$U_{12} \cap U = \big(U_1 \cap U_2\big) \cap U = \mathrm{U}_1 \cap \big(U_2 \cap U\big).$$

Since U_2 is dense and open, $U_2 \cap U$ is nonempty and open. Since U_1 is dense, $U_1 \cap \big(U_2 \cap U\big)$ is nonempty. So $U_{12} \cap U \neq \varnothing$, hence U_{12} is dense. □

Exercise 2.26. (*) Give explicit continuous surjective functions $f\colon \mathbf{R} \longrightarrow I$, where I is:

(a) $\mathbf{R}$ (b) $(0,\infty)$ (c) $(-\infty,0)$ (d) $(-\infty,0]$ (e) $[-1,1]$
(f) $(0,1]$ (g) $[0,1)$ (h) $(-\pi/2,\pi/2)$ (i) $\{0\}$.

[*Hint*: Draw some functions you know from calculus and see what their ranges are.]

Solution. These are of course not the only possible answers (well, except for the last one).

(a) $x \longmapsto x$;

(b) $x \longmapsto e^x$;

(c) $x \longmapsto -e^x$;

(d) $x \longmapsto -x^2$;

(e) $x \longmapsto \sin(x)$;

(f) $x \longmapsto \min\{e^x, 1\}$;

(g) $x \longmapsto \max\{-e^x, -1\} + 1$;

(h) $x \longmapsto \arctan(x)$;

(i) $x \longmapsto 0$. □

Exercise 2.27. Let A be a subset of a topological space X. Prove that

$$X \setminus A^\circ = \overline{X \setminus A}.$$

Solution. Since $A^\circ \subseteq A$, we have $\big(X \setminus A\big) \subseteq \big(X \setminus A^\circ\big)$. But A° is open, so $X \setminus A^\circ$ is a closed set containing $X \setminus A$, hence

$$\overline{X \setminus A} \subseteq X \setminus A^\circ.$$

For the opposite inclusion, note that $(X \setminus A) \subseteq \overline{X \setminus A}$, so

$$X \setminus \overline{X \setminus A} \subseteq X \setminus (X \setminus A) = A,$$

therefore $X \setminus \overline{X \setminus A}$ is an open set contained in A, so that

$$X \setminus \overline{X \setminus A} \subseteq A^\circ,$$

which implies that $X \setminus A^\circ \subseteq \overline{X \setminus A}$. □

Exercise 2.28. (*)

(a) Show that a topological group G is Hausdorff if and only if $\{e\}$ is a closed subset of G.

(b) Show that if G is a Hausdorff topological group then its centre Z is a closed subgroup.

(c) Show that if $f: G \longrightarrow H$ is a continuous group homomorphism and H is Hausdorff, then $\ker(f)$ is a closed normal subgroup of G.

Solution.

(a) By Exercise 2.17, if G is Hausdorff then the singleton $\{e\}$ is closed.

Conversely, suppose $\{e\}$ is a closed subset of G. Consider the map $f: G \times G \longrightarrow G$ given by $f(g,h) = g^{-1}h$, then f is continuous and

$$f^{-1}(e) = \{(g,g) : g \in G\} = \Delta(G)$$

(see Tutorial Question 3.9). Since f is continuous and $\{e\}$ is closed, $\Delta(G)$ is closed in $G \times G$, so by Tutorial Question 3.9, G is Hausdorff.

(b) We have

$$Z = \{g \in G : gxg^{-1}x^{-1} = e \text{ for all } x \in G\} = \bigcap_{x \in G} \{g \in G : gxg^{-1}x^{-1} = e\}$$

which is an intersection of closed sets, since each of the sets is the inverse image of $\{e\}$ under the continuous map $g \longmapsto gxg^{-1}x^{-1}$.

(c) The assertion is immediate from $\ker(f) = f^{-1}(e)$. □

Exercise 2.29. Let (X,d) be a metric spaces. Prove that

$$(x_n) \sim (y_n) \qquad \text{if } \big(d(x_n,y_n)\big) \longrightarrow 0 \text{ as } n \longrightarrow \infty$$

defines an equivalence relation on the set of sequences in X.

Solution. The reflexivity $(x_n) \sim (x_n)$ and symmetry $(x_n) \sim (y_n) \iff (y_n) \sim (x_n)$ are very clear. For the transitivity, suppose $(x_n) \sim (y_n)$ and $(y_n) \sim (z_n)$. Let $\varepsilon > 0$. There exists $N_1 \in \mathbf{N}$ such that $d(x_n,y_n) < \varepsilon/2$ for all $n \geqslant N_1$. There exists $N_2 \in \mathbf{N}$ such that $d(y_n,z_n) < \varepsilon/2$ for all $n \geqslant N_2$. Letting $N = \max\{N_1,N_2\}$ we have (by the triangle inequality)

$$d(x_n, z_n) \leqslant d(x_n, y_n) + d(y_n, z_n) < \varepsilon \quad \text{for all } n \geqslant N.$$

So $(x_n) \sim (z_n)$. □

Exercise 2.30. Let X be a topological space. Suppose $\{C_n : n \in \mathbf{N}\}$ is a countable collection of connected subsets of X such that $C_n \cap C_{n+1} \neq \varnothing$ for all $n \in \mathbf{N}$. Then

$$\bigcup_{n\in\mathbf{N}} C_n$$

is a connected subset of X.

Solution. Let $f: \bigcup_{n\in\mathbf{N}} C_n \longrightarrow \{0,1\}$ be a continuous function, where $\{0,1\}$ is given the discrete topology. Pick an element x_0 of C_0. We use induction to prove that $f(C_n) = \{f(x_0)\}$ for every natural number n.

The base case when $n = 0$ follows from the connectedness of C_0 and Proposition 2.29.

For the induction step, suppose the statement is true for a natural number n and consider an element x of C_{n+1}. Since $C_n \cap C_{n+1} \neq \varnothing$, we can pick an element x' of $C_n \cap C_{n+1}$. By the induction hypothesis, we have $f(x') = f(x_0)$. It then follows from the connectedness of C_{n+1} and Proposition 2.29 that $f(x) = f(x') = f(x_0)$.

Hence f is constant, which implies that $\bigcup_{n\in\mathbf{N}} C_n$ is connected. □

Exercise 2.31. Give $\mathbf{N} \subseteq \mathbf{R}$ the subspace topology. Let X be a topological space and (x_n) a sequence in X. Prove that (x_n) is a continuous function $\mathbf{N} \longrightarrow X$.

Solution. First note that the subspace topology on $\mathbf{N} \subseteq \mathbf{R}$ is the discrete topology: for any $n \in \mathbf{N}$, we have $\{n\} = (n-1, n+1) \cap \mathbf{N}$, so $\{n\}$ is open in $\mathbf{N}$. Therefore every subset of $\mathbf{N}$ is open, hence every function $\mathbf{N} \longrightarrow X$ is continuous. □

Exercise 2.32. Any sequence has at most one limit.

Solution. Suppose x and x' are two limits of a sequence (x_n). For any $\varepsilon > 0$, there exist $N, N' \in \mathbf{N}$ such that

$$x_n \in \mathbf{B}_{\varepsilon/2}(x) \qquad \text{for all } n \geqslant N \quad \text{and} \quad x_n \in \mathbf{B}_{\varepsilon/2}(x') \qquad \text{for all } n \geqslant N'.$$

Therefore, for $n = \max\{N, N'\}$ we have $x_n \in \mathbf{B}_{\varepsilon/2}(x) \cap \mathbf{B}_{\varepsilon/2}(x')$, which (via the triangle inequality) implies that $d(x, x') < \varepsilon$.

Since this holds for all $\varepsilon > 0$, we conclude that $d(x, x') = 0$ so that $x = x'$. □

Exercise 2.33. Show that any distance-preserving function $f: X \longrightarrow Y$ is continuous.

In particular, any isometry is a homeomorphism.

Solution. Let $x \in X$. Given $\varepsilon > 0$, if $x' \in \mathbf{B}_\varepsilon(x)$ then $d_X(x, x') < \varepsilon$, so

$$d_Y\big(f(x), f(x')\big) = d_X(x, x') < \varepsilon,$$

hence $f(x') \in \mathbf{B}_\varepsilon(f(x))$. □

Exercise 2.34. A map $f\colon X \longrightarrow Y$ between topological spaces is said to be *open* if for every open subset $U \subseteq X$, the image $f(U) \subseteq Y$ is an open subset.

(a) Show that an open continuous bijective map $f\colon X \longrightarrow Y$ is a homeomorphism.

(b) Suppose S generates the topology on X and let S' denote the set of all finite intersections of elements of S. Show that f is open if and only if $f(U) \subseteq Y$ is an open subset for all $U \in S'$.

(Compare this to Tutorial Question 2.6. Where is the difference coming from?)

(c) Show that the projection maps $\pi_1\colon X_1 \times X_2 \longrightarrow X_1$ and $\pi_2\colon X_1 \times X_2 \longrightarrow X_2$ are open maps.

Solution.

(a) We need to check that $f^{-1}\colon Y \longrightarrow X$ is continuous; let $U \subseteq X$ be open, then $\left(f^{-1}\right)^{-1}(U) = f(U)$ is open in Y since f is an open map.

(b) One direction is trivial. For the other direction, we are told that every open subset U of X is of the form

$$U = \bigcup_{i\in I} U_i, \qquad U_1 \in S'.$$

Then

$$f(U) = \bigcup_{i\in I} f(U_i).$$

By assumption each $f(U_i)$ is open in Y, so their union must also be an open subset.

(c) By part (b) and Example 2.18, we only need to check the open condition on open rectangles $U_1 \times U_2 \subseteq X_1 \times X_2$: we have $\pi_1(U_1 \times U_2) = U_1$, clearly open in X_1. Same for π_2. □

Exercise 2.35. Give $\mathbf{Q} \subseteq \mathbf{R}$ the induced metric and consider the sequence (x_n) defined recursively by

$$x_1 = 1, \qquad x_{n+1} = \frac{x_n}{2} + \frac{1}{x_n} \quad \text{for } n \in \mathbf{N}.$$

(a) Prove that $1 \leqslant x_n \leqslant 2$ for all $n \in \mathbf{N}$ and breathe a sigh of relief that the recursive definition does not accidentally divide by 0.

(b) For $n \in \mathbf{N}$, let $y_n = x_{n+1} - x_n$. Prove that

$$y_{n+1} = -\frac{y_n^2}{2x_{n+1}} \qquad \text{for all } n \in \mathbf{N}.$$

(c) Prove that

$$|y_n| \leqslant \frac{1}{2^n} \qquad \text{for all } n \in \mathbf{N}.$$

(d) Show that (x_n) is Cauchy.

(e) Show that (x_n) converges to $\sqrt{2}$ in $\mathbf{R}$, and conclude that $\mathbf{Q}$ is not complete.

Solution.

(a) Induction on n. Base case $x_1 = 1$ clear.

Fix $n \in \mathbf{N}$ and suppose $1 \leqslant x_n \leqslant 2$. Then

$$\frac{1}{2} \leqslant \frac{x_n}{2} \leqslant 1 \qquad \text{and} \qquad \frac{1}{2} \leqslant \frac{1}{x_n} \leqslant 1,$$

so $1 \leqslant x_{n+1} \leqslant 2$.

(b) Fix $n \in \mathbf{N}$. Noting that $2x_n x_{n+1} = x_n^2 + 2$, we have

$$y_n^2 = \left(x_{n+1} - x_n\right)^2 = x_{n+1}^2 - 2x_{n+1}x_n + x_n^2 = x_{n+1}^2 - 2$$
$$2x_{n+1}y_{n+1} = 2x_{n+1}\left(\frac{1}{x_{n+1}} - \frac{x_{n+1}}{2}\right) = 2 - x_{n+1}^2 = -y_n^2.$$

(c) From part (b) we have

$$|y_{n+1}| = \frac{|y_n|^2}{2x_{n+1}} \qquad \text{for all } n \in \mathbf{N}.$$

We can use this, part (a), and induction by n.

For the base case we have $y_1 = \frac{1}{2}$.

For the induction step, fix $n \in \mathbf{N}$ and suppose $|y_n| \leqslant \frac{1}{2^n}$, then

$$|y_{n+1}| = \frac{|y_n|^2}{2x_{n+1}} \leqslant \frac{|y_n|^2}{2} \leqslant \frac{1}{2^{2n+1}} \leqslant \frac{1}{2^{n+1}}.$$

(d) Let $\varepsilon > 0$ and let $N \in \mathbf{N}$ be such that $2^{N-1} > 1/\varepsilon$. If $n \geqslant m \geqslant N$ then

$$\begin{aligned}
|x_n - x_m| &= |y_{n-1} + y_{n-2} + \cdots + y_m| \\
&\leqslant |y_{n-1}| + \cdots + |y_m| \\
&\leqslant \frac{1}{2^{n-1}} + \cdots + \frac{1}{2^m} \\
&= \left(\frac{1}{2^{n-m-1}} + \frac{1}{2^{n-m-2}} + \cdots + 1\right)\frac{1}{2^m} \\
&\leqslant \frac{2}{2^m} \leqslant \frac{1}{2^N} < \varepsilon.
\end{aligned}$$

Here we used the fact that the geometric series with ratio $1/2$ sums up to 2.

(e) Thinking of (x_n) as a sequence in $\mathbf{R}$, it converges to some limit $x \in \mathbf{R}$ by the completeness of $\mathbf{R}$. We can therefore take limits as $n \longrightarrow \infty$ on both sides of the defining relation

$$x_{n+1} = \frac{x_n}{2} + \frac{1}{x_n} \quad \text{for } n \in \mathbf{N}$$

to get

$$x = \frac{x}{2} + \frac{1}{x} \Rightarrow x^2 = 2.$$

Throwing in the fact that $x \geqslant 1$, we conclude that $x = \sqrt{2}$.

The conclusion that $\mathbf{Q}$ is not complete now follows from the fact that $\sqrt{2} \notin \mathbf{Q}$. □

Exercise 2.36. Let X be a complete metric space and let $S \subseteq X$. Then the closure $\overline{S}$ (with the metric induced from $\overline{S} \subseteq X$) is a completion of S (with the metric induced from $S \subseteq X$).

Solution. Of course, $\overline{S}$ is complete: if (x_n) is a Cauchy sequence in $\overline{S}$, then it is a Cauchy sequence in X, so $(x_n) \longrightarrow x \in X$ since X is complete. But $\overline{S}$ is closed, so $(x_n) \longrightarrow x \in \overline{S}$.

We let $\iota\colon S \longrightarrow \overline{S}$ be the inclusion map: $\iota(s) = s$ for all $s \in S$. It is injective and distance-preserving (as d_S and $d_{\overline{S}}$ are both induced from d_X).

Finally, S is dense in $\overline{S}$: by Proposition 2.50, for every $x \in \overline{S}$ there exists a sequence (s_n) in S such that $(s_n) \longrightarrow x$. □

Exercise 2.37. Let (x_n) be a sequence in a metric space X, let $\varphi\colon \mathbf{N} \longrightarrow \mathbf{N}$ be an injective function, and consider the sequence $(y_n) = \big(x_{\varphi(n)}\big)$ in X. Prove that if (x_n) converges to x, then so does (y_n).

Does the converse hold?

Solution. Suppose $(x_n) \longrightarrow x$. Given $\varepsilon > 0$, let $N \in \mathbf{N}$ be such that $x_n \in \mathbf{B}_\varepsilon(x)$ for all $n \geqslant N$.

Since $\varphi\colon \mathbf{N} \longrightarrow \mathbf{N}$ is injective, the inverse image $\varphi^{-1}\big(\{1, \dots, N-1\}\big)$ is a finite set, so it has a maximal element M. (If the set is empty, just take $M = 0$.) For all $n \geqslant M+1$, we have $\varphi(n) \geqslant N$, so $y_n = x_{\varphi(n)} \in \mathbf{B}_\varepsilon(x)$.

The converse certainly does not hold. For instance, take $(x_n) = (1, 0, 1, 0, 1, 0, \dots)$ and $\varphi(n) = 2n$, then the sequence $(y_n) = (0, 0, 0, \dots)$ converges to 0 but (x_n) does not converge. □

Exercise 2.38. Show that if $f\colon X \longrightarrow Y$ is a continuous map between topological spaces and $A \subseteq X$ then $f\big(\overline{A}\big) \subseteq \overline{f(A)}$.

Solution. Let $x \in \overline{A}$, let $y = f(x)$, and suppose that $y \notin \overline{f(A)}$. Then there exists an open neighbourhood $V \subseteq \big(Y \smallsetminus f(A)\big)$ with $y \in V$. As f is continuous, there exists an open neighbourhood $U \subseteq X$ of x with $f(U) \subseteq V$; as V does not intersect $f(A)$, we get that U does not intersect A, contradicting the fact that $x \in \overline{A}$. □

Exercise 2.39. Let X be a topological space. We say that a collection of closed subsets of X has the *finite intersection property* if every finite subcollection has nonempty intersection.

Prove that X is compact if and only if every collection of closed sets with the finite intersection property has nonempty intersection.

Solution. Suppose X is compact and $\{C_i\colon i \in I\}$ is a collection of closed sets with the finite intersection property. Suppose that

$$\bigcap_{i \in I} C_i = \varnothing.$$

Then

$$X = \bigcup_{i \in I} U_i, \qquad \text{where } U_i \coloneqq X \smallsetminus C_i,$$

is an open covering of X. Since X is compact, there exists a finite subset $J \subseteq I$ such that

$$X = \bigcup_{j \in J} U_j,$$

which implies that

$$\bigcap_{j \in J} C_j = \varnothing,$$

contradicting the finite intersection property of the collection $\{C_i \colon i \in I\}$.

Conversely, suppose every collection of closed sets of X with the finite intersection property has nonempty intersection. Suppose that X is not compact, so there exists an open cover of X:

$$X = \bigcup_{i \in I} U_i$$

with no finite subcover.

For each $i \in I$, let $C_i = X \smallsetminus U_i$. Then for every finite $J \subseteq I$, $\{U_i \colon i \in J\}$ is not a cover of X, which means that the collection $\{C_i \colon i \in J\}$ has nonempty intersection. Hence the collection $\{C_i \colon i \in I\}$ has the finite intersection property, but note that the collection itself has empty intersection, since $\{U_i \colon i \in I\}$ is a cover of X, so we have reached a contradiction. □

Exercise 2.40. Check (directly from the definition of uniform continuity) that $f \colon \mathbf{R}_{>0} \longrightarrow \mathbf{R}_{>0}$ given by $f(x) = \frac{1}{x}$ is not uniformly continuous.

Solution. First make sure that you negate the condition in the definition correctly: there exists $\varepsilon > 0$ such that for all $\delta > 0$ there exist x, x' such that $x' \in \mathbf{B}_\delta(x)$ and $f(x') \notin \mathbf{B}_\varepsilon\big(f(x)\big)$.

And now, to work: let $\varepsilon = 1$. Take an arbitrary $\delta > 0$. Set $x = \min\{\delta, 1\}$. I claim that $x' \coloneqq x/2$ satisfies the desired condition. Let's check:

$$|x - x'| = \frac{x}{2} \leqslant \frac{\delta}{2} < \delta,$$

so indeed $x' \in \mathbf{B}_\delta(x)$.

Also

$$|f(x) - f(x')| = \left|\frac{1}{x} - \frac{1}{x'}\right| = \left|\frac{1}{x} - \frac{2}{x}\right| = \frac{1}{x} \geqslant 1 = \varepsilon,$$

so indeed $f(x') \notin \mathbf{B}_\varepsilon\big(f(x)\big)$. □

Exercise 2.41. Let $f \colon X \longrightarrow Y$ be a uniformly continuous function between two metric spaces and suppose $(x_n) \sim (x'_n)$ are equivalent sequences in X. Prove that $\big(f(x_n)\big) \sim \big(f(x'_n)\big)$ as sequences in Y.

Does the conclusion hold if f is only assumed to be continuous?

Solution. Let $\varepsilon > 0$. As f is uniformly continuous, there exists $\delta > 0$ such that for all $x, x' \in X$, if $d_X(x, x') < \delta$ then $d_Y(f(x), f(x')) < \varepsilon$. As $(x_n) \sim (x'_n)$, there exists $N \in \mathbf{N}$ such that $d_X(x_n, x'_n) < \delta$ for all $n \geqslant N$. Hence for all $n \geqslant N$ we have $d_Y(f(x_n), f(x'_n)) < \varepsilon$.

The result does not hold in general for continuous functions; for instance one can take $f\colon \mathbf{R}_{>0} \longrightarrow \mathbf{R}_{>0}$ given by $f(x) = \frac{1}{x}$, and $(1/n) \sim (1/n^2)$ but $(f(1/n)) = (n)$, $(f(1/n^2)) = (n^2)$ and $(n) \not\sim (n^2)$. □

Exercise 2.42. In the context of the proof of Theorem 2.62, show that if $(x_n) \sim (x'_n)$ and $(y_n) \sim (y'_n)$, then

$$\lim_{n \longrightarrow \infty} d(x'_n, y'_n) = \lim_{n \longrightarrow \infty} d(x_n, y_n).$$

Solution. This uses the same approach as Proposition 2.56: we have

$$\left| d(x'_n, y'_n) - d(x_n, y_n) \right| \leqslant d(x'_n, x_n) + d(y'_n, y_n).$$

But by assumption the two distances on the RHS can be made arbitrarily small, so we conclude that $d(x'_n, y'_n)$ and $d(x_n, y_n)$ can be made arbitrarily close, hence they have the same limit.

(This explanation shouldn't keep you from writing a more rigorous proof.) □

Exercise 2.43. Let $X = \mathbf{R}_{>0}$, $Y = \mathbf{R}$, $f\colon X \longrightarrow Y$ given by $f(x) = \frac{1}{x}$. For $\widehat{X} = \mathbf{R}_{\geqslant 0}$ and $\widehat{Y} = Y = \mathbf{R}$, prove that there is no continuous function $\widehat{f}\colon \widehat{X} \longrightarrow \widehat{Y}$ such that $\widehat{f}|_X = f$.

Solution. Suppose that a continuous extension $\widehat{f}\colon \mathbf{R}_{\geqslant 0} \longrightarrow \mathbf{R}_{\geqslant 0}$ exists. Consider the sequence $(x_n) = \left(\frac{1}{n}\right) \longrightarrow 0 \in \mathbf{R}_{\geqslant 0}$. By continuity of $\widehat{f}$ we must have

$$\widehat{f}(0) = \widehat{f}\left(\lim_{n \longrightarrow \infty} \frac{1}{n}\right) = \lim_{n \longrightarrow \infty} \widehat{f}\left(\frac{1}{n}\right) = \lim_{n \longrightarrow \infty} f\left(\frac{1}{n}\right) = \lim_{n \longrightarrow \infty} n.$$

But the rightmost limit does not exist (in $\mathbf{R}_{\geqslant 0}$), contradiction. □

Exercise 2.44. Prove that any contraction is uniformly continuous.

Solution. Suppose $f\colon X \longrightarrow Y$ is a contraction with constant C.

Let $\varepsilon > 0$ and set $\delta = \frac{\varepsilon}{C+1}$, then for all $x_1, x_2 \in X$ such that $d_X(x_1, x_2) < \delta$, we have

$$d_Y(f(x_1), f(x_2)) \leqslant C d_X(x_1, x_2) \leqslant C\delta = \frac{C}{C+1}\varepsilon < \varepsilon.$$ □

Exercise 2.45. Show that a subset $S \subseteq X$ is bounded if and only if $S \subseteq \mathbf{D}_r(x)$ for some $r \geqslant 0$ and some $x \in X$.

Solution. If $S \subseteq \mathbf{D}_r(x)$ then $\operatorname{diam}(S) \leqslant \operatorname{diam}\left(\mathbf{D}_r(x)\right) = 2r$ so S is bounded.

Conversely, suppose S is bounded and let $r = \operatorname{diam}(S)$. Let $x \in S$ be any point, then $d(x, y) \leqslant r$ for all $y \in S$, so that $S \subseteq \mathbf{D}_r(x)$. □

Exercise 2.46. Let (X, d) be a metric space and let A, B be bounded sets. Then $A \cup B$ is bounded.

Solution. Let $a \in A$, $b \in B$, and $r = d(a, b)$. I claim that the diameter of $A \cup B$ is at most $\operatorname{diam}(A) + r + \operatorname{diam}(B)$. If $x, y \in A \cup B$ then

$$d(x,y) \leqslant \begin{cases} \operatorname{diam}(A) & \text{if } x, y \in A \\ \operatorname{diam}(B) & \text{if } x, y \in B \\ d(x,a) + d(a,b) + d(b,y) \leqslant \operatorname{diam}(A) + r + \operatorname{diam}(B) & \text{if } x \in A, y \in B \\ d(y,a) + d(a,b) + d(b,x) \leqslant \operatorname{diam}(A) + r + \operatorname{diam}(B) & \text{if } x \in B, y \in A. \end{cases}$$ □

Exercise 2.47. In any metric space (X, d), any totally bounded set S is bounded.

Solution. Take $\varepsilon = 1$ and let $B_1, \dots, B_N$ be a cover of S by open balls of radius 1. Each B_n is bounded, so by Exercise 2.46 the finite union $B_1 \cup \cdots \cup B_N$ is bounded, hence so is its subset S. □

Exercise 2.48. Prove that a function $f \colon X \longrightarrow Y$ between metric spaces is bounded if and only if $f(X)$ is a bounded subset of Y.

Solution. The function f is bounded if and only if there exist $y \in Y$, $M \in \mathbf{R}$ be such that

$$d_Y\big(y, f(x)\big) \leqslant M \qquad \text{for all } x \in X.$$

On the other hand, this is equivalent to saying that $f(X) \subseteq \mathbf{D}_M(y)$, so by Exercise 2.45 equivalent to $f(X)$ being a bounded subset of Y. □

Exercise 2.49. Given metric spaces X, Y, prove that a sequence (f_n) in $B(X, Y)$ converges uniformly to $f \in B(X, Y)$ if and only if $(f_n) \longrightarrow f$ with respect to the uniform metric d_∞ on $B(X, Y)$.

Solution. Suppose (f_n) converges uniformly to f. Given $\varepsilon > 0$, there exists $N \in \mathbf{N}$ such that for all $n \geqslant N$ we have

$$d_Y(f_n(x), f(x)) < \frac{\varepsilon}{2} \qquad \text{for all } x \in X.$$

So for all $n \geqslant N$ we have

$$d_\infty(f_n, f) = \sup_{x \in X}\{d_Y(f_n(x), f(x))\} \leqslant \frac{\varepsilon}{2} < \varepsilon,$$

in other words $(f_n) \longrightarrow f$ w.r.t. d_∞.

Conversely, suppose $(f_n) \longrightarrow f$. Given $\varepsilon > 0$, there exists $N \in \mathbf{N}$ such that for all $n \geqslant N$ we have

$$\sup_{x \in X}\{d_Y(f_n(x), f(x))\} = d_\infty(f_n, f) < \varepsilon,$$

hence for all $n \geqslant N$

$$d_Y(f_n(x), f(x)) < \varepsilon \qquad \text{for all } x \in X,$$

in other words (f_n) converges uniformly to f. □

Exercise 2.50. (*) Let $f\colon G \longrightarrow H$ be a group homomorphism between topological groups. Prove that the following are equivalent:

(a) f is continuous;

(b) f is continuous at some element of G;

(c) f is continuous at the identity element e_G of G.

Solution. In this proof, we will keep using the following fact: if U is a neighbourhood of some element g of G, and if g' is another element of G, then $g'U$ is a neighbourhood of $g'g$. This follows from the equation $g'U = L_{g'^{-1}}(U)$ and the continuity of $L_{g'^{-1}}$ (see Proposition 2.44).

(a)⇒(b): This follows from Tutorial Question 3.3.

(b)⇒(c): Suppose f is continuous at some element g of G. Since f is a group homomorphism, $f(e_G) = e_H$. If U is a neighbourhood of e_H, then $f(g)U$ is a neighbourhood of g, so $f^{-1}\big(f(g)U\big)$ is a neighbourhood of g. Since

$$x \in f^{-1}(U) \iff f(x) \in U \iff f(gx) \in f(g)U \iff gx \in f^{-1}\big(f(g)U\big),$$

it follows that $f^{-1}(U) = g^{-1}f^{-1}\big(f(g)U\big)$, so $f^{-1}(U)$ is a neighbourhood of e_G.

(c)⇒(a): Using similar arguments as in the proof for (b)⇒(c), we can prove that continuity at e_G implies continuity at every element of G. Hence f is continuous by Tutorial Question 3.3. □

Exercise 2.51. (*)

(a) Let V be a $\mathbf{Q}$-vector space. Prove that every group homomorphism $f\colon \mathbf{Q} \longrightarrow V$ is $\mathbf{Q}$-linear.

(b) What can you say (and prove) about **continuous** group homomorphisms $\mathbf{R} \longrightarrow \mathbf{R}$?

(c) Suppose that a group homomorphism $f\colon \mathbf{R} \longrightarrow \mathbf{R}$ is continuous at some real number. Prove that f is continuous on $\mathbf{R}$, and conclude that f is $\mathbf{R}$-linear.

(d) Let B be a basis for $\mathbf{R}$ as a $\mathbf{Q}$-vector space. (Recall from Exercise 1.4 that B is uncountable.) Use two distinct irrational elements of B to construct a $\mathbf{Q}$-linear transformation $f\colon \mathbf{R} \longrightarrow \mathbf{R}$ that is not $\mathbf{R}$-linear.

If you would (and why wouldn't you?), follow the rabbit:

`https://en.wikipedia.org/wiki/Cauchy%27s_functional_equation`

Solution.

(a) Let $v = f(1) \in V$.

For $n \in \mathbf{N}$ we have

$$f(n) = f(1 + 1 + \cdots + 1) = f(1) + \cdots + f(1) = nv.$$

For $m \in \mathbf{N}$ we have

$$v = f(1) = f\left(\frac{1}{m} + \cdots + \frac{1}{m}\right) = mf\left(\frac{1}{m}\right),$$

so $f(1/m) = (1/m)\,v$.

Therefore, for any $n, m \in \mathbf{N}$ we have

$$f\left(\frac{n}{m}\right) = nf\left(\frac{1}{m}\right) = \frac{n}{m}\,v.$$

Combining this with $f(-a) = -f(a)$ and $f(0) = 0$, we conclude that $f(x) = xv = xf(1)$ for all $x \in \mathbf{Q}$.

(b) Let $f\colon \mathbf{R} \longrightarrow \mathbf{R}$ be additive. Let $g\colon \mathbf{Q} \longrightarrow \mathbf{R}$ be the restriction of f to $\mathbf{Q} \subseteq \mathbf{R}$. Let $a = g(1) = f(1)$.

By part (b), $g(q) = q\,g(1) = qa$ for all $q \in \mathbf{Q}$. Let $x \in \mathbf{R}$. As $\mathbf{Q}$ is dense in $\mathbf{R}$, there is some sequence $(q_n) \longrightarrow x$ with $q_n \in \mathbf{Q}$; since f is continuous we have

$$f(x) = f\left(\lim_{n\longrightarrow\infty} q_n\right) = \lim_{n\longrightarrow\infty} f(q_n) = \lim_{n\longrightarrow\infty} g(q_n) = \lim_{n\longrightarrow\infty} (q_n a) = xa = xf(1).$$

Hence f is $\mathbf{R}$-linear.

(c) It follows from Exercise 2.50 that f is continuous, so by part (c) f is $\mathbf{R}$-linear.

(d) Let B be a $\mathbf{Q}$-basis for $\mathbf{R}$. Exactly one element of B is a nonzero rational, and without loss of generality we may assume it is 1. Since B is uncountable, it also contains uncountably many irrationals. Let $b, c \in B \cap (\mathbf{R} \smallsetminus \mathbf{Q})$. Consider the bijective function $\sigma\colon B \longrightarrow B$ given by

$$\sigma(b) = c, \qquad \sigma(c) = b, \qquad \sigma(x) = x \text{ for all } x \in B \smallsetminus \{b, c\}.$$

Since B is a $\mathbf{Q}$-basis of $\mathbf{R}$, σ extends by $\mathbf{Q}$-linearity to a $\mathbf{Q}$-linear transformation $f\colon \mathbf{R} \longrightarrow \mathbf{R}$. In particular, f is additive.

Suppose that f is $\mathbf{R}$-linear, then:

$$c = f(b) = bf(1) = b\,1 = b,$$

contradicting the fact that $b \neq c$. □

Exercise 2.52. If (X, d_X) and (Y, d_Y) are two metric spaces, a metric d on $X \times Y$ is said to be *conserving* if

$$d_\infty\big((x_1, y_1), (x_2, y_2)\big) \leqslant d\big((x_1, y_1), (x_2, y_2)\big) \leqslant d_1\big((x_1, y_1), (x_2, y_2)\big)$$

for all $(x_1, y_1), (x_2, y_2) \in X \times Y$.

(For the definitions of d_1 and d_∞, see Examples 2.4 and 2.5.)

Prove that any conserving metric d defines the product topology on $X \times Y$. (In particular, all conserving metrics on $X \times Y$ are equivalent.)

Solution. Let $\mathcal{T}$ denote the product topology on $X \times Y$ and $\mathcal{T}_d$ the topology defined by the metric d.

We start by proving that any open rectangle $U \times V \in \mathcal{T}$ is also open in $\mathcal{T}_d$, which will imply that $\mathcal{T} \subseteq \mathcal{T}_d$. Consider an arbitrary element $(u,v) \in U \times V$. Since U is open in X, there exists $s > 0$ such that $\mathbf{B}_s(u) \subseteq U$. Similarly, there exists $t > 0$ such that $\mathbf{B}_t(v) \subseteq V$. Let $r = \min\{s,t\} > 0$. I claim that the d-open ball $B := \mathbf{B}_r\big((u,v)\big) \subseteq U \times V$. Why? If $(x,y) \in B$ then since d is conserving,

$$\max\left\{d_X(x,u), d_Y(y,v)\right\} = d_\infty\big((x,y),(u,v)\big) \leqslant d\big((x,y),(u,v)\big) < r,$$

so $d_X(x,u) < r \leqslant s$ hence $x \in U$, and $d_Y(y,v) < r \leqslant t$ hence $y \in V$.

Now we prove that any d-open ball $B := \mathbf{B}_\varepsilon((x,y))$ is also open in the product topology $\mathcal{T}$, which will imply that $\mathcal{T}_d \subseteq \mathcal{T}$. Let $w = (u,v) \in B$, then there exists $r > 0$ such that $\mathbf{B}_r(w) \subseteq B$. Let U_w be the d_X-open ball $\mathbf{B}_{r/2}(u) \subseteq X$, and let V_w be the d_Y-open ball $\mathbf{B}_{r/2}(v) \subseteq Y$. I claim that $U_w \times V_w \subseteq \mathbf{B}_r(w) \subseteq B$. Why? If $(s,t) \in U_w \times V_w$, since d is conserving,

$$d\big((s,t),(u,v)\big) \leqslant d_X(s,u) + d_Y(t,v) < \frac{r}{2} + \frac{r}{2} = r. \qquad \square$$

Exercise 2.53. Let X be a set and let d_1, d_2 be two metrics on X.

(a) Suppose that there exist $m, M \in \mathbf{R}_{>0}$ such that

$$m\, d_1(x,y) \leqslant d_2(x,y) \leqslant M\, d_1(x,y) \qquad \text{for all } x,y \in X. \tag{2.1}$$

Show that d_1 and d_2 are equivalent.

(b) Prove that the converse of (a) does not hold.

In other words, find a set X and two equivalent metrics d_1 and d_2 with the property that there **do not** exist positive real numbers m and M such that Equation (2.1) holds.

Solution.

(a) Let $\mathcal{T}_1$ be the topology defined by d_1, $\mathcal{T}_2$ the topology defined by d_2. We know that each topology is generated by the corresponding open balls.

Consider an open ball $\mathbf{B}^{d_2}_r(x)$ of $\mathcal{T}_2$. I claim that the open ball $\mathbf{B}^{d_1}_{r/M}(x)$ of $\mathcal{T}_1$ is contained in $\mathbf{B}^{d_2}_r(x)$: if $y \in \mathbf{B}^{d_1}_{r/M}(x)$ then $d_1(x,y) < r/M$, so that

$$d_2(x,y) \leqslant M\, d_1(x,y) < r.$$

So $\mathcal{T}_1$ is finer than $\mathcal{T}_2$.

Now consider an open ball $\mathbf{B}^{d_1}_r(x)$ of $\mathcal{T}_1$. I claim that the open ball $\mathbf{B}^{d_2}_{rm}(x)$ of $\mathcal{T}_2$ is contained in $\mathbf{B}^{d_1}_r(x)$: if $y \in \mathbf{B}^{d_2}_{rm}(x)$ then $d_2(x,y) < rm$, so that

$$d_1(x,y) \leqslant \frac{1}{m} d_2(x,y) < r.$$

So $\mathcal{T}_2$ is finer than $\mathcal{T}_1$, in conclusion $\mathcal{T}_1 = \mathcal{T}_2$.

(b) Let $X = \mathbf{Z}$. Let d_1 be the discrete metric on $\mathbf{Z}$. Let d_2 be the induced Euclidean metric from $\mathbf{R}$, that is $d_2(x,y) = |x-y|$ for all $x, y \in \mathbf{Z}$.

First we note that d_1 and d_2 are equivalent metrics. It suffices to show that every singleton $\{x\} \subseteq \mathbf{Z}$ is open with respect to d_2:

$$\mathbf{B}_1^{d_2}(x) = \{y \in \mathbf{Z} \colon |y-x| < 1\} = \{y \in \mathbf{Z} \colon x-1 < y < x+1\} = \{x\}.$$

Suppose that d_1 and d_2 satisfy Equation (2.1) for some $m, M > 0$. In particular, if $x \neq y$ we would have

$$m \leqslant |x-y| \leqslant M \qquad \text{for all } x \neq y \in \mathbf{Z},$$

which is blatantly false (take $y = 0$, $x = \lceil M \rceil + 1$). □

Exercise 2.54. Let X, Y be metric spaces. Show that for any $z_1, z_2 \in X \times Y$ we have

$$\frac{1}{2} d_1(z_1, z_2) \leqslant d_\infty(z_1, z_2) \leqslant d_1(z_1, z_2) \leqslant 2 d_\infty(z_1, z_2).$$

Conclude that for any conserving metric d on $X \times Y$, any $z \in X \times Y$ and any $\varepsilon > 0$ we have

$$\mathbf{B}_{\varepsilon/2}^{d_\infty}(z) \subseteq \mathbf{B}_{\varepsilon}^{d_1}(z) \subseteq \mathbf{B}_{\varepsilon}^{d}(z) \subseteq \mathbf{B}_{\varepsilon}^{d_\infty}(z) \subseteq \mathbf{B}_{2\varepsilon}^{d_1}(z).$$

Solution. The inequalities involving d_1 and d_∞ follow simply from

$$\frac{a+b}{2} \leqslant \max\{a, b\} \leqslant a + b \leqslant 2 \max\{a, b\},$$

which hold for any $a, b \in \mathbf{R}_{\geqslant 0}$.

The inclusions between open balls now follow by the same reasoning as in part (a) of Exercise 2.53. □

Exercise 2.55. Let X, Y be metric spaces and $S \subseteq X$, $T \subseteq Y$ totally bounded subsets. Prove that $S \times T$ is a totally bounded subset of $X \times Y$ (where the latter is equipped with a conserving metric d).

Solution. Let $\varepsilon > 0$ and let

$$S \subseteq \bigcup_{i=1}^{n} \mathbf{B}_{\varepsilon/2}^{X}(x_i) \qquad \text{and} \qquad T \subseteq \bigcup_{j=1}^{m} \mathbf{B}_{\varepsilon/2}^{Y}(y_j)$$

be corresponding covers of S, respectively T.

Then

$$S \times T \subseteq \bigcup_{i=1}^{n} \bigcup_{j=1}^{m} \mathbf{B}_{\varepsilon/2}^{X}(x_i) \times \mathbf{B}_{\varepsilon/2}^{Y}(y_j).$$

It remains to note that for any $(x, y) \in X \times Y$ we have

$$\mathbf{B}_{\varepsilon/2}^{X}(x) \times \mathbf{B}_{\varepsilon/2}^{Y}(y) = \mathbf{B}_{\varepsilon/2}^{d_\infty}\big((x,y)\big) \subseteq \mathbf{B}_{\varepsilon}^{d}\big((x,y)\big),$$

where $\mathbf{B}^{d_\infty}$ denotes an open ball with respect to the d_∞ metric, $\mathbf{B}^d$ denotes an open ball with respect to the d metric, and the last inclusion comes from the fact that d is conserving and Exercise 2.54. □

Exercise 2.56. Suppose X and Y are metric spaces with the property that every bounded subset of either of them is totally bounded. Prove that the same is true in the product $X \times Y$ (equipped with a conserving metric).

Solution. Let $Z \subseteq X \times Y$ be bounded, then there exists $(x,y) \in X \times Y$ and $r > 0$ such that

$$Z \subseteq \mathbf{B}^d_r\big((x,y)\big) \subseteq \mathbf{B}^{d_\infty}_r\big((x,y)\big) = \mathbf{B}^X_r(x) \times \mathbf{B}^Y_r(y).$$

Since $\mathbf{B}^X_r(x)$ and $\mathbf{B}^Y_r(y)$ are bounded in X and in Y, they are totally bounded. Therefore by Exercise 2.55 so is their product, hence so is its subset Z. □

Exercise 2.57. Let K be a sequentially compact subset of a metric space X. Prove that any open cover of K has a Lebesgue number.

Solution. Take an open cover

$$K \subseteq \bigcup_{i \in I} U_i.$$

Suppose that this has no Lebesgue number. This means that for every $n \in \mathbf{N}$, there exists a subset $A_n \subseteq K$ such that $\operatorname{diam}(A_n) < \frac{1}{n}$ and $A_n \nsubseteq U_i$ for all $i \in I$. Pick $a_n \in A_n$ to form a sequence (a_n) in K. By assumption this has a subsequence (a_{n_j}) that converges to some $x \in K$.

There exists $i \in I$ such that $x \in U_i$. Let $\varepsilon > 0$ be such that $\mathbf{B}_\varepsilon(x) \subseteq U_i$. There exists $J_1 \in \mathbf{N}$ such that $1/n_j < \varepsilon/2$ for all $j \geqslant J_1$, so that $A_{n_j} \subseteq \mathbf{B}_{\varepsilon/2}(a_{n_j})$. There exists $J_2 \in \mathbf{N}$ such that $d(a_{n_j}, x) < \varepsilon/2$ for all $j \geqslant J_2$. Letting $J = \max\{J_1, J_2\}$ we get $A_{n_j} \subseteq \mathbf{B}_\varepsilon(x) \subseteq U_i$, contradiction. □

Exercise 2.58. Let X, Y be metric spaces and let (f_n) be a sequence in $C_0(X,Y)$ that converges uniformly to $f \in C_0(X,Y)$. If $(x_n) \longrightarrow x$ in X, then $\big(f_n(x_n)\big) \longrightarrow f(x)$ in Y.

Solution. Let $\varepsilon > 0$. Since f is continuous, there exists $\delta > 0$ such that if $d_X(x', x) < \delta$ then $d_Y(f(x'), f(x)) < \varepsilon/2$.

Since $(x_n) \longrightarrow x$, there exists $N_1 \in \mathbf{N}$ such that if $n \geqslant N_1$ then $d_Y(x_n, x) < \delta$.

Since $(f_n) \longrightarrow f$, there exists $N_2 \in \mathbf{N}$ such that if $n \geqslant N_2$ then $d_Y(f_n(x'), f(x')) < \varepsilon/2$ for all $x' \in X$.

Let $N = \max\{N_1, N_2\}$, then if $n \geqslant N$ we have

$$d_Y(f_n(x_n), f(x)) \leqslant d_Y(f_n(x_n), f(x_n)) + d_Y(f(x_n), f(x)) < \varepsilon.$$ □

Exercise 2.59. If X and Y are metric spaces with X compact and $K \subseteq C_0(X,Y)$ is compact, then K is bounded, closed, and equicontinuous.

(This is a converse to the Arzelà–Ascoli Theorem, see Theorem 2.84.)

Solution. We know that K is bounded (since every compact subset is totally bounded, hence bounded by Exercise 2.47) and that K is closed by Proposition 2.35.

Suppose K is not equicontinuous: there exists $\varepsilon > 0$ such that for all $\delta > 0$ there exists $f \in K$ and $x, x' \in X$ with $d_X(x, x') < \delta$ and $d_Y(f(x), f(x')) \geqslant \varepsilon$.

In particular, we can take $\delta = 1/n$ for $n \in \mathbf{N}$ and obtain a sequence (f_n) in K and two equivalent sequences $(x_n) \sim (x'_n)$ in X such that

$$d_Y(f_n(x_n), f_n(x'_n)) \geqslant \varepsilon.$$

But K is compact so (f_n) has a subsequence $(f_{n_k}) \longrightarrow f \in K$.

The corresponding subsequence (x_{n_k}) of (x_n) is a sequence in X, which is compact, so itself has a subsequence $(x_{n_{k_j}}) \longrightarrow x \in X$. Since $(x'_n) \sim (x_n)$, we also have $(x'_{n_{k_j}}) \longrightarrow x$.

Now Exercise 2.58 tells us that $(f_{n_{k_j}}(x_{n_{k_j}}))$ and $(f_{n_{k_j}}(x'_{n_{k_j}}))$ both converge to $f(x)$, contradicting the fact that their terms stay at least ε apart. □

Exercise 2.60. Let $\mathbf{S}^1 = \mathbf{S}_1\big((0,0)\big) = \big\{x, y \in \mathbf{R} \colon x^2 + y^2 = 1\big\}$ be the unit circle in $\mathbf{R}^2$.

Consider the function $f \colon [0,1) \longrightarrow \mathbf{S}^1$ given by the parametrisation

$$f(t) = \big(\cos(2\pi t), \sin(2\pi t)\big).$$

Endow $[0,1)$ with the induced metric from $\mathbf{R}$ and $\mathbf{S}^1$ with the induced metric from $\mathbf{R}^2$.

Prove that f is a bijective continuous function, but not a homeomorphism.

(You may use without proof whatever properties of the functions sin and cos you manage to remember from previous subjects.)

Solution.

(a) We know that $t \longmapsto 2\pi t$, $t \longmapsto \cos(t)$ and $t \longmapsto \sin(t)$ are continuous, so by Tutorial Question 3.7 so is f.

(b) Suppose $t_1 \neq t_2 \in [0,1)$ are such that $f(t_1) = f(t_2)$. Then $\cos(2\pi t_1) = \cos(2\pi t_2)$, which implies that $t_2 = 1 - t_1$. In that case $\sin(2\pi t_2) = \sin(2\pi - 2\pi t_1) = \sin(-2\pi t_1) = -\sin(2\pi t_1)$. But we also have $\sin(2\pi t_2) = \sin(2\pi t_1)$, so $\sin(2\pi t_1) = 0$, hence $t_1 = 0$ and $t_2 = 1 - t_1 = 1$, contradicting $t_2 \in [0,1)$.

We conclude that f is injective.

For surjectivity, let $(x,y) \in \mathbf{S}^1$, in other words $x^2 + y^2 = 1$. Define $\theta \in [0, 2\pi)$ by

$$\theta = \begin{cases} \arccos(x) & \text{if } y \geqslant 0 \\ 2\pi - \arccos(x) & \text{if } y < 0. \end{cases}$$

Letting $t = \theta/(2\pi)$, we have $f(t) = (x,y)$.

(c) At this point we know that f is a homeomorphism iff $f^{-1} \colon \mathbf{S}^1 \longrightarrow [0,1)$ is continuous. Note that $\mathbf{S}^1 \subseteq \mathbf{R}^2$ is compact: it is clearly bounded as any two points are at distance at most 2 of each other, so we just need to check that it is a closed subset of $\mathbf{R}^2$.

But $\mathbf{S}^1 = \mathbf{D}_1\big((0,0)\big) \cap C$ is the intersection of two closed sets, where

$$C = \{x, y \in \mathbf{R} \colon x^2 + y^2 \geqslant 1\} = \mathbf{R}^2 \smallsetminus \mathbf{B}_1\big((0,0)\big).$$

Since $\mathbf{S}^1$ is compact, if f^{-1} were continuous then $[0,1) = f^{-1}(\mathbf{S}^1)$ would be compact, hence closed in $\mathbf{R}$. This is a contradiction, because 1 is an accumulation point of $[0,1)$ but does not lie in the set. □

Exercise 2.61. Prove that any Cauchy sequence (x_n) in a metric space (X,d) is *bounded*, that is there exists $C \geqslant 0$ such that $d(x_n, x_m) \leqslant C$ for all $n, m \in \mathbf{N}$.

Solution. Let $N \in \mathbf{N}$ be such that for all $m, n \geqslant N$ we have $d(x_m, x_n) < 1$.

Let $B = \max\{d(x_m, x_N) : 1 \leqslant m < N\}$. Let $C = 2B + 1$, then we have

$$d(x_m, x_n) \leqslant \begin{cases} 1 \leqslant C & \text{if } m, n \geqslant N \\ d(x_m, x_N) + d(x_N, x_n) \leqslant B + 1 \leqslant C & \text{if } m < N, n \geqslant N \\ d(x_m, x_N) + d(x_N, x_n) \leqslant 2B \leqslant C & \text{if } m, n < N. \end{cases}$$

□

Exercise 2.62. Let X be a topological space and define $x \sim x'$ if there exists a connected subset $C \subset X$ such that $x, x' \in C$.

Prove that this is an equivalence relation on the set X, thereby partitioning X into a disjoint union of maximal connected subsets (these are called the *connected components* of X).

[*Hint*: Recall that an equivalence relation has three defining axioms: (a) $x \sim x$ for all $x \in X$; (b) if $x \sim x'$ then $x' \sim x$; (c) if $x \sim x'$ and $x' \sim x''$ then $x \sim x''$.]

Solution.

(a) $x \sim x$: for any $x \in X$, the set $C = \{x\}$ is connected and contains x, so $x \sim x$.

(b) if $x \sim x'$ then $x' \sim x$: clear from the definition, which does not distinguish x and x'.

(c) if $x \sim x'$ and $x' \sim x''$ then $x \sim x''$: since $x \sim x'$ there exists a connected set C_1 such that $x, x' \in C_1$; since $x' \sim x''$ there exists a connected set C_2 such that $x', x'' \in C_2$; by Tutorial Question 4.2, since C_1 and C_2 are connected and $x' \in C_1 \cap C_2$, the union $C_1 \cup C_2$ is connected, and it contains both x and x'', so that $x \sim x''$. □

Exercise 2.63. Let (X,d) be a metric space.

If A and B are bounded sets with $A \cap B \neq \varnothing$, then

$$\operatorname{diam}(A \cup B) \leqslant \operatorname{diam}(A) + \operatorname{diam}(B).$$

Solution. It suffices to show that for any $x, y \in A \cup B$ we have

$$d(x,y) \leqslant \operatorname{diam}(A) + \operatorname{diam}(B).$$

If $x, y \in A$, this is obvious as $d(x,y) \leqslant \operatorname{diam}(A)$. Similarly if $x, y \in B$.

It remains to see what happens if $x \in A$ and $y \in B$. Let $t \in A \cap B$. We have

$$d(x,y) \leqslant d(x,t) + d(t,y) \leqslant \operatorname{diam}(A) + \operatorname{diam}(B),$$

as desired. □

Exercise 2.64. Consider the equation

$$x^3 - x - 1 = 0. \tag{2.2}$$

(a) Show that the equation must have **at least one solution** in the interval $[1,2]$.

(b) Show that the function $f\colon [1,2] \longrightarrow [1,2]$ given by

$$f(x) = (1+x)^{1/3}$$

is a contraction.

(c) Show that Equation (2.2) has **a unique solution** ξ in the interval $[1,2]$ and describe a sequence of real numbers that converges to ξ.

Solution.

(a) We can use the Intermediate Value Theorem: at $x = 1$, $x^3 - x - 1 = -1 < 0$, while at $x = 2$, $x^3 - x - 1 = 5 > 0$, so there must be at least one point x in $[1,2]$ such that $x^3 - x - 1 = 0$.

(b) The derivative of f is

$$f'(x) = \frac{1}{3}(1+x)^{-2/3} = \frac{1}{3}\frac{1}{(1+x)^{2/3}}.$$

As $x \in [1,2]$, we have $f'(x) > 0$ and

$$1 \leqslant x \Rightarrow 2 \leqslant 1 + x \Rightarrow \frac{1}{1+x} \leqslant \frac{1}{2} \Rightarrow \frac{1}{(1+x)^{2/3}} \leqslant \frac{1}{2^{2/3}} \leqslant 1,$$

so that

$$f'(x) \leqslant \frac{1}{3}.$$

Now let x, y be such that $1 \leqslant x < y \leqslant 2$ and apply the Mean Value Theorem to f on $[x,y]$ to deduce that there exists $c \in (x,y)$ such that

$$\frac{f(y) - f(x)}{y - x} = f'(c) \Rightarrow |f(y) - f(x)| = |f'(c)|\,|y - x| \leqslant \frac{1}{3}|y - x|.$$

We conclude that f is a contraction.

(c) Observe that $x^3 - x - 1 = 0$ is equivalent to $f(x) = x$, so the solutions of Equation (2.2) are precisely the fixed points of f. As f is a contraction and $[1,2]$ is complete, the Banach Fixed Point Theorem says that there is a unique fixed point ξ in $[1,2]$. It also tells us that we can start with any $x_1 \in [1,2]$, for instance $x_1 = 1$, and iteratively apply f to get a sequence (x_n) converging to ξ:

$$x_1 = 1, \qquad x_2 = f(x_1) = 2^{1/3}, \qquad x_3 = f(x_2) = \left(1 + 2^{1/3}\right)^{1/3}, \ldots$$

□

Exercise 2.65. Let $f\colon \mathbf{R} \longrightarrow \mathbf{R}$ be a contraction and define $F\colon \mathbf{R} \longrightarrow \mathbf{R}$ by

$$F(x) = x + f(x).$$

(a) Use the Banach Fixed Point Theorem to show that the equation $x + f(x) = a$ has a unique solution for any $a \in \mathbf{R}$.

(b) Deduce that F is a bijection.

(c) Show that F is continuous.

(d) Show that F^{-1} is continuous (so it is a homeomorphism).

Solution.

(a) Given $a \in \mathbf{R}$, let $f_a\colon \mathbf{R} \longrightarrow \mathbf{R}$ be given by

$$f_a(x) = a - f(x).$$

Note that f_a is a contraction:

$$\left|f_a(x) - f_a(y)\right| = \left|a - f(x) - a + f(y)\right| = \left|f(y) - f(x)\right| \leqslant c\left|x - y\right| \qquad \text{for all } x, y \in \mathbf{R}.$$

Next note that $F(x) = a$ if and only if $a = x + f(x)$ if and only if $x = f_a(x)$ if and only if x is a fixed point of f_a.

By the Banach Fixed Point Theorem, f_a has a unique fixed point; therefore $F(x) = a$ has a unique solution.

(b) $F(x) = a$ having a unique solution for every $a \in \mathbf{R}$ is saying precisely that $F\colon \mathbf{R} \longrightarrow \mathbf{R}$ is bijective.

(c) If $c = 0$ then f is a constant function $f(x) = b$ so $F(x) = x + b$, clearly continuous with continuous inverse $F^{-1}(x) = x - b$.

So we may assume $c > 0$ (also in part (d)).

Given $\varepsilon > 0$, let $\delta = \varepsilon/c$, then if $|x - y| < \delta$ we have

$$|f(x) - f(y)| < c\delta = c\frac{\varepsilon}{c} = \varepsilon.$$

We conclude that f is (uniformly) continuous, so F is continuous, being the sum of the continuous functions $x \longmapsto x$ and $x \longmapsto f(x)$.

(d) The Banach Fixed Point Theorem tells us that the unique fixed point of f_a is the limit of the iterates of f_a evaluated at any starting point in $\mathbf{R}$, for instance at 0:

$$F^{-1}(a) = \lim_{n \longrightarrow \infty}\left(f_a^{\circ n}(0)\right).$$

Let $a, b \in \mathbf{R}$. I claim that for any $n \in \mathbf{N}$ we have

$$\left|f_a^{\circ n}(0) - f_b^{\circ n}(0)\right| \leqslant \left(1 + c + \cdots + c^{n-1}\right)\left|a - b\right|. \tag{2.3}$$

We prove this by induction on n. The base case is $n = 1$, where we have

$$\left|f_a(0) - f_b(0)\right| = \left|a - f(0) - b + f(0)\right| = |a - b|.$$

Fix $n \in \mathbf{N}$ and assume that the inequality (2.3) holds for n. We have

$$\begin{aligned}\left|f_a^{\circ(n+1)}(0) - f_b^{\circ(n+1)}(0)\right| &= \left|a - f\big(f_a^{\circ n}(0)\big) - b + f\big(f_b^{\circ n}(0)\big)\right| \\ &\leqslant \left|a - b\right| + c\big(1 + c + \cdots + c^{n-1}\big)\left|a - b\right| \\ &= \big(1 + c + \cdots + c^n\big)\left|a - b\right|,\end{aligned}$$

where in the second to last step we used the contractive property of f and the inequality (2.3) for n.

Finally, we have

$$\left|F^{-1}(a) - F^{-1}(b)\right| = \lim_{n \to \infty} \left|f_a^{\circ n}(0) - f_b^{\circ n}(0)\right| \leqslant \frac{1}{1-c}\left|a - b\right|.$$

So for any $\varepsilon > 0$ we can take $\delta < (1-c)\varepsilon$ and deduce that F^{-1} is continuous. □

Exercise 2.66. Let X be the interval $(0, 1/3)$ in $\mathbf{R}$ with the Euclidean metric. Show that $f\colon X \longrightarrow X$ defined by $f(x) = x^2$ is a contraction, but does not have a fixed point in X. Why does this not contradict the Banach Fixed Point Theorem?

Solution. First we check that f does take values in X: if $x \in (0, 1/3)$ then $0 < x < 1/3$ so $0 < x^2 < 1/9 < 1/3$.

Next we note that $f(x) = x^2$ is differentiable with continuous derivative on $(0, 1/3)$ so the Mean Value Theorem applies on any subinterval $(x, y) \subseteq (0, 1/3)$:

$$|f(x) - f(y)| = |f'(\xi)|\,|x - y| \qquad \text{for some } \xi \in (x, y) \subseteq (0, 1/3).$$

Of course $f'(\xi) = 2\xi$ so if $\xi \in (0, 1/3)$ then $f'(\xi) \in (0, 2/3)$, proving that f is a contraction with constant (at most) $2/3$.

What are the fixed points of f? They satisfy $x = f(x) = x^2$, so $x = 0$ or $x = 1$, but neither of these is in $X = (0, 1/3)$.

The Banach Fixed Point Theorem is not contradicted: one of the assumptions is that X is complete, but $(0, 1/3) \subseteq \mathbf{R}$ is not complete since it is not closed in the complete metric space $\mathbf{R}$. □

Exercise 2.67. Let (X, d) be a complete metric space and $f\colon X \longrightarrow X$ be a function. Let $g = f \circ f$, that is, $g(x) = f\big(f(x)\big)$. Suppose that $g\colon X \longrightarrow X$ is a contraction. Prove that f has a unique fixed point.

Solution. By the Banach Fixed Point Theorem, g has a unique fixed point $x_0 \in X$. I claim that x_0 is also the unique fixed point of f. For uniqueness, note that if $f(x) = x$ then $g(x) = f\big(f(x)\big) = f(x) = x$ so x is a fixed point of g, hence $x = x_0$. To show that $f(x_0) = x_0$, note that $f(x_0) = f\big(g(x_0)\big) = g\big(f(x_0)\big)$, so $f(x_0)$ is a fixed point of g, hence $f(x_0) = x_0$. □

Exercise 2.68. Prove that no two of the following spaces are homeomorphic:

(a) the interval $X = [-1,1]$ in $\mathbf{R}$;

(b) the open unit disc Y in $\mathbf{R}^2$;

(c) the closed unit disc Z in $\mathbf{R}^2$.

Solution. Y is not compact since it is not closed in $\mathbf{R}^2$, for instance the point $(0,1)$ is in the closure of Y but not in Y. On the other hand, Z is compact since it is closed and bounded in $\mathbf{R}^2$. Similarly, X is compact.

So Y and Z are not homeomorphic, and X and Y are not homeomorphic.

Suppose $f\colon X \longrightarrow Z$ is a homeomorphism. Let $x \in X^\circ$, then $f(x) \in Z^\circ$. The restriction of f to $X \setminus \{x\} \longrightarrow Z \setminus \{f(x)\}$ is then also a homeomorphism, but this is impossible since $X \setminus \{x\} = [-1,x) \cup (x,1]$ is disconnected, while $Z \setminus \{f(x)\}$ is connected. □

Exercise 2.69. Are the following pairs of spaces homeomorphic or not?

(a) the unit circle in $\mathbf{R}^2$ and the unit interval $[0,1]$ in $\mathbf{R}$;

(b) the intervals $[0,1]$ and $(0,1)$ in $\mathbf{R}$;

(c) the intervals $[0,1]$ and $[0,2]$ in $\mathbf{R}$.

Solution.

(a) No: removing an interior point of $[0,1]$ gives a disconnected set, but removing any point from the unit circle gives a set that is connected.

(b) No: $[0,1]$ is compact, being closed and bounded in $\mathbf{R}$, while $(0,1)$ is not compact, since it is not closed in $\mathbf{R}$.

(c) Yes: $f\colon [0,1] \longrightarrow [0,2]$ given by $f(x) = 2x$ is clearly a homeomorphism. □

Exercise 2.70. Which of the following metric spaces are compact?

(a) The unit circle in $\mathbf{R}^2$.

(b) The unit open disk in $\mathbf{R}^2$.

(c) The closed unit ball in the space ℓ^∞ of bounded real sequences $\big(a_1, a_2, \dots\big)$.

Solution.

(a) Compact: closed and bounded in $\mathbf{R}^2$.

(b) Not compact: not closed, since $(1,0)$ is in the closure of the open disk but not in the open disk itself.

(c) Not compact: the sequence (e_n) of standard vectors has no convergent subsequence, since $d(e_n, e_m) = 1$ whenever $n \neq m$. □

Exercise 2.71. Let C be a nonempty compact subset of a metric space (X, d). Prove that there exist points $a, b \in C$ such that

$$d(a,b) = \sup\{d(x,y) : x, y \in C\}.$$

In other words, the diameter of C is realised as the distance between two points of C.

Solution. As you know from Tutorial Question 6.7, the distance function $d\colon X \times X \longrightarrow \mathbf{R}$ is continuous. By Theorem 2.39, $C \times C$ is compact, so by Proposition 2.70 there exists $(a_{\max}, b_{\max}) \in C \times C$ such that

$$d(a,b) \leqslant d(a_{\max}, b_{\max}) \qquad \text{for all } (a,b) \in C \times C.$$

Therefore $a_{\max}, b_{\max} \in C$ realise the diameter of C. □

Exercise 2.72. Let (X, d) be a metric space and let $S \subseteq X$ be a nonempty subset. Define $d_S\colon X \longrightarrow \mathbf{R}_{\geqslant 0}$ by

$$d_S(x) = \inf_{s \in S} d(x,s).$$

(a) Prove that d_S is uniformly continuous.
[*Hint*: Show that $|d_S(x) - d_S(y)| \leqslant d(x,y)$ for all $x, y \in X$.]

(b) Prove that $d_S(x) = 0$ if and only if $x \in \overline{S}$.

(c) Prove that if $U \subseteq X$ is an open neighbourhood of x, then $d_{X \setminus U}(x) > 0$.

Solution.

(a) We start with the hint. Let $x, y \in X$. For all $s \in S$ we have

$$d_S(x) \leqslant d(x,s) \leqslant d(x,y) + d(y,s),$$

hence

$$d_S(x) \leqslant d(x,y) + d_S(y).$$

We can swap the roles of x and y to get

$$d_S(y) \leqslant d(y,x) + d_S(x),$$

and the two inequalities together give

$$|d_S(x) - d_S(y)| \leqslant d(x,y).$$

Uniform continuity is now clear: for any $\varepsilon > 0$ we take $\delta = \varepsilon$ and use the above inequality.

(b) If $d_S(x) = 0$ then $\inf d(x,s) = 0$ so for any $\varepsilon > 0$ there exists $s \in S$ such that $d(x,s) < \varepsilon$. In particular, for $n \in \mathbf{N}$ we can set $\varepsilon = 1/n$ and get $s_n \in S$ such that $d(x, s_n) < \varepsilon$. This gives us a sequence (s_n) in S that converges to x, so $x \in \overline{S}$.

Conversely, if $x \in \overline{S}$ then there exists a sequence (s_n) in S that converges to x. Given $\varepsilon > 0$, there exists $N \in \mathbf{N}$ such that $d(x, s_N) < \varepsilon$, therefore $\inf d(x,s) = 0$.

(c) If $d_{X\smallsetminus U}(x) = 0$ then by part (b) we have $x \in \overline{X \smallsetminus U} = X \smallsetminus U$, the latter equality due to U being open. But then $x \in U \cap (X \smallsetminus U)$, contradiction. □

Exercise 2.73. Give an example of a metric space X and an open ball $\mathbf{B}_\varepsilon(x)$ such that

$$\overline{\mathbf{B}_\varepsilon(x)} \neq \mathbf{D}_\varepsilon(x).$$

Solution. Take $X = \{0,1\}$ with the discrete metric, $x = 0$ and $\varepsilon = 1$. Then

$$\overline{\mathbf{B}_1(0)} = \overline{\{0\}} = \{0\} \neq \{0,1\} = \mathbf{D}_1(0).$$ □

Exercise 2.74.

(a) Suppose $f\colon \mathbf{R}^n \longrightarrow \mathbf{R}^n$ is a continuous function and S is a bounded subset of $\mathbf{R}^n$. Prove that $f(S)$ is bounded.

(b) Find a uniformly continuous function $f\colon X \longrightarrow Y$ between metric spaces and a bounded subset B of X such that $f(B)$ is unbounded.

Solution.

(a) Let $f\colon \mathbf{R}^n \longrightarrow \mathbf{R}^n$ be a continuous function and let B be a bounded subset of $\mathbf{R}^n$. It follows from Exercise 2.45 that B is contained in some closed ball $\mathbf{D}_r(v)$, which is compact by part (b) of Tutorial Question 7.6. Hence $f\big(\mathbf{D}_r(v)\big)$ is compact by Proposition 2.37, and therefore bounded by part (c) of Tutorial Question 7.6. Since $f(B) \subseteq f\big(\mathbf{D}_r(v)\big)$, it follows that

$$\begin{aligned}\operatorname{diam}\big(f(B)\big) &= \sup\{d(x,y)\colon x,y \in f(B)\}\\ &\leqslant \sup\{d(x,y)\colon x,y \in \mathbf{D}_r(v)\}\\ &= \operatorname{diam}\big(\mathbf{D}_r(v)\big) < \infty.\end{aligned}$$

Hence $f(B)$ is bounded.

(b) Let $X = (\mathbf{N}, d_1)$ and $Y = (\mathbf{N}, d_2)$, where d_1 is the discrete metric on $\mathbf{N}$ and d_2 is the Euclidean metric on $\mathbf{N}$.

We claim that the identity function $\mathrm{id}_{\mathbf{N}}\colon X \longrightarrow Y$ is uniformly continuous. Indeed, for every positive real number ϵ, put $\delta = 1$. If $d_1(x,y) < 1$, then $x = y$, and therefore $d_2\big(\mathrm{id}_X(x), \mathrm{id}_X(y)\big) = 0 < \epsilon$.

Since $\mathbf{B}_2^{d_1}(0) = \mathbf{N}$, it follows that $\mathbf{N}$ is bounded in X. However, $\mathrm{id}_{\mathbf{N}}(\mathbf{N}) = \mathbf{N}$ is not bounded because

$$\operatorname{diam}_{d_2}(\mathbf{N}) = \sup\{d_2(m,n)\colon m,n \in \mathbf{N}\} = \sup \mathbf{Z} = \infty.$$ □

Exercise 2.75.

(a) Suppose $f\colon \mathbf{R}^n \longrightarrow \mathbf{R}^n$ is a continuous function and S is a totally bounded subset of $\mathbf{R}^n$. Prove that $f(S)$ is totally bounded.

(b) Find a continuous function $f\colon X \longrightarrow Y$ between metric spaces and a totally bounded subset S of X such that $f(S)$ is not totally bounded.

Solution.

(a) The subset S of $\mathbf{R}^n$ is bounded because of Exercise 2.47, and therefore $f(S)$ is bounded by part (a) of Exercise 2.74. It then follows from part (d) of Tutorial Question 7.6 that $f(S)$ is totally bounded.

(b) Let $X = (-\pi/2, \pi/2)$, $Y = \mathbf{R}$, and let $f\colon X \longrightarrow Y$ be the continuous function defined by $f(x) = \tan(x)$. The domain $(-\pi/2, \pi/2)$ is bounded because its diameter is π, but its image is the unbounded set $\mathbf{R}$. $\square$

Exercise 2.76. Let $f\colon X \longrightarrow Y$ be a function between metric spaces.

(a) Prove that f is a contraction if and only if $\operatorname{diam}\big(f(S)\big) < \operatorname{diam}(S)$ for every subset S of X.

(b) Suppose f is a contraction and B is a bounded subset of X. Prove that $f(B)$ is bounded.

Solution.

(a) If f is a contraction and S is a subset of X, then

$$\begin{aligned}\operatorname{diam}\big(f(S)\big) &= \sup\big\{d_Y\big(f(x_1), f(x_2)\big)\colon x_1, x_2 \in S\big\}\\ &< \sup\{d_X(x_1, x_2)\colon x_1, x_2 \in S\}\\ &= \operatorname{diam}(S)\end{aligned}$$

Conversely, suppose $\operatorname{diam}\big(f(S)\big) < \operatorname{diam}(S)$ for every subset S of X. If x_1 and x_2 are elements of X, then

$$d_Y\big(f(x_1), f(x_2)\big) = \operatorname{diam}\big(\{f(x_1), f(x_2)\}\big) < \operatorname{diam}(\{x_1, x_2\}) = d_X(x_1, x_2).$$

Hence f is a contraction.

(b) It follows from part (a) that

$$\operatorname{diam}\big(f(B)\big) < \operatorname{diam}(B) < \infty.$$

Hence $f(B)$ is bounded. $\square$

3. Normed and Hilbert spaces

Exercise 3.1. Let $(V, \|\|)$ be a normed vector space. Prove that the norm function $\|\cdot\| \colon V \longrightarrow \mathbf{R}_{\geqslant 0}$ is uniformly continuous.

Solution. Given $\varepsilon > 0$, let $\delta = \varepsilon$. I claim that if $d_V(v, w) < \varepsilon$ then

$$d_{\mathbf{R}}\big(\|v\|, \|w\|\big) = \big|\|v\| - \|w\|\big| < \varepsilon.$$

To prove this, note that

$$\begin{aligned} \|v\| &= \|v - w + w\| \leqslant \|v - w\| + \|w\| \Rightarrow \|v\| - \|w\| \leqslant \|v - w\| \\ \|w\| &= \|v + w - v\| \leqslant \|v\| + \|w - v\| \Rightarrow -\|v - w\| \leqslant \|v\| - \|w\|, \end{aligned}$$

so that

$$d_{\mathbf{R}}\big(\|v\|, \|w\|\big) = \big|\|v\| - \|w\|\big| \leqslant \|v - w\| = d_V(v, w),$$

and the rest follows. □

Exercise 3.2. If $\|\cdot\|_1$ and $\|\cdot\|_2$ are equivalent norms on V, then the corresponding metrics d_1 and d_2 (as in Proposition 3.1) are equivalent.

Solution. By Proposition 2.21 we know that d_2 is coarser than d_1 if and only if the function $(V, d_1) \longrightarrow (V, d_2)$ given by $v \longmapsto v$ is continuous. By Theorem 2.52, this in turn is equivalent to showing that for every $v \in V$, every sequence that converges to v in (V, d_1) also converges to v in (V, d_2).

So let (v_n) be a sequence that converges to v in (V, d_1), that is $\big(d_1(v_n, v)\big) \longrightarrow 0$, so $\big(\|v_n - v\|_1\big) \longrightarrow 0$, hence $\big(m\|v_n - v\|_1\big) \longrightarrow 0$ and $\big(M\|v_n - v\|_1\big) \longrightarrow 0$. Since by assumption

$$m\|v_n - v\|_1 \leqslant \|v_n - v\|_2 \leqslant M\|v_n - v\|_1,$$

this implies by the Sandwich Theorem that $\big(\|v_n - v\|_2\big) \longrightarrow 0$, in other words that $(v_n) \longrightarrow v$ in (V, d_2).

The fact that d_1 is coarser than d_2 follows because

$$\frac{1}{M}\|v\|_2 \leqslant \|v\|_1 \leqslant \frac{1}{m}\|v\|_2 \qquad \text{for all } v \in V,$$

so we can interchange the roles of d_1 and d_2 in the previous argument. □

Exercise 3.3. Let V be a vector space over $\mathbf{F}$. Show that the intersection of an arbitrary collection of convex subsets of V is convex.

Solution. Suppose I is an arbitrary set and S_i is a convex subset of V for all $i \in I$. Let

$$S = \bigcap_{i \in I} S_i$$

and let $v, w \in S$, $a, b \in \mathbf{R}_{\geqslant 0}$ such that $a + b = 1$. Then for all $i \in I$ we have $v, w \in S_i$, so that $av + bw \in S_i$ since S_i is convex. Therefore $av + bw \in S$. □

Exercise 3.4. Prove that, if $(V, \|\cdot\|)$ is a normed space, then $f : V \longrightarrow \mathbf{R}$ given by $f(v) = \|v\|$ is a convex function.

Solution. Suppose $v, w \in S$ and $a, b \in \mathbf{R}_{\geqslant 0}$ such that $a + b = 1$. Then

$$f(av + bw) = \|av + bw\| \leqslant \|av\| + \|bw\| = |a|\,\|v\| + |b|\,\|w\| = a\|v\| + b\|w\| = af(v) + bf(w). \quad □$$

Exercise 3.5. (*) Let $I \subseteq \mathbf{R}$ be an interval and let $f : I \longrightarrow \mathbf{R}$ be a twice-differentiable function.

The aim of this Exercise is to check the familiar calculus fact: f is convex if and only if $f''(x) \geqslant 0$ for all $x \in I$.

It was heavily inspired by Alexander Nagel's Wisconsin notes [1]:

https://people.math.wisc.edu/~ajnagel/convexity.pdf

(a) For any $s, t \in I$ with $s < t$, define the linear function $L_{s,t} : [s, t] \longrightarrow \mathbf{R}$ by

$$L_{s,t}(x) = f(s) + \left(\frac{x - s}{t - s}\right)\big(f(t) - f(s)\big).$$

Convince yourself that this is the equation of the secant line joining $(s, f(s))$ to $(t, f(t))$.

Prove that f is convex on I if any only if

$$f(x) \leqslant L_{s,t}(x) \qquad \text{for all } s, t \in I \text{ such that } s < t \text{ and all } s \leqslant x \leqslant t.$$

(b) Check that for all $s, t \in I$ such that $s < t$ we have

$$L_{s,t}(x) - f(x) = \frac{x - s}{t - s}\big(f(t) - f(x)\big) - \frac{t - x}{t - s}\big(f(x) - f(s)\big).$$

(c) Use the Mean Value Theorem for f twice to prove that there exist ξ, ζ with $x < \xi < t$ and $s < \zeta < x$ such that

$$L_{s,t}(x) - f(x) = \frac{(t - x)(x - s)}{t - s}\big(f'(\xi) - f'(\zeta)\big).$$

(d) Use the Mean Value Theorem once more to conclude that if $f''(x) \geqslant 0$ for all $x \in I$, then f is convex on I.

(e) Now we prove the converse. From this point on, assume that $f : I \longrightarrow \mathbf{R}$ is twice-differentiable and convex, and let $s, t \in I^\circ$.

1. Show that if $s < x < t$ then
$$\frac{f(x) - f(s)}{x - s} \leqslant \frac{f(t) - f(x)}{t - x}.$$
2. Conclude that if $s < x_1 < x_2 < t$ then
$$\frac{f(x_1) - f(s)}{x_1 - s} \leqslant \frac{f(t) - f(x_2)}{t - x_2}.$$
3. Conclude that if $s < t$ then $f'(s) \leqslant f'(t)$, and finally that $f''(x) \geqslant 0$ on I.

Solution. Parts (b)–(d) are pretty thoroughly discussed in the above reference if you need more guidance, so I'll just do parts (a) and (e).

(a) In the definition of convex function, take $v = s$, $w = t$, $a = (t - x)/(t - s)$, $b = (x - s)/(t - s)$, so that $av + bw = x$. Then we know that
$$f(x) \leqslant \frac{t - x}{t - s} f(s) + \frac{x - s}{t - s} f(t) = f(s) + \frac{x - s}{t - s}\big(f(t) - f(s)\big) = L_{s,t}(x).$$
The other direction is straightforward.

(e) 1. From part (a) we have
$$\frac{f(x) - f(s)}{x - s} \leqslant \frac{f(t) - f(s)}{t - s}.$$
Cross-multiplying, we end up with
$$x\big(f(t) - f(s)\big) - s\big(f(t) - f(x)\big) - t\big(f(x) - f(s)\big) \geqslant 0,$$
which is also equivalent to the inequality we are trying to prove.

2. Apply the previous part twice, first with $s < x_1 < x_2$ and then with $x_1 < x_2 < t$, to get
$$\frac{f(x_1) - f(s)}{x_1 - s} \leqslant \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leqslant \frac{f(t) - f(x_2)}{t - x_2}.$$

3. Following from the previous part, we have
$$f'(s) = \lim_{x_1 \searrow s} \frac{f(x_1) - f(s)}{x_1 - s} \leqslant \lim_{x_2 \nearrow t} \frac{f(t) - f(x_2)}{t - x_2} = f'(t).$$
This implies that f' is an increasing function on I°, therefore $f''(x) \geqslant 0$ on I°. □

Exercise 3.6. Let $(V, \|\cdot\|)$ be a normed space and take $r, s > 0$, $u, v \in V$, $\alpha \in \mathbf{F}^\times$. Show that

(a) $\mathbf{B}_r(u + v) = \mathbf{B}_r(u) + \{v\}$;

(b) $\alpha\,\mathbf{B}_1(0) = \mathbf{B}_{|\alpha|}(0)$;

(c) $\mathbf{B}_r(v) = r\mathbf{B}_1(0) + \{v\}$;

(d) $r\mathbf{B}_1(0) + s\mathbf{B}_1(0) = (r+s)\mathbf{B}_1(0)$;

(e) $\mathbf{B}_r(u) + \mathbf{B}_s(v) = \mathbf{B}_{r+s}(u+v)$;

(f) $\mathbf{B}_1(0)$ is a convex subset of V;

(g) any open ball in V is convex.

Solution.

(a)

$$\begin{aligned} w \in \mathbf{B}_r(u+v) &\iff \|(u+v)-w\| < r \\ &\iff \|u-(w-v)\| < r \\ &\iff w-v \in \mathbf{B}_r(u) \\ &\iff w \in \mathbf{B}_r(u) + \{v\}. \end{aligned}$$

(b)

$$\begin{aligned} w \in \alpha\mathbf{B}_1(0) &\iff \frac{1}{\alpha} w \in \mathbf{B}_1(0) \\ &\iff \left\|\frac{1}{\alpha} w\right\| < 1 \\ &\iff \|w\| < |\alpha| \\ &\iff w \in \mathbf{B}_{|\alpha|}(0). \end{aligned}$$

(c) From (a) and (b):

$$\mathbf{B}_r(v) = \mathbf{B}_r(0) + \{v\} = r\mathbf{B}_1(0) + \{v\}.$$

(d) If $\|u\| < r$ and $\|v\| < s$ then $\|u+v\| < r+s$, so $r\mathbf{B}_1(0) + s\mathbf{B}_1(0) \subseteq (r+s)\mathbf{B}_1(0)$. Conversely, if $\|w\| < r+s$, then

$$w = \frac{r}{r+s} w + \frac{s}{r+s} w \in r\mathbf{B}_1(0) + s\mathbf{B}_1(0).$$

(e) From (c) and (d):

$$\mathbf{B}_r(u) + \mathbf{B}_s(v) = r\mathbf{B}_1(0) + s\mathbf{B}_1(0) + \{u\} + \{v\} = (r+s)\mathbf{B}_1(0) + \{u+v\} = \mathbf{B}_{r+s}(u+v).$$

(f) If $u, v \in \mathbf{B}_1(0)$ and $0 \leqslant a \leqslant 1$, then by (d)

$$au + (1-a)v \in a\mathbf{B}_1(0) + (1-a)\mathbf{B}_1(0) = (a+1-a)\mathbf{B}_1(0) = \mathbf{B}_1(0).$$

(g) $\mathbf{B}_r(u) = r\mathbf{B}_1(0) + \{u\}$ is the translate of a convex set, hence is itself convex. □

Exercise 3.7. Let $(V, \|\cdot\|)$ be a normed space and let S, T be subsets of V and $\alpha \in \mathbf{F}$. Prove that

(a) If S and T are bounded, so are $S+T$ and αS.

(b) If S and T are totally bounded, so are $S+T$ and αS.

(c) If S and T are compact, so are $S+T$ and αS.

Solution.

(a) A subset S of V is bounded if and only if $S \subseteq \mathbf{B}_s(0) = s\mathbf{B}_1(0)$ for some $s \geqslant 0$. So $S \subseteq s\mathbf{B}_1(0)$ and $T \subseteq t\mathbf{B}_1(0)$, hence $S+T \subseteq s\mathbf{B}_1(0) + t\mathbf{B}_1(0) = (s+t)\mathbf{B}_1(0)$.

Similarly $\alpha S \subseteq s\alpha\mathbf{B}_1(0) = s\mathbf{B}_{|\alpha|}(0) = (s|\alpha|)\mathbf{B}_1(0)$.

(b) Let $\varepsilon > 0$. Since S and T are totally bounded, they can each be covered by finitely many open balls of radius $\varepsilon/2$:

$$S \subseteq \bigcup_{n=1}^{N} \mathbf{B}_{\varepsilon/2}(s_n)$$
$$T \subseteq \bigcup_{m=1}^{M} \mathbf{B}_{\varepsilon/2}(t_m),$$

but then

$$S+T \subseteq \bigcup_{n=1}^{N} \mathbf{B}_{\varepsilon/2}(s_n) + \bigcup_{m=1}^{M} \mathbf{B}_{\varepsilon/2}(t_m) = \bigcup_{n=1}^{N}\bigcup_{m=1}^{M} \big(\mathbf{B}_{\varepsilon/2}(s_n) + \mathbf{B}_{\varepsilon/2}(t_m)\big) = \bigcup_{n=1}^{N}\bigcup_{m=1}^{M} \mathbf{B}_{\varepsilon}(s_n + t_m).$$

For αS, note that S can be covered by finitely many open balls of radius $\varepsilon/|\alpha|$:

$$S \subseteq \bigcup_{n=1}^{N} \mathbf{B}_{\varepsilon/|\alpha|}(s_n),$$

so that

$$\alpha S \subseteq \bigcup_{n=1}^{N} \alpha\mathbf{B}_{\varepsilon/|\alpha|}(s_n) = \bigcup_{n=1}^{N} \mathbf{B}_{\varepsilon}(s_n).$$

(c) Consider the addition map $a\colon V \times V \longrightarrow V$, $a(v,w) = v+w$. We know that it is continuous, so its restriction

$$a|_{S\times T}\colon S \times T \longrightarrow V, \qquad a(s,t) = s+t$$

is also continuous, and its image is $S+T$. Since S and T are compact, so is $S \times T$, and so is $S+T = a(S\times T)$.

The same argument with scalar multiplication gives compactness of αS. □

Exercise 3.8. Let $f\colon V \longrightarrow W$ is a linear transformation between vector spaces.

(a) If U is a subspace of V, then its image $f(U)$ is a subspace of W.

(b) If U is a subspace of W, then its preimage $f^{-1}(U)$ is a subspace of V.

(c) If S is a convex subset of V, then its image $f(S)$ is a convex subset of W.

(d) If S is a convex subset of W, then its preimage $f^{-1}(S)$ is a convex subset of V.

Solution.

(a) If w_1 and w_2 are vectors in $f(U)$, then there exists vectors v_1 and v_2 in U such that $w_1 = f(v_1)$ and $w_2 = f(v_2)$. Since U is a vector space, it follows that $v_1 + v_2 \in U$, so

$$w_1 + w_2 = f(v_1) + f(v_2) = f(v_1 + v_2) \in f(U).$$

If α is a scalar and w is a vector in $f(U)$, then there exists a vector v in U such that $w = f(v)$. Since U is a vector space, it follows that $\alpha v \in U$, so

$$\alpha w = \alpha f(v) = f(\alpha v) \in f(U).$$

(b) If v_1 and v_2 are vectors in $f^{-1}(U)$, then

$$f(v_1 + v_2) = f(v_1) + f(v_2) \in f(U)$$

because U is a vector space and both $f(v_1)$ and $f(v_2)$ belong to U.
If α is a scalar and v is a vector in $f^{-1}(U)$, then

$$f(\alpha v) = \alpha f(v) \in f(U)$$

because U is a vector space and $f(v)$ belongs to U.

(c) Let $f(s), f(t) \in f(S)$ and let $a, b \geqslant 0$ such that $a + b = 1$. We have

$$af(s) + bf(t) = f(as + bt) \in f(S),$$

where we used the convexity of S to conclude that $as + bt \in S$.

(d) Let $u, v \in f^{-1}(S)$ and let $a, b \geqslant 0$ such that $a + b = 1$. Then

$$f(au + bv) = af(u) + bf(v) \in S,$$

where we used the convexity of S. We conclude that $au + bv \in f^{-1}(S)$. □

Exercise 3.9. For any $n \in \mathbf{N}$, give a linear distance-preserving map $\mathbf{F}^n \longrightarrow \ell^2$. (Take the Euclidean norm on $\mathbf{F}^n$.)

Solution. Consider $f : \mathbf{F}^n \longrightarrow \ell^2$ given by

$$f(a) = f(a_1, a_2, \ldots, a_n) = (a_1, a_2, \ldots, a_n, 0, 0, \ldots).$$

We have

$$\left\|(a_1, a_2, \ldots, a_n, 0, 0, \ldots)\right\|_{\ell^2} = \left(\sum_{k=1}^{n} |a_k|^2\right)^{1/2} = \|(a_1, a_2, \ldots, a_n)\|_{\mathbf{F}^n},$$

so $f(a) \in \ell^2$, and f is distance-preserving.
Linearity is straightforward. □

Exercise 3.10. Consider the maps $H_{\text{even}}, H_{\text{odd}}\colon \mathbf{F}^{\mathbf{N}} \longrightarrow \mathbf{F}^{\mathbf{N}}$ defined by

$$H_{\text{even}}\big((a_n)\big) = \big(a_{2n}\big), \qquad H_{\text{odd}}\big((a_n)\big) = \big(a_{2n-1}\big)$$

and construct $f\colon \mathbf{F}^{\mathbf{N}} \longrightarrow \mathbf{F}^{\mathbf{N}} \times \mathbf{F}^{\mathbf{N}}$ as

$$f(a) = \big(H_{\text{even}}(a), H_{\text{odd}}(a)\big).$$

(a) Prove that the restriction of H_{even} and H_{odd} to ℓ^p gives continuous linear functions $H_{\text{even}}, H_{\text{odd}}\colon \ell^p \longrightarrow \ell^p$ for all $p \in \mathbf{R}_{\geqslant 1}$ and for $p = \infty$.

(b) Prove that f is an invertible linear map.

In the next two parts, recall that on the product $V \times W$ of two normed spaces we can work with the norm given by

$$\|(v, w)\| := \|v\|_V + \|w\|_W.$$

(c) Take $p = 1$ and show that the restriction $f\colon \ell^1 \longrightarrow \ell^1 \times \ell^1$ is a linear isometry.

(Recall that we can work with the norm on $\ell^1 \times \ell^1$ given by

$$\|(x, y)\| := \|x\|_{\ell^1} + \|y\|_{\ell^1}.)$$

(d) Show that the statement from part (c) does not hold for the space ℓ^∞; prove the strongest statement that you can for ℓ^∞.

Solution. (a) Linearity is straightforward, even on all of $\mathbf{F}^{\mathbf{N}}$:

$$\begin{aligned} H_{\text{even}}\big(\lambda a + \mu b\big) &= H_{\text{even}}\big((\lambda a_n + \mu b_n)\big) \\ &= \big(\lambda a_{2n} + \mu b_{2n}\big) \\ &= \lambda\big(a_{2n}\big) + \mu\big(b_{2n}\big) \\ &= \lambda H_{\text{even}}(a) + \mu H_{\text{even}}(b) \end{aligned}$$

and similarly for H_{odd}.

If $a = (a_n) \in \ell^p$ then

$$\big\|H_{\text{even}}(a)\big\|_{\ell^p}^p = \sum_{n=1}^{\infty} |a_{2n}|^p \leqslant \sum_{n=1}^{\infty} |a_n|^p = \|a\|_{\ell^p}^p,$$

so $H_{\text{even}}(a) \in \ell^p$ and $H_{\text{even}}\colon \ell^p \longrightarrow \ell^p$ is continuous. The same argument works for H_{odd}.

Similarly, if $a = (a_n) \in \ell^\infty$ then

$$\big\|H_{\text{even}}\big\|_{\ell^\infty} = \sup_{n \in \mathbf{N}} |a_{2n}| \leqslant \sup_{n \in \mathbf{N}} |a_n| = \|a\|_{\ell^\infty}$$

and the same for H_{odd}.

(b) The map f is linear because its two components are linear.

We construct an explicit inverse $g\colon \mathbf{F}^{\mathbf{N}} \times \mathbf{F}^{\mathbf{N}} \longrightarrow \mathbf{F}^{\mathbf{N}}$: given $b, c \in \mathbf{F}^{\mathbf{N}}$, define

$$g(b,c) := a := (a_n) \in \mathbf{F}^{\mathbf{N}} \qquad \text{by} \qquad a_n = \begin{cases} b_{n/2} & \text{if } n \text{ is even} \\ c_{(n+1)/2} & \text{if } n \text{ is odd.} \end{cases}$$

It is clear that g is the inverse of f.

(c) We have

$$\begin{aligned} \|f(a)\| &= \big\|(H_{\text{even}}(a), H_{\text{odd}}(a))\big\| \\ &= \big\|H_{\text{even}}(a)\big\|_{\ell^1} + \big\|H_{\text{odd}}(a)\big\|_{\ell^1} \\ &= \sum_{n=1}^{\infty} |a_{2n}| + \sum_{n=1}^{\infty} |a_{2n-1}| \\ &= \sum_{n=1}^{\infty} |a_n| \\ &= \|a\|_{\ell^1}, \end{aligned}$$

so that f is a distance-preserving map.

To prove surjectivity of f, we show that the restriction of the function g from part (b) maps to ℓ^1: for $b, c \in \ell^1$, we have $a := g(b,c)$.

The fact that $a \in \ell^1$ follows from

$$\sum_{n=1}^{2m} |a_n| = \sum_{k=1}^{m} |a_{2k}| + \sum_{k=1}^{m} |a_{2k-1}| = \sum_{k=1}^{m} |b_k| + \sum_{k=1}^{m} |c_k|.$$

As $b, c \in \ell^1$, the limit of the RHS as $m \longrightarrow \infty$ exists and equals $\|b\|_{\ell^1} + \|c\|_{\ell^1}$, so $a \in \ell^1$, $f(a) = (b,c)$, and (of course) $\|a\|_{\ell^1} = \|(b,c)\|$.

(d) We try to use the same approach as in (b):

$$\begin{aligned} \|f(a)\| &= \big\|(H_{\text{even}}(a), H_{\text{odd}}(a))\big\| \\ &= \big\|H_{\text{even}}(a)\big\|_{\ell^\infty} + \big\|H_{\text{odd}}(a)\big\|_{\ell^\infty} \\ &= \sup_{n\in\mathbf{N}} |a_{2n}| + \sup_{n\in\mathbf{N}} |a_{2n-1}| \\ &\leqslant \sup_{n\in\mathbf{N}} |a_n| + \sup_{n\in\mathbf{N}} |a_n| \\ &= 2\|a\|_{\ell^\infty}, \end{aligned}$$

which shows that f is continuous.

It also indicates that f is not distance-preserving: take $(a) = (1, 1, \dots)$ then

$$\|f(a)\| = 2 \neq 1 = \|a\|_{\ell^\infty}.$$

So far we know that f is linear and continuous. It is also injective because it is the restriction of the injective map from part (b).

To prove surjectivity, we show that the restriction of the function g from part (b) maps to ℓ^∞: for $b, c \in \ell^\infty$, we have $a := g(b,c)$. But

$$\sup_{n\in\mathbf{N}} |a_n| = \sup\left\{\sup_{n\in\mathbf{N}} |a_{2n}|, \sup_{n\in\mathbf{N}} |a_{2n-1}|\right\} = \sup\left\{\|b\|_{\ell^\infty}, \|c\|_{\ell^\infty}\right\},$$

which is finite because it is the maximum of two finite quantities.

Finally, the last equation tells us that

$$\|g(b,c)\| = \|a\| = \sup\left\{\|b\|_{\ell^\infty}, \|c\|_{\ell^\infty}\right\} \leqslant \|b\|_{\ell^\infty} + \|c\|_{\ell^\infty} = \|(b,c)\|,$$

so g is also a continuous function.

We conclude that f is a linear homeomorphism. □

Exercise 3.11. Consider the map $f\colon \ell^1 \longrightarrow \mathbf{F}^{\mathbf{N}}$ given by

$$f\big((a_n)\big) = \left(\frac{a_n}{n}\right).$$

(a) Prove that f maps to ℓ^1 and $f\colon \ell^1 \longrightarrow \ell^1$ is linear, continuous, and injective.

(b) Prove that the image W of f is not closed in ℓ^1.

Solution. (a) For all $n \in \mathbf{N}$ we have

$$\left|\frac{a_n}{n}\right| \leqslant |a_n|,$$

so that for $m \in \mathbf{N}$:

$$\sum_{n=1}^{m}\left|\frac{a_n}{n}\right| \leqslant \sum_{n=1}^{m}|a_n|.$$

As $(a_n) \in \ell^1$, the RHS has a finite limit as $m \longrightarrow \infty$, hence so does the LHS, so $f\big((a_n)\big) \in \ell^1$.

Linearity is clear:

$$\begin{aligned} f\big(\lambda(a_n) + \mu(b_n)\big) &= f\big((\lambda a_n + \mu b_n)\big) \\ &= \left(\frac{\lambda a_n + \mu b_n}{n}\right) \\ &= \lambda\left(\frac{a_n}{n}\right) + \mu\left(\frac{b_n}{n}\right) \\ &= \lambda f\big((a_n)\big) + \mu f\big((b_n)\big). \end{aligned}$$

We've seen already that $\big\|f\big((a_n)\big)\big\|_{\ell^1} \leqslant \big\|(a_n)\big\|_{\ell^1}$, so f is continuous.

Suppose $f\big((a_n)\big) = f\big((b_n)\big)$, then for all $n \in \mathbf{N}$ we have $a_n/n = b_n/n$, therefore $a_n = b_n$. So f is injective.

(b) For each $n \in \mathbf{N}$ let $v_n = \big(1, 1/2, \ldots, 1/n, 0, 0, \ldots\big) \in \mathbf{F}^{\mathbf{N}}$. Since v_n has only finitely many nonzero terms, it is in ℓ^1. Letting $w_n = f(v_n)$, we have $w_n \in W$.

Set

$$w = \left(1, \frac{1}{2^2}, \frac{1}{3^2}, \dots\right).$$

Since

$$\sum_{n=1}^{\infty} \frac{1}{n^2}$$

converges, we have $w \in \ell^1$.

However, $w \notin W$: if $w \in W$ then $w = f(v)$ where $v = (1, 1, \dots)$, but $v \notin \ell^1$.

Finally

$$\|w - w_n\|_{\ell^1} = \left\|(0, 0, \dots, 0, \frac{1}{(n+1)^2}, \frac{1}{(n+2)^2}, \dots\right\|_{\ell^1} = \sum_{k=n+1}^{\infty} \frac{1}{k^2},$$

which is the tail of a convergent series, hence converges to 0. Therefore $(w_n) \longrightarrow w$, but $w \notin W$, so W is not closed in ℓ^1. □

Exercise 3.12. **(*)** Let U, V, W be normed spaces over $\mathbf{F}$ and let $\beta \colon U \times V \longrightarrow W$ be a bilinear map.

We say that β is *Lipschitz* if there exists $c > 0$ such that

$$\|\beta(u, v)\|_W \leqslant c\, \|u\|_U\, \|v\|_V \qquad \text{for all } u \in U, v \in V.$$

Prove that β is continuous at $(0, 0)$ if and only if β is Lipschitz if and only if β is continuous on $U \times V$.

Solution. Suppose β is continuous at $(0, 0)$ but not Lipschitz. Then for every $n \in \mathbf{N}$ there exist vectors $u_n \in U$ and $v_n \in V$ such that

$$\|\beta(u_n, v_n)\|_W > n^2\, \|u_n\|_U\, \|v_n\|_V.$$

This forces u_n, v_n to be nonzero. Let

$$u'_n = \frac{1}{n\,\|u_n\|_U}\, u_n \text{ and } v'_n = \frac{1}{n\,\|v_n\|_V}\, v_n.$$

We now prove $(u'_n, v'_n) \longrightarrow (0, 0)$ but $\beta(u'_n, v'_n) \not\longrightarrow 0 = \beta(0, 0)$ as $n \longrightarrow \infty$, which contradicts the continuity of β.

Since $\|u'_n\|_U = \|v'_n\|_V = 1/n$, it follows that

$$\|(u'_n, v'_n)\|_{U \times V} = \|u'_n\|_U + \|v'_n\|_V = \frac{1}{2n}.$$

Therefore, $\|(u'_n, v'_n)\| \longrightarrow 0$ and thus $(u'_n, v'_n) \longrightarrow (0, 0)$ as $n \longrightarrow \infty$.

On the other hand, we have

$$\|\beta(u'_n, v'_n)\|_W = \left\|\beta\left(\frac{1}{n\,\|u_n\|_U}\, u_n, \frac{1}{n\,\|v_n\|_V}\, v_n\right)\right\|_W = \frac{\|\beta(u_n, v_n)\|_W}{n^2\, \|u_n\|_U\, \|v_n\|_V} > 1.$$

Hence $\beta(u'_n, v'_n) \not\longrightarrow 0$ as $n \longrightarrow \infty$.

Now suppose β is Lipschitz; we prove that it is continuous at any $(u,v) \in U \times V$. Given $\varepsilon > 0$, let

$$\delta = \min\left\{1, \frac{\varepsilon}{2c(\|u\|_U + 1)}, \frac{\varepsilon}{2c(\|v\|_V + 1)}\right\}.$$

If $(u',v') \in \mathbf{B}_\delta(u,v)$, then

$$\|u' - u\|_U + \|v' - v\|_V = \|(u' - u, v' - v)\|_{U\times V} = \|(u',v') - (u,v)\|_{U\times V} < \delta$$

and it follows that $\|u' - u\| < \delta$ and $\|v' - v\| < \delta$. Now we have

$$\begin{aligned}
\|\beta(u',v') - \beta(u,v)\|_W &= \|\beta(u',v') - \beta(u',v) + \beta(u',v) - \beta(u,v)\|_W \\
&= \|\beta(u',v'-v) + \beta(u'-u,v)\|_W \\
&\leqslant \|\beta(u',v'-v)\|_W + \|\beta(u'-u,v)\|_W \\
&\leqslant c\,\|u'\|_U\,\|v'-v\|_V + c\,\|u'-u\|_U\,\|v\|_V \\
&\leqslant c\big(\|u\|_U + \|u'-u\|_U\big)\|v'-v\|_V + c\,\|u'-u\|_U\,\|v\|_V \\
&< c\,(\|u\|_U + 1)\delta + c\delta\|v\|_V \\
&\leqslant c(\|u\|_U + 1)\frac{\varepsilon}{2c(\|u\|_U + 1)} + c\|v\|_V\frac{\varepsilon}{2c(\|v\|_V + 1)} \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\
&= \varepsilon.
\end{aligned}$$

Therefore, $\mathbf{B}_\delta(u,v) \subseteq \beta^{-1}(\mathbf{B}_\varepsilon(\beta(u,v)))$ and thus β is continuous.

Obviously, if β is continuous on $U \times V$ then it is continuous at $(0,0)$, closing the cycle of equivalences. □

Exercise 3.13. (*) Let U, V, W be nonzero normed spaces over $\mathbf{F}$ and let $\beta\colon U\times V \longrightarrow W$ be a nonzero bilinear map. Then β is **not** uniformly continuous.

Solution. Since U, V, W are nonzero and β is nonzero, there exist vectors $u \in U$ and $v \in V$ such that $\beta(u,v) \neq 0$. This forces u and v to be nonzero.

Take $\varepsilon = 1$. Given $\delta > 0$, put

$$a = \frac{\delta}{2\|u\|_U}, \qquad b = \frac{3\|u\|_U}{\delta\,\|\beta(u,v)\|_W}.$$

It follows that

$$\big\|(0,bv) - (au,bv)\big\|_{U\times V} = \big\|(-au,0)\big\|_{U\times V} = a\|u\|_U = \frac{\delta}{2} < \delta,$$

but

$$\big\|\beta(0,bv) - \beta(au,bv)\big\|_W = \big\|\beta(-au,bv)\big\|_W = ab\big\|\beta(u,v)\big\|_W = \frac{3}{2} > 1 = \varepsilon.$$

Therefore, β is not uniformly continuous.

(In fact, the proof shows that β is not even uniformly continuous on the subspace $\mathbf{F}u \times \mathbf{F}v \subseteq U \times V$.) □

Exercise 3.14. (*) Let U, V, W be normed spaces over $\mathbf{F}$.
Suppose $\beta\colon U \times V \longrightarrow W$ is a continuous bilinear map.
Consider the linear function $\beta_U\colon U \longrightarrow \operatorname{Hom}(V, W)$ given by $\beta_U(u) = f_u$, where

$$f_u\colon V \longrightarrow W \qquad \text{is defined by } f_u(v) = \beta(u, v).$$

(a) Prove that for any $u \in U$, $f_u \in L(V, W)$, in other words f_u is continuous.

(b) By part (a) we can think of β_U as a function $U \longrightarrow L(V, W)$.
Prove that $\beta_U\colon U \longrightarrow L(V, W)$ is continuous.

Solution.

(a) **First approach (direct):** Let $v \in V$. We prove that $f_u\colon V \longrightarrow W$ is continuous at v. (Note that, crucially, u remains fixed.)

Let $\varepsilon > 0$; as β is continuous at (u, v), there exists $\delta > 0$ such that

$$\text{if } \|(u, v_1) - (u, v)\|_{U\times V} < \delta, \text{ then } \|\beta(u, v_1) - \beta(u, v)\|_W < \varepsilon.$$

Therefore, if $\|v_1 - v\|_V < \delta$, then

$$\|(u, v_1) - (u, v)\|_{U\times V} = \|v_1 - v\|_V < \delta,$$

so that

$$\|f_u(v_1) - f_u(v)\|_W = \|\beta(u, v_1) - \beta(u, v)\|_W < \varepsilon.$$

Second approach (using Lipschitz): Let $\varepsilon > 0$; as β is continuous, it is Lipschitz, so there exists $c > 0$ such that

$$\|\beta(u, v)\|_W \leqslant c\,\|u\|_U\,\|v\|_V \qquad \text{for all } u \in U, v \in V.$$

It follows that

$$\|f_u(v)\|_W = \|\beta(u, v)\|_W \leqslant c\,\|u\|_U\,\|v\|_V.$$

Since $c\|u\|_U$ is a constant independent of v, the linear transformation f_u is Lipschitz and thus continuous.

(b) Let $\varepsilon > 0$; as β is continuous, it is Lipschitz, so there exists $c > 0$ such that

$$\|\beta(u, v)\|_W \leqslant c\,\|u\|_U\,\|v\|_V \qquad \text{for all } u \in U, v \in V.$$

It follows that

$$\|\beta_U(u)\|_{L(V,W)} = \|f_u\|_{L(V,W)} = \sup_{\|v\|_V=1} \|\beta(u, v)\|_W \leqslant c\,\|u\|_U.$$

Therefore, β_U is Lipschitz and thus continuous. □

Exercise 3.15. Prove directly that any Cauchy sequence in ℓ^∞ converges, so that ℓ^∞ is a Banach space.

Solution. Let $(x^{(n)})$ be a Cauchy sequence in ℓ^∞. Each element $x^{(n)}$ is a sequence

$x^{(n)} = (x_k^{(n)})$ in $\mathbf{F}$. For $\varepsilon > 0$, there exists $N \in \mathbf{N}$ such that

$$\|x^{(m)} - x^{(n)}\|_{\ell^\infty} < \frac{\varepsilon}{2} \qquad \text{for all } m, n \geqslant N.$$

Fixing $k \in \mathbf{N}$, consider the sequence $(x_k^{(n)})$ (as n varies) in $\mathbf{F}$. It is Cauchy since

$$|x_k^{(m)} - x_k^{(n)}| \leqslant \|x^{(m)} - x^{(n)}\|_{\ell^\infty} < \frac{\varepsilon}{2}.$$

As $\mathbf{F}$ is complete, $(x_k^{(n)})$ has some limit $y_k \in \mathbf{F}$.

Set $y = (y_k)$. It remains to prove that $y \in \ell^\infty$ and that $(x^{(n)}) \longrightarrow y$ in ℓ^∞.

As $(x^{(n)})$ is a Cauchy sequence in ℓ^∞, it is bounded in ℓ^∞, so there exists a constant C such that

$$\|x^{(n)}\|_{\ell^\infty} \leqslant C \qquad \text{for all } n \in \mathbf{N}.$$

Therefore

$$|x_k^{(n)}| \leqslant \|x^{(n)}\|_{\ell^\infty} \leqslant C \qquad \text{for all } k, n \in \mathbf{N}.$$

As we take the limit as $n \longrightarrow \infty$ we get

$$|y_k| \leqslant C \qquad \text{for all } k \in \mathbf{N},$$

in other words $y = (y_k) \in \ell^\infty$.

Let $\varepsilon > 0$ and $N \in \mathbf{N}$ be as above. I claim that

$$|x_k^{(n)} - y_k| < \varepsilon \qquad \text{for all } n \geqslant \mathbf{N}, k \in \mathbf{N}.$$

Let $k \in \mathbf{N}$. As $(x_k^{(m)}) \longrightarrow y_k$ as $m \longrightarrow \infty$, we can choose $m \geqslant N$ large enough that

$$|x_k^{(m)} - y_k| < \frac{\varepsilon}{2}.$$

Therefore, given any $n \geqslant N$ we have

$$|x_k^{(n)} - y_k| \leqslant |x_k^{(n)} - x_k^{(m)}| + |x_k^{(m)} - y_k| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

The conclusion holds for all $k \in \mathbf{N}$, so we are done. □

Exercise 3.16. In Theorem 3.35 we saw that the function

$$\beta \colon \ell^\infty \times \ell^1 \longrightarrow \mathbf{F} \qquad \text{defined by } \beta(u, v) \longmapsto \sum_{n=1}^{\infty} u_n v_n$$

is a continuous bilinear map.

Show that there is a continuous linear function $\ell^\infty \longrightarrow (\ell^1)^\vee$ that is an isometry.

Conclude that ℓ^∞ is a Banach space.

Solution. By Exercise 3.12, $\beta_U \colon \ell^\infty \longrightarrow \left(\ell^1\right)^\vee$ is linear and continuous, where

$$\beta_U(u) = u^\vee, \qquad u^\vee(v) = \beta(u, v) = \sum_{n=1}^{\infty} u_n v_n.$$

To see that $u \longmapsto u^\vee$ is surjective, let $\varphi \in \left(\ell^1\right)^\vee$. Since ℓ^1 has Schauder basis $\{e_1, e_2, \ldots\}$, for any $v = (v_n) \in \ell^1$ we have

$$v = \sum_{n=1}^\infty v_n e_n,$$

so that

$$\varphi(v) = \sum_{n=1}^\infty v_n \varphi(e_n).$$

Setting $u_n = \varphi(e_n)$ and $u = (u_n)$, if we show that $u \in \ell^\infty$ then $\varphi = u^\vee$. But since $\varphi \in \left(\ell^1\right)^\vee = L(\ell^1, \mathbf{F})$, it is Lipschitz, so for all $v \in \ell^1$ we have

$$|\varphi(v)| \leqslant \|\varphi\| \, \|v\|_{\ell^1}.$$

In particular, for all $n \in \mathbf{N}$ we get

$$|u_n| = |\varphi(e_n)| \leqslant \|\varphi\|,$$

hence $u \in \ell^\infty$, and also $\|u\|_{\ell^\infty} \leqslant \|\varphi\| = \|u^\vee\|$.

Hölder's Inequality gives us

$$|u^\vee(v)| \leqslant \sum_{n=1} |u_n v_n| \leqslant \|u\|_{\ell^\infty} \|v\|_{\ell^1},$$

so for $v \in \ell^1 \setminus \{0\}$ we get

$$\frac{|u^\vee(v)|}{\|v\|_{\ell^1}} \leqslant \|u\|_{\ell^\infty},$$

so $\|u^\vee\| \leqslant \|u\|_{\ell^\infty}$.

As we had already established the opposite inequality, we conclude that $\|u^\vee\| = \|u\|_{\ell^\infty}$.

Since ℓ^∞ is isometric to $\left(\ell^1\right)^\vee$ and all dual spaces as Banach, ℓ^∞ is Banach. □

Exercise 3.17. Flip the factors in Exercise 3.16:

In Theorem 3.35 we saw that the function

$$\ell^1 \times \ell^\infty \longrightarrow \mathbf{F} \qquad \text{defined by } (u, v) \longmapsto \sum_{n=1}^\infty u_n v_n$$

is a continuous bilinear map.

(a) Show that there is a continuous linear function $\ell^1 \longrightarrow (c_0)^\vee$ that is an isometry. (Recall that $c_0 \subseteq \ell^\infty$ consists of all convergent sequences with limit 0.)

[*Hint*: It may be useful to prove surjectivity first, and then the distance-preserving property.]

(b) Conclude that ℓ^1 is a Banach space.

(c) Where in your proof for (a) did you make use of the fact that you are working with c_0 rather than ℓ^∞?

Solution.

(a) If we restrict the bilinear map from the statement to $\ell^1 \times c_0$, we get a continuous bilinear map

$$\beta\colon \ell^1 \times c_0 \longrightarrow \mathbf{F}.$$

By Exercise 3.12, β_U is linear and continuous. In our notation, this is the function $u \longmapsto u^\vee : \ell^1 \longrightarrow (c_0)^\vee$, where

$$u^\vee(v) = \beta(u,v) = \sum_{n=1}^{\infty} u_n v_n.$$

For surjectivity, we need to show that each $\varphi \in (c_0)^\vee$ is of the form $\varphi = u^\vee$ for some $u \in \ell^1$. Take such φ. Recall that c_0 has Schauder basis $\{e_1, e_2, \dots\}$, so for any $v = (v_n) \in c_0$ we have

$$\varphi(v) = \sum_{n=1}^{\infty} v_n \varphi(e_n).$$

Let $u_n = \varphi(e_n)$ and $u = (u_n)$. We need to show that $u \in \ell^1$. For this, fix $m \in \mathbf{N}$ and let (ignoring the n's for which $u_n = 0$)

$$x = \sum_{n=1}^{m} \frac{|u_n|}{u_n} e_n = \left(\frac{|u_1|}{u_1}, \dots, \frac{|u_m|}{u_m}, 0, 0, \dots\right),$$

so that

$$\|x\|_{\ell^\infty} = 1.$$

Then

$$\begin{aligned}
\sum_{n=1}^{m} |u_n| &= \left|\sum_{n=1}^{m} \frac{|u_n|}{u_n} u_n\right| \\
&= \left|\sum_{n=1}^{m} \varphi\left(\frac{|u_n|}{u_n} e_n\right)\right| \\
&= |\varphi(x)| \leqslant \|\varphi\| \, \|x\|_{\ell^\infty} = \|\varphi\|.
\end{aligned}$$

Taking the limit as $m \longrightarrow \infty$ we conclude that $u \in \ell^1$ and that $\|u\|_{\ell^1} \leqslant \|\varphi\| = \|u^\vee\|$.

So $u \longmapsto u^\vee$ is surjective.

We have the Hölder Inequality

$$\sum_{n=1}^{\infty} |u_n v_n| \leqslant \|u\|_{\ell^1} \|v\|_{\ell^\infty},$$

valid for all $u \in \ell^1$ and all $v \in \ell^\infty$, so certainly for all $v \in c_0$.

Hence for $v \neq 0$:

$$\frac{|u^\vee(v)|}{\|v\|_{\ell^\infty}} \leqslant \|u\|_{\ell^1},$$

so taking supremum we get $\|u^\vee\| \leqslant \|u\|_{\ell^1}$.

As we had already established the other inequality, we conclude that $\|u^\vee\| = \|u\|_{\ell^1}$, so $u \longmapsto u^\vee$ is distance-preserving.

Putting it all together, we have a linear isometry $\ell^1 \longrightarrow (c_0)^\vee$.

(b) We know that duals of normed spaces are complete, so $(c_0)^\vee$ is complete, so ℓ^1, being isometric to it, also is complete.

(c) We used the Schauder basis $\{e_1, e_2, \dots\}$ for c_0 to prove surjectivity as well as the distance-preserving property. □

Exercise 3.18. Consider the subset c of $\mathbf{F}^{\mathbf{N}}$ consisting of all convergent sequences (with any limit).

(a) Convince yourself that c is a vector subspace of ℓ^∞.

(b) Prove that $\lim\colon c \longrightarrow \mathbf{F}$ given by

$$(a_n) \longmapsto \lim_{n\longrightarrow\infty} (a_n)$$

is a continuous surjective linear map.

(c) Prove that the formula

$$J\big((a_n)\big) = R\big((a_n)\big) - \Big(\lim_{n\longrightarrow\infty} a_n\Big)(1, 1, \dots)$$

defines a linear homeomorphism $J\colon c \longrightarrow c_0$. (Here R denotes the right shift map.)

(d) Conclude that c is Banach.

[*Hint*: Tutorial Question 9.6 should come in handy here and in the following part.]

(e) Show that c is separable and find a Schauder basis for c.

Solution. (a) We know that convergent sequences are bounded, so $c \subseteq \ell^\infty$. We also know that the sum of two convergent sequences is convergent, and that a scalar multiple of a convergent sequence is convergent, and that the constant sequence $(0, 0, \dots)$ is convergent.

(b) We know that lim is linear, as a consequence of the continuity of addition and of scalar multiplication.

It is certainly surjective, as given any $a \in \mathbf{F}$ the constant sequence $(a, a, \dots)$ converges to a.

Finally, if $a = (a_n) \in c$ then (a_n) is a bounded sequence and

$$\Big|\lim_{n\longrightarrow\infty} a_n\Big| \leqslant \sup_{n\in\mathbf{N}} |a_n| = \|a\|_{\ell^\infty},$$

so lim is a continuous linear map.

(c) It is clear that J is linear and continuous, as R and lim are linear and continuous.

We exhibit an explicit inverse of J: let $K\colon c_0 \longrightarrow c$ be given by

$$K\big((b_n)\big) = L\big((b_n)\big) - b_1(1, 1, \dots).$$

Note that K is linear and continuous, as L and $(b_n) \longmapsto b_1$ are linear and continuous.

We check that K and J and inverses. If $b \in c_0$ and $a \in c$ then:

$$\begin{aligned} J\big(K(b)\big) &= J\big(L(b)\big) - b_1 J(1,1,\dots) \\ &= R\big(L(b)\big) - 0(1,1,\dots) - b_1\big(R(1,1,\dots) - (1,1,\dots)\big) \\ &= (0, b_2, b_3, \dots) - b_1(-1,0,0,\dots) \\ &= b, \\ K\big(J(a)\big) &= K\big(R(a)\big) - \big(\lim a_n\big) K(1,1,\dots) \\ &= L\big(R(a)\big) - \big(\lim a_n\big)\big(L(1,1,\dots) - (1,1,\dots)\big) \\ &= a. \end{aligned}$$

(d) We know from Tutorial Question 9.6 that c_0 is closed in ℓ^∞, so c must also be closed in ℓ^∞ as it is homeomorphic to c_0. But ℓ^∞ is complete, so c is complete.

(e) We know that $\{e_1, e_2, e_3, \dots\}$ is a Schauder basis for c_0, so we apply $K\colon c_0 \longrightarrow c$ to this to get:

$$\begin{aligned} K(e_1) &= L(e_1) - (1,1,\dots) = -(1,1,\dots) \\ K(e_2) &= L(e_2) - 0(1,1,\dots) = e_1 \\ K(e_3) &= L(e_3) - 0(1,1,\dots) = e_2 \\ &\vdots \\ K(e_n) &= L(e_n) - 0(1,1,\dots) = e_{n-1} \qquad \text{for } n \geqslant 2 \\ &\vdots \end{aligned}$$

We suspect then that $\{(1,1,\dots), e_1, e_2, e_3, \dots\}$ is a Schauder basis for c.

This is of course true whenever we have a linear homeomorphism $f\colon V \longrightarrow W$ between normed spaces: If $\{b_1, b_2, \dots\}$ is a Schauder basis for V, then $\{f(b_1), f(b_2), \dots\}$ is a Schauder basis for W.

Let $w \in W$ and let $v = f^{-1}(w) \in V$. Write

$$v = \sum_{j \in \mathbf{N}} \alpha_j b_j,$$

then

$$w = f(v) = \sum_{j \in \mathbf{N}} \alpha_j f(b_j).$$

Uniqueness follows from the uniqueness of the expansion for v. □

Exercise 3.19. Consider the left shift map $L\colon \mathbf{F}^{\mathbf{N}} \longrightarrow \mathbf{F}^{\mathbf{N}}$ given by $L\big((a_n)\big) = (a_{n+1})$, that is

$$L(a_1, a_2, a_3, \dots) = (a_2, a_3, \dots).$$

(a) Prove that L is a surjective linear map. What is the kernel of L?

(b) Prove that for all $1 \leqslant p \leqslant \infty$, the restriction of L to ℓ^p is a surjective continuous map onto ℓ^p.

(c) Define the right shift map $R\colon \mathbf{F}^{\mathbf{N}} \longrightarrow \mathbf{F}^{\mathbf{N}}$ and prove that it is an injective linear map, the restriction of which is distance-preserving for any ℓ^p with $1 \leqslant p \leqslant \infty$.

(d) Check that $L \circ R = \mathrm{id}_{\mathbf{F}^{\mathbf{N}}} \neq R \circ L$.

Solution.

(a) It is clear that L is surjective. Linearity is pretty straightforward, and it's also clear that $\ker(L) = \mathrm{Span}\{e_1\}$.

(b) We have

$$\|L(a_1, a_2, a_3, \dots)\|_{\ell^p} = \left(\sum_{n=2}^{\infty} |a_n|^p\right)^{1/p} \leqslant \left(\sum_{n=1}^{\infty} |a_n|^p\right)^{1/p} = \big\|(a_1, a_2, \dots)\big\|_{\ell^p},$$

so L is bounded, and $L\big((a_n)\big) \in \ell^p$ if $(a_n) \in \ell^p$.

For the surjectivity note that if $b = (b_1, b_2, \dots) \in \ell^p$, then

$$b = L(a) \qquad \text{for } a = (0, b_1, b_2, \dots)$$

and $\|a\|_{\ell^p} = \|b\|_{\ell^p}$, so $a \in \ell^p$.

The case of ℓ^∞ is done in a similar way.

(c) To get a linear map we need to set

$$R(a_1, a_2, a_3, \dots) = (0, a_1, a_2, a_3, \dots).$$

Both injectivity and linearity are straightforward.

We have, for $p \geqslant 1$ or $p = \infty$:

$$\|R(a_1, a_2, \dots)\|_{\ell^p} = \|(0, a_1, a_2, \dots)\|_{\ell^p} = \|(a_1, a_2, \dots)\|_{\ell^p},$$

so R is distance-preserving and $R(a) \in \ell^p$ if $a \in \ell^p$.

(d) Clear. For any $a = (a_n) \in \mathbf{F}^{\mathbf{N}}$ we have

$$\begin{aligned} L(R(a)) &= L(R(a_1, a_2, \dots)) = L(0, a_1, a_2, \dots) = (a_1, a_2, \dots) = a, \\ R(L(a)) &= R(L(a_1, a_2, \dots)) = R(a_2, a_3, \dots) = (0, a_2, a_3, \dots) \neq a \text{ unless } a_1 = 0. \end{aligned}$$ □

Exercise 3.20. Let $\big(V, \langle\cdot,\cdot\rangle\big)$ be an inner product space. Prove that the inner product is a continuous function.

Solution. One way is to use the Polarisation Identity and the fact that the norm is continuous.

But we can also proceed more directly: suppose $(x_n, y_n) \longrightarrow (x, y)$, then $(x_n) \longrightarrow x$ and $(y_n) \longrightarrow y$. As (y_n) converges, it is bounded, so there exists $C \geqslant 0$ such that $\|y_n\| \leqslant C$ for all $n \in \mathbf{N}$.

Given $\varepsilon > 0$, let $N \in \mathbf{N}$ be such that

$$\|x_n - x\| < \frac{\varepsilon}{2C} \qquad \text{and} \qquad \|y_n - y\| < \frac{\varepsilon}{2\|x\|} \text{ for all } n \geqslant N.$$

Then

$$\begin{aligned}\left|\langle x_n, y_n\rangle - \langle x, y\rangle\right| &= \left|\langle x_n, y_n\rangle - \langle x, y_n\rangle + \langle x, y_n\rangle - \langle x, y\rangle\right| \\ &= \left|\langle x_n - x, y_n\rangle + \langle x, y_n - y\rangle\right| \\ &\leqslant \left|\langle x_n - x, y_n\rangle\right| + \left|\langle x, y_n - y\rangle\right| \\ &\leqslant \|x_n - x\| \, \|y_n\| + \|x\| \, \|y_n - y\| \\ &\leqslant C\|x_n - x\| + \|x\| \, \|y_n - y\| \\ &< \varepsilon.\end{aligned}$$

We conclude that $\big(\langle x_n, y_n\rangle\big) \longrightarrow \langle x, y\rangle$. □

Exercise 3.21. Let V be an inner product space. For any $v \in V$ we have

$$\|v\| = \sup_{\|w\|=1} \left|\langle v, w\rangle\right|.$$

The supremum is in fact achieved by a well-chosen w.

Solution. If $v = 0$ then the equality is obvious.

So assume now that $v \neq 0$. By Cauchy–Schwarz we have for all $w \in V$:

$$\left|\langle v, w\rangle\right| \leqslant \|v\| \, \|w\|.$$

Therefore for all $w \in V$ with $\|w\| = 1$ we have

$$\left|\langle v, w\rangle\right| \leqslant \|v\|,$$

so that

$$\sup_{\|w\|=1} \left|\langle v, w\rangle\right| \leqslant \|v\|.$$

To get equality, take $w = \frac{1}{\|v\|} v$ and see that the LHS is indeed $\|v\|$. □

Exercise 3.22. Let V, W be inner product spaces and let $f \in L(V, W)$. Prove that

$$\|f\| = \sup_{\|v\|_V = \|w\|_W = 1} \left|\langle f(v), w\rangle_W\right|.$$

[*Hint*: Use Exercise 3.21.]

Solution. Recall from Exercise 3.21 that

$$\|u\|_W = \sup_{\|w\|_W=1} \left|\langle u, w\rangle_W\right| \qquad \text{for all } u \in W.$$

Setting $u = f(v)$ for some $v \in V$, we get

$$\|f(v)\|_W = \sup_{\|w\|_W=1} \left|\langle f(v), w\rangle_W\right| \qquad \text{for all } v \in V.$$

Therefore

$$\|f\| = \sup_{\|v\|_V=1} \|f(v)\|_W = \sup_{\|v\|_V=\|w\|_W=1} \left|\langle f(v), w\rangle_W\right|. \qquad \square$$

Exercise 3.23. Let $(V, \langle\cdot,\cdot\rangle)$ be an inner product space and let R, S be subsets of V.

(a) Prove that $S \cap S^\perp = 0$.

(b) Prove that if $R \subseteq S$ then $S^\perp \subseteq R^\perp$.

(c) Prove that $S \subseteq (S^\perp)^\perp$.

(d) Prove that $S^\perp = \overline{\operatorname{Span}(S)}^\perp$.

Solution.

(a) If $x \in S^\perp \cap S$ then $\langle x, s\rangle = 0$ for all $s \in S$, in particular $\langle x, x\rangle = 0$ so $x = 0$.

(b) Suppose $R \subseteq S$ and $x \in S^\perp$. For any $r \in R$ we have $r \in S$ so $\langle x, r\rangle = 0$, hence $x \in R^\perp$.

(c) Let $s \in S$. For any $x \in S^\perp$, we have

$$\langle s, x\rangle = \overline{\langle x, s\rangle} = 0,$$

so $s \in (S^\perp)^\perp$.

(d) Since $S \subseteq \operatorname{Span}(S) \subseteq \overline{\operatorname{Span}(S)}$, by part (b) we get

$$\overline{\operatorname{Span}(S)}^\perp \subseteq S^\perp.$$

In the other direction, suppose $x \in S^\perp$. For any $v \in \operatorname{Span}(S)$ we have

$$\langle x, v\rangle = \langle x, \alpha_1 s_1 + \cdots + \alpha_n s_n\rangle = \overline{\alpha}_1\langle x, s_1\rangle + \cdots + \overline{\alpha}_n\langle x, s_n\rangle = 0.$$

Now if $(v_n) \longrightarrow w \in \overline{\operatorname{Span}(S)}$ with $v_n \in \operatorname{Span}(S)$, we have

$$\langle x, w\rangle = \langle x, \lim v_n\rangle = \lim\langle x, v_n\rangle = \lim 0 = 0. \qquad \square$$

Exercise 3.24. Let S be a subset of a Hilbert space H. Prove that $\operatorname{Span}(S)$ is dense in H if and only if $S^\perp = 0$.

Solution. If $S^\perp = 0$ then (using the Hilbert Projection Theorem Part II)

$$\overline{\operatorname{Span}(S)} = (S^\perp)^\perp = 0^\perp = H.$$

Conversely, if S is dense in H then

$$S^\perp = \overline{\operatorname{Span}(S)}^\perp = H^\perp = 0. \qquad \square$$

Exercise 3.25. **(*)** Let U, V, W be normed spaces.

Define the norm of a continuous bilinear map $\beta\colon U \times V \longrightarrow W$, and show that it is a norm on the vector space $\mathrm{Bil}(U, V; W)$ of continuous bilinear maps $U \times V \longrightarrow W$.

[*Hint*: Have a look at Exercise 3.12 to remember what it says.]

Solution. Since β is continuous and bilinear, it is Lipschitz: there exists $c > 0$ such that

$$\|\beta(u,v)\|_W \subseteq c\|u\|_U\|v\|_V \qquad \text{for all } u \in U, v \in V.$$

We can then define

$$\|\beta\| := \sup_{u \in U \setminus \{0\}, v \in V \setminus \{0\}} \frac{\|\beta(u,v)\|_W}{\|u\|_U\|v\|_V}.$$

By the bilinearity of β, we have

$$\|\beta\| = \sup_{\|u\|_U = 1, \|v\|_V = 1} \|\beta(u,v)\|_W.$$

The triangle inequality for this norm follows from this last equality and the triangle inequality for $\|\cdot\|_W$. The same is true for the property $\|a\beta\| = |a|\,\|\beta\|$ for all $a \in \mathbf{F}$.

If $\|\beta\| = 0$ then $\|\beta(u,v)\|_W = 0$ for all nonzero u and v, so by the non-degeneracy of $\|\cdot\|_W$ we get $\beta(u,v) = 0$ for all nonzero u and v. The bilinearity of β means that $\beta(0,v) = 0$ and $\beta(u,0) = 0$ for all $u \in U$, $v \in V$, so we conclude that $\beta = 0$. □

Exercise 3.26. Let $V = \mathbf{R}^2$ viewed as a normed space with the Euclidean norm. Compute the norm of each of the following elements $M \in L(V)$ directly from the description of the operator norm:

$$\|M\| = \sup_{\|v\|=1} \|M(v)\|.$$

(a) $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$;

(b) $B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$;

(c) $C = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$ for $a, b \in \mathbf{R}$.

Solution. In all cases we will denote $v = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbf{R}^2$ with $x_1^2 + x_2^2 = 1$.

(a) We have

$$\|Av\| = \left\|\begin{pmatrix} x_2 \\ 0 \end{pmatrix}\right\| = |x_2|.$$

Maximising this under the constraint $x_1^2 + x_2^2 = 1$ gives $\|A\| = 1$.

(b) We have

$$\|Bv\| = \left\|\begin{pmatrix} x_2 \\ -x_1 \end{pmatrix}\right\| = \sqrt{x_2^2 + x_1^2} = 1,$$

so $\|B\| = 1$.

(c) We have

$$\|Cv\| = \left\|\begin{pmatrix} ax_1 \\ bx_2 \end{pmatrix}\right\| = \sqrt{a^2x_1^2 + b^2x_2^2},$$

so we are looking to maximise, under the constraint $x_1^2 + x_2^2 = 1$, the quantity

$$S = a^2x_1^2 + b^2x_2^2 = a^2x_1^2 + b^2\big(1 - x_1^2\big) = b^2 + (a^2 - b^2)x_1^2.$$

If $|a| \geqslant |b|$ then $a^2 - b^2 \geqslant 0$ so to maximise S we must maximise x_1^2, which happens when $x_1^2 = 1$, so that $S = a^2$.

Otherwise we have $|a| < |b|$ so $a^2 - b^2 < 0$ so to maximise S we must minimise x_1^2, which happens when $x_1 = 0$, so that $S = b^2$.

Hence the maximum value is $S = \max\{a^2, b^2\}$ and so $\|C\| = \sqrt{S} = \max\{|a|, |b|\}$. □

For the next few questions, recall that the *adjoint* of a continuous linear map $f\colon X \longrightarrow Y$ of Hilbert spaces is the unique continuous linear map $f^*\colon Y \longrightarrow X$ satisfying

$$\big\langle f(x), y\big\rangle_Y = \big\langle x, f^*(y)\big\rangle_X \qquad \text{for all } x \in X, y \in Y.$$

Exercise 3.27. Prove that $f \longmapsto f^*$ is conjugate-linear, in other words that

$$\big(\alpha f + \beta g\big)^* = \overline{\alpha} f^* + \overline{\beta} g^* \qquad \text{for all } \alpha, \beta \in \mathbf{F}, f, g \in L(X, Y).$$

Solution. We have

$$\begin{aligned}\big\langle x, (\alpha f + \beta g)^*(y)\big\rangle &= \big\langle (\alpha f + \beta g)(x), y\big\rangle \\ &= \alpha\big\langle f(x), y\big\rangle + \beta\big\langle g(x), y\big\rangle \\ &= \alpha\big\langle x, f^*(y)\big\rangle + \beta\big\langle x, g^*(y)\big\rangle \\ &= \big\langle x, \overline{\alpha}\, f^*(y) + \overline{\beta}\, g^*(y)\big\rangle.\end{aligned}$$

□

Exercise 3.28. Prove that $f \longmapsto f^*$ is an involution, in other words that

$$\big(f^*\big)^* = f \qquad \text{for all } f \in L(X, Y).$$

Solution. We have

$$\begin{aligned}\big\langle x, \big(f^*\big)^*(y)\big\rangle &= \big\langle f^*(x), y\big\rangle \\ &= \overline{\big\langle y, f^*(x)\big\rangle} \\ &= \overline{\big\langle f(y), x\big\rangle} \\ &= \big\langle x, f(y)\big\rangle.\end{aligned}$$

□

Exercise 3.29. Let X, Y, Z be Hilbert spaces.

(a) Prove that $\big(f \circ g\big)^* = g^* \circ f^*$ for all $g \in L(X, Y)$, $f \in L(Y, Z)$.

(b) Prove that $\mathrm{id}_X^* = \mathrm{id}_X$.

Solution.

(a) We have

$$\begin{aligned}\big\langle x, (f\circ g)^*(y)\big\rangle &= \big\langle (f\circ g)(x), y\big\rangle \\ &= \big\langle f\big(g(x)\big), y\big\rangle \\ &= \big\langle g(x), f^*(y)\big\rangle \\ &= \big\langle x, g^*\big(f^*(y)\big)\big\rangle \\ &= \big\langle x, \big(g^*\circ f^*\big)(y)\big\rangle.\end{aligned}$$

(b) Tautological:

$$\big\langle \mathrm{id}_X(x), y\big\rangle = \big\langle x, y\big\rangle = \big\langle x, \mathrm{id}_X(y)\big\rangle.$$ □

Exercise 3.30. Let $f \in L(X,Y)$ with X, Y Hilbert spaces.

(a) Prove that $\|f^*\| = \|f\|$, so $f \longmapsto f^*$ is distance-preserving.

(b) Prove that $\|f^* \circ f\| = \|f\|^2$.

Solution.

(a) By Exercise 3.22 we know that

$$\begin{aligned}\|f^*\| &= \sup_{\|x\|=1=\|y\|} \big|\langle f^*(y), x\rangle_X\big| \\ &= \sup_{\|x\|=1=\|y\|} \big|\overline{\langle x, f^*(y)\rangle_X}\big| \\ &= \sup_{\|x\|=1=\|y\|} \big|\langle x, f^*(y)\rangle_X\big| \\ &= \sup_{\|x\|=1=\|y\|} \big|\langle f(x), y\rangle_Y\big| \\ &= \|f\|.\end{aligned}$$

(b) Using again Exercise 3.22 we have

$$\begin{aligned}\big\|f^*\circ f\big\| &= \sup_{\|x\|=1=\|y\|} \big|\big\langle f^*\big(f(x)\big), y\big\rangle\big| \\ &= \sup_{\|x\|=1=\|y\|} \big|\big\langle f(x), f(y)\big\rangle\big| \\ &\leqslant \sup_{\|x\|=1=\|y\|} \big\|f(x)\big\|\,\big\|f(y)\big\| \\ &= \left(\sup_{\|x\|=1} \big\|f(x)\big\|\right)\left(\sup_{\|y\|=1} \big\|f(y)\big\|\right) \\ &= \|f\|^2.\end{aligned}$$

The inequality in the above calculation comes from Cauchy–Schwarz. We note that taking $y = x$ gives an equality, so that equality of suprema actually holds, and we conclude that

$$\|f^* \circ f\| = \|f\|^2. \qquad \square$$

Exercise 3.31. Let $f\colon X \longrightarrow Y$ be a continuous linear map of Hilbert spaces. Prove that

$$\ker\left(f^*\right) = \left(\operatorname{im} f\right)^\perp \qquad \text{and} \qquad \overline{\operatorname{im}\left(f^*\right)} = \left(\ker f\right)^\perp.$$

Solution. We have

$$\begin{aligned} y \in \left(\operatorname{im} f\right)^\perp &\iff y \perp f(x) \qquad \text{for all } x \in X \\ &\iff \left\langle f(x), y\right\rangle = 0 \qquad \text{for all } x \in X \\ &\iff \left\langle x, f^*(y)\right\rangle = 0 \qquad \text{for all } x \in X \\ &\iff f^*(y) = 0 \\ &\iff y \in \ker f^*. \end{aligned}$$

From this and Exercise 3.28 we have

$$\ker f = \ker\left(f^*\right)^* = \left(\operatorname{im} f^*\right)^\perp,$$

so that

$$\left(\ker f\right)^\perp = \left(\left(\operatorname{im} f^*\right)^\perp\right)^\perp = \overline{\operatorname{im} f^*},$$

where the last equality comes from Corollary 3.43. $\square$

Exercise 3.32. Let X be a Hilbert space, $f \in L(X)$, and W a closed subspace of X. Then W is f-invariant if and only if $W^\perp$ is $\left(f^*\right)$-invariant.

Solution. Suppose W is f-invariant. Let $y \in W^\perp$. For any $x \in W$ we have $f(x) \in W$ so that

$$\left\langle x, f^*(y)\right\rangle = \left\langle f(x), y\right\rangle = 0.$$

As this holds for all $x \in W$, we conclude that $f^*(y) \in W^\perp$, so $W^\perp$ is f^*-invariant.

Conversely, suppose $W^\perp$ is f^*-invariant, then by the above

$$W = \left(W^\perp\right)^\perp \qquad \text{is } \left(f^*\right)^* = f\text{-invariant}. \qquad \square$$

Exercise 3.33. Let $a = (a_n) \in \ell^\infty$ and consider $f\colon \ell^2 \longrightarrow \mathbf{F}^{\mathbf{N}}$ given by

$$f(x) = \left(a_1 x_1, a_2 x_2, \ldots, a_n x_n, \ldots\right).$$

(a) Prove that the image of f is contained in ℓ^2 and that $f\colon \ell^2 \longrightarrow \ell^2$ is linear and continuous.

(b) Find the norm $\|f\|$.

(c) Show that if $a_n \in \mathbf{R}$ for all $n \in \mathbf{N}$ then f is self-adjoint.

Solution. (a) We have

$$\|f(x)\|_{\ell^2}^2 = \sum_{n=1}^{\infty} |a_n|^2 |x_n|^2 \leqslant \|a\|_{\ell^\infty}^2 \sum_{n=1}^{\infty} |x_n|^2 = \|a\|_{\ell^\infty}^2 \|x\|_{\ell^2}^2,$$

so if $x \in \ell^2$ then $f(x) \in \ell^2$.

It is straightforward that f is linear. It is clear that f is continuous from the inequality above.

(b) We have

$$\|f\| = \sup_{\|x\|=1} \|f(x)\| \leqslant \|a\|_{\ell^\infty}$$

from the previous part.

Let $\varepsilon > 0$. Let $n \in \mathbf{N}$ be such that $|a_n| > \|a\|_{\ell^\infty} - \varepsilon$. Then

$$\|f(e_n)\| = \|a_n e_n\| = |a_n| > \|a\|_{\ell^\infty} - \varepsilon,$$

therefore $\|f\| = \|a\|_{\ell^\infty}$.

(c) We have

$$\langle f(x), y\rangle = \sum_{n=1}^{\infty} a_n x_n \overline{y}_n = \sum_{n=1}^{\infty} x_n \overline{(a_n y_n)} = \langle x, f(y)\rangle,$$

where we used the fact that $a_n \in \mathbf{R}$ for all $n \in \mathbf{N}$. □

Exercise 3.34. (*) Every nonzero Hilbert space H has an orthonormal basis.

[*Hint*: Use Zorn's Lemma (Lemma 1.3) and mimic the proof of the existence of bases for arbitrary vector spaces (Theorem 1.2).]

Solution. Let X be the set of all orthonormal systems in H. This is a poset under inclusion (it is the restriction of the poset structure on the power set of H to subsets of H that are orthonormal systems). It is nonempty: if y is any nonzero element of H, let $u = \frac{1}{\|y\|} y$, then $\{u\} \in X$.

Let C be a nonempty chain in X, in other words $C = \{S_i : i \in I\}$ where each S_i is an orthonormal system, and for any $i, j \in I$ we have $S_i \subseteq S_j$ or $S_j \subseteq S_i$.

Let

$$S = \bigcup_{i \in I} S_i.$$

If $s, t \in S$, then there exist $i, j \in I$ such that $s \in S_i$ and $t \in S_j \subseteq S_i$ (without loss of generality). Since $s, t \in S_i$ and S_i is orthonormal, we get that

$$\langle s, t\rangle = \begin{cases} 0 & \text{if } s \neq t \\ 1 & \text{if } s = t. \end{cases}$$

So S is orthonormal, hence is an upper bound for the chain C.

By Zorn's Lemma, X has a maximal element B. Let $Y = \overline{\operatorname{Span}(B)}$. If $Y = H$ then B is an orthonormal basis for H and we are done.

So assume that $Y \neq H$. Since H is a Hilbert space and Y is a closed subspace we have

$$H = Y \oplus Y^\perp,$$

so that $Y^\perp \neq 0$. Let $z \in Y^\perp$ be a nonzero element and let $u = \frac{1}{\|z\|} z$. Then $B \cup \{u\}$ is an orthonormal system (since u is a unit vector and it is in $Y^\perp$, hence in $B^\perp$) that strictly contains B, contradicting the maximality of B. □

Exercise 3.35. (*) Let $p_1(x) = 0$ and

$$p_{n+1}(x) = p_n(x) - \frac{p_n(x)^2 - x^2}{2} = p_n(x) - \frac{\big(p_n(x) - |x|\big)\big(p_n(x) + |x|\big)}{2} \qquad \text{for } n \geqslant 1.$$

Prove that, for all $x \in [-1,1]$ and all $n \geqslant 1$:

(a) $0 \leqslant p_n(x) \leqslant |x|$;

(b) $p_n(x) \leqslant p_{n+1}(x)$;

(c) $|x| - p_{n+1}(x) \leqslant |x|\left(1 - \frac{|x|}{2}\right)^n$.

Solution.

(a) We proceed by induction on n. Clearly $0 \leqslant p_1(x) \leqslant |x|$ for all $x \in [-1,1]$ since $p_1(x) = 0$.

Fix $n \geqslant 1$ and suppose $0 \leqslant p_n(x) \leqslant |x|$. Then

$$\begin{aligned} -|x| &\leqslant p_n(x) - |x| \leqslant 0 \\ |x| &\leqslant p_n(x) + |x| \leqslant 2|x|, \end{aligned}$$

so that

$$0 \geqslant \frac{p_n(x)^2 - x^2}{2} \geqslant -|x|^2,$$

and finally

$$0 \leqslant p_n(x) - \frac{p_n(x)^2 - x^2}{2} \leqslant |x| - |x|^2.$$

We are done because the middle expression is precisely $p_{n+1}(x)$, and

$$|x| - |x|^2 = |x|\big(1 - |x|\big) \leqslant |x| \quad \text{for } x \in [-1,1].$$

(b) We have

$$2\big(p_{n+1}(x) - p_n(x)\big) = x^2 - p_n(x)^2 \geqslant 0$$

by part (a).

(c) Note that

$$|x| - p_{n+2}(x) = |x| - p_{n+1}(x) - \frac{\big(|x| - p_{n+1}(x)\big)\big(|x| + p_{n+1}(x)\big)}{2} \leqslant \big(|x| - p_{n+1}(x)\big)\left(1 - \frac{|x|}{2}\right),$$

at which point the claim follows by a simple induction argument. □

Exercise 3.36. **(*)** Fix $n \geqslant 1$ and consider the function $f\colon [0,1] \longrightarrow \mathbf{R}$ given by

$$f(t) = t\left(1 - \frac{t}{2}\right)^n.$$

Prove that

$$f(t) < \frac{2}{n+1} \qquad \text{for all } t \in [0,1].$$

Solution. We have

$$f'(t) = \left(1 - \frac{t}{2}\right)^{n-1}\left(1 - \frac{(n+1)t}{2}\right),$$

with a stationary point at $t_0 = 2/(n+1) \in [0,1]$ and another at $2 \notin [0,1]$. So f attains its maximum either at t_0 or at one of the boundary points 0 or 1. But

$$f(0) = 0, \qquad f(1) = \frac{1}{2^n}, \qquad f(t_0) = \frac{2}{n+1}\left(1 - \frac{1}{n+1}\right)^n < \frac{2}{n+1},$$

and certainly $1/2^n < 2/(n+1)$ for all $n \geqslant 1$.

We conclude that the maximum value of f on $[0,1]$ is less than $2/(n+1)$. □

Exercise 3.37. Show that Lemma 3.58 holds more generally for the intervals $[-a,a]$ for any $a > 0$.

Solution. Let (p_n) be a sequence in $x\mathbf{R}[x]$ such that $(p_n) \longrightarrow |x|$ uniformly on $[-1,1]$.

Define $q_n(x) = a\,p_n(x/a)$, then I claim that $(q_n) \longrightarrow |x|$ uniformly on $[-a,a]$.

Let $\varepsilon > 0$ and let $N \in \mathbf{N}$ be such that for all $n \geqslant N$ we have

$$\big|p_n(t) - |t|\big| < \frac{\varepsilon}{a} \qquad \text{for all } t \in [-1,1].$$

Then for all $n \geqslant N$

$$\big|q_n(x) - |x|\big| = \big|a\,p_n(x/a) - a\,|x/a|\big| = a\,\big|p_n(x/a) - |x/a|\big| < \varepsilon \qquad \text{for all } x \in [-a,a]. \quad \square$$

Exercise 3.38. For each $n \in \mathbf{N}$ define $f_n\colon [0,1] \longrightarrow \mathbf{R}$ by

$$f_n(x) = \frac{nx^2}{1+nx}.$$

Convince yourself that f_n is continuous.

Find the pointwise limit f of the sequence (f_n) and determine whether the sequence converges uniformly to f.

Solution. The function f_n is the quotient of two continuous functions, and the denominator $1 + nx$ is nonzero on $[0,1]$, so f_n is continuous on $[0,1]$.

The pointwise limit is given by

$$f_n(x) = \frac{nx^2}{1+nx} = \frac{x^2}{\frac{1}{n} + x} \longrightarrow \frac{x^2}{0+x} = x \qquad \text{as } n \longrightarrow \infty,$$

so $f(x) = x$ for all $x \in [0,1]$.

The uniform norm of $f_n - f$ is given by

$$\|f_n - f\| = \left\| -\frac{x}{1+nx} \right\| = \sup_{x\in[0,1]} \frac{x}{1+nx} = \frac{1}{1+n} \longrightarrow 0 \qquad \text{as } n \longrightarrow \infty,$$

so the convergence is uniform.

To justify the above statement about the supremum, let $g\colon [0,1] \longrightarrow \mathbf{R}$ be given by

$$g(x) = \frac{x}{1+nx},$$

then

$$g'(x) = \frac{1}{(1+nx)^2}.$$

This shows that g has no stationary points in $(0,1)$, so its extremal values must occur at the boundary points $x = 0$ and $x = 1$. We have $g(0) = 0$ and $g(1) = 1/(1+n)$, so the latter is the maximum value. □

Exercise 3.39. Let V, W be normed spaces, with V Banach, and let $f \in L(V, W)$. Suppose that there exists a constant $c > 0$ such that

$$\|f(v)\|_W \geqslant c\,\|v\|_V \qquad \text{for all } v \in V.$$

Then $\operatorname{im}(f)$ is a closed subspace of W.

Solution. Let $w \in W$ and let (v_n) be a sequence in V such that $\big(f(v_n)\big) \longrightarrow w$ in W. We need to prove that $w \in \operatorname{im}(f)$.

For all $n, m \in \mathbf{N}$ we have

$$\|f(v_n) - f(v_m)\|_W = \|f(v_n - v_m)\|_W \geqslant c\,\|v_n - v_m\|_V.$$

But the sequence $\big(f(v_n)\big)$ converges, hence is Cauchy in W. Therefore the above inequality says that the sequence (v_n) is Cauchy in V. As V is Banach, we have $(v_n) \longrightarrow v \in V$. Since f is continuous, we have $w = \lim f(v_n) = f(v)$ and $w \in \operatorname{im}(f)$. □

Exercise 3.40. Consider the Hilbert space ℓ^2 of square-summable complex sequences $(a_1, a_2, \dots)$.

Let $\big(\lambda_n\big)$ be a bounded sequence of complex numbers and define $T\colon \ell^2 \longrightarrow \ell^2$ by

$$T\big(a_1, a_2, \dots\big) = \big(\lambda_1 a_2, \lambda_2 a_4, \dots, \lambda_n a_{2n}, \dots\big).$$

(a) Show that T is a continuous linear operator.

(b) Compute the norm $\|T\|$.

(c) Find the adjoint operator T^*.

Solution. Let $\lambda = (\lambda_n)$. As it is bounded, $\|\lambda\|_{\ell^\infty} < \infty$.

(a) Linearity of T is straightforward. For the continuity let $a = (a_n) \in \ell^2$, then

$$\|T(a)\|_{\ell^2}^2 = \sum_{n=1}^{\infty} \left|\lambda_n a_{2n}\right|^2 \leqslant \|\lambda\|_{\ell^\infty}^2 \sum_{n=1}^{\infty} |a_{2n}|^2 \leqslant \|\lambda\|_{\ell^\infty}^2 \|a\|_{\ell^2}^2.$$

(b) From part (a) we see that

$$\|T\| \leqslant \|\lambda\|_{\ell^\infty}.$$

We claim that this is actually an equality. Note that $T(e_{2n-1}) = 0$ and $T(e_{2n}) = \lambda_n e_n$, where $\{e_n : n \in \mathbf{N}\}$ is the Schauder basis of ℓ^2.

Let $\varepsilon > 0$. Since $\|\lambda\|_{\ell^\infty} = \sup_n |\lambda_n|$, there exists $n \in \mathbf{N}$ such that $|\lambda_n| > \|\lambda\|_{\ell^\infty} - \varepsilon$. Then

$$\frac{\|T(e_{2n})\|}{\|e_{2n}\|} = |\lambda_n| > \|\lambda\|_{\ell^\infty} - \varepsilon,$$

so we conclude that

$$\|\lambda\|_{\ell^\infty} = \sup_{a\neq 0} \frac{\|T(a)\|}{\|a\|} = \|T\|.$$

(c) For $a = (a_n), b = (b_n) \in \ell^2$ we have

$$\left\langle T(a), b\right\rangle = \sum_{n=1}^{\infty} \lambda_n a_{2n} \bar{b}_n = \sum_{n=1}^{\infty} a_{2n} \overline{\left(\overline{\lambda}_n\, b_n\right)} = \left\langle a, T^*(b)\right\rangle$$

where

$$T^*(b) = \left(0, \overline{\lambda}_1 b_1, 0, \overline{\lambda}_2 b_2, 0, \dots\right). \qquad \square$$

Exercise 3.41. Let $X = C_0\big([0,1], \mathbf{R}\big)$ be the Banach space of continuous functions $f : [0,1] \longrightarrow \mathbf{R}$ with the supremum norm.

Define $\phi : X \longrightarrow \mathbf{R}$ by $\phi(f) = f(0)$ for all $f \in X$.

Prove that ϕ is a continuous linear map.

Solution. It is clear that ϕ is linear:

$$\phi(f+g) = (f+g)(0) = f(0) + g(0) = \phi(f) + \phi(g)$$

and

$$\phi(\alpha f) = (\alpha f)(0) = \alpha f(0) = \alpha\phi(f).$$

It is also clearly continuous:

$$|\phi(f)| = |f(0)| \leqslant \|f\|,$$

as $\|f\|$ is the supremum of $|f(x)|$ for $x \in [0,1]$. $\square$

Exercise 3.42. We explore the Hilbert Projection Theorem when V is a Banach space but not a Hilbert space.

(a) Let $V = \mathbf{R}^2$ with the ℓ^1-norm, that is

$$\big\|(x_1, x_2)\big\| = |x_1| + |x_2|.$$

Let $Y = \mathbf{B}_1(0)$, the closed unit ball around 0. Find two distinct closest points in Y to $x = (-1,1) \in V$.

(b) Can you find a similar example for $V = \mathbf{R}^2$ with the ℓ^∞-norm:

$$\|(x_1, x_2)\| = \max\{|x_1|, |x_2|\}?$$

(c) Let V be a normed space and Y a convex subset of V. Fix $x \in V$. Let $Z \subseteq Y$ be the set of all closest points in Y to x. Prove that Z is convex.

Solution.

(a) Let $y = (y_1, y_2) \in Y$, then $d(y,0) \leqslant 1$.

Note that $d(x,0) = 2$. By the triangle inequality

$$d(x,y) + d(y,0) \geqslant d(x,0) \Rightarrow d(x,y) \geqslant d(x,0) - d(y,0) \geqslant 2 - 1 = 1.$$

Since this holds for all $y \in Y$, we have $d_Y(x) \geqslant 1$.

But there are (uncountably many) points of Y at distance 1 from x: take any point $y = (y_1, y_2)$ on the line segment joining $(-1,0)$ to $(0,1)$, then $y_2 = y_1 + 1$ with $-1 \leqslant y_1 \leqslant 0$ and

$$d(x,y) = |-1-y_1| + |y_1| = 1 + y_1 - y_1 = 1.$$

We conclude that $d_Y(x) = 1$ and all the points on that line segment are closest points to x.

(b) We can recreate a similar scenario for the ℓ^∞-norm on $V = \mathbf{R}^2$ by taking $Y = \mathbf{B}_1(0)$ and $x = (2,0)$, for instance.

The same argument as in (a) gives us $d_Y(x) = 1$ and every point on the line segment joining $(1,-1)$ to $(1,1)$ is at this distance from x.

(c) (Let's note that the conclusion definitely holds for parts (a) and (b), as well as in the Hilbert case covered by the Projection Theorem.)

Let $D = d_Y(x)$.

If Z is empty it is certainly convex.

Otherwise let $z_1, z_2 \in Z$ and let $a \in [0,1]$. Consider $y = az_1 + (1-a)z_2$. Since $z_1, z_2 \in Z \subseteq Y$ and Y is convex, we have that $y \in Y$. We have

$$\begin{aligned} d(y,x) = \|y - x\| &= \|az_1 + (1-a)z_2 - x\| = \|az_1 - ax + (1-a)z_2 - (1-a)x\| \\ &= \|a(z_1 - x) + (1-a)(z_2 - x)\| \leqslant \|a(z_1 - x)\| + \|(1-a)(z_2 - x)\| \\ &= a\|z_1 - x\| + (1-a)\|z_2 - x\| = aD + (1-a)D = D. \end{aligned}$$

So $d(y,x) \leqslant D$, but also $d(y,x) \geqslant D = d_Y(x)$, so we must have $d(y,x) = D$ and $y \in Z$. □

Exercise 3.43. Let $H = \ell^2$ over $\mathbf{R}$ and consider the subset

$$W = \{y = (y_n) \in \ell^2 : y_n \geqslant 0 \text{ for all } n \in \mathbf{N}\}.$$

(a) Prove that W is a closed, convex subset of H. Is it a vector subspace?

(b) Find the closest point $y_{\min} \in W$ to

$$x = (x_n) = \left(\frac{(-1)^n}{n}\right) = \left(-1, \frac{1}{2}, -\frac{1}{3}, \dots\right)$$

and compute $d_W(x)$.

[*Hint*: You may use without proof the identity $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.]

Solution.

(a) If $y, z \in W$ and $a \in [0,1]$ then $ay + (1-a)z = \big(ay_n + (1-a)z_n\big)$ and it is clear that $ay_n + (1-a)z_n \geqslant 0$, so W is convex.

To show that W is closed we note that

$$W = \bigcap_{n \in \mathbf{N}} \pi_n^{-1}\big([0,\infty)\big),$$

where $\pi_n : \ell^2 \longrightarrow \mathbf{R}$ is given by $\pi_n\big((a_n)\big) = a_n$. We've seen in Tutorial Question 10.6 that π_n is continuous, so since $[0,\infty)$ is closed in $\mathbf{R}$, W is the intersection of a collection of closed subsets, hence it is closed.

Not a vector subspace because not closed under multiplication by $-1 \in \mathbf{R}$.

(b) Let $y = (y_n) \in W$, then

$$\begin{aligned} \|x - y\|^2 &= \sum_{n=1}^{\infty} \left| \frac{(-1)^n}{n} - y_n \right|^2 \\ &= \sum_{n \text{ odd}} \left| -\frac{1}{n} - y_n \right|^2 + \sum_{n \text{ even}} \left| \frac{1}{n} - y_n \right|^2 \\ &= \sum_{n \text{ odd}} \left| \frac{1}{n} + y_n \right|^2 + \sum_{n \text{ even}} \left| \frac{1}{n} - y_n \right|^2 \end{aligned}$$

Note that since $y_n \geqslant 0$:

$$\begin{aligned} &\text{if } n \text{ is odd then} && \left| \frac{1}{n} + y_n \right|^2 \geqslant \frac{1}{n^2} \\ &\text{if } n \text{ is even then} && \left| \frac{1}{n} - y_n \right|^2 \geqslant 0. \end{aligned}$$

Putting this together with the previous result, we get

$$d(x,y)^2 = \|x - y\|^2 \geqslant \sum_{n \text{ odd}} \frac{1}{n^2}.$$

As this holds for all $y \in W$, we get that

$$d_W(x) \geqslant \sqrt{\sum_{n \text{ odd}} \frac{1}{n^2}}.$$

But following the calculations above it is easy to put together an element $y_{\min} = (y_n) \in W$ that achieves this lower bound:

$$y_n = \begin{cases} \frac{1}{n} & \text{if } n \text{ is even} \\ 0 & \text{if } n \text{ is odd.} \end{cases}$$

Finally, to compute $d_W(x)$, note

$$\sum_{n \text{ odd}} \frac{1}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n^2} - \sum_{n \text{ even}} \frac{1}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n^2} - \sum_{k=1}^{\infty} \frac{1}{(2k)^2} = \frac{3}{4} \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{8},$$

hence

$$d_W(x) = \frac{\pi}{2\sqrt{2}}.$$

□

Exercise 3.44 (Pythagorean theorem)**.** Let v and w be two orthogonal vectors in an inner product space V. Prove that

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

Solution. Straightforward computation:

$$\|v + w\|^2 = \langle v + w, v + w \rangle = \langle v, v \rangle + \langle v, w \rangle + \langle w, v \rangle + \langle w, w \rangle = \|v\|^2 + \|w\|^2.$$

□

Exercise 3.45. (*) For $n \in \mathbf{N}$, consider the function $f_n \colon [0, 2] \longrightarrow \mathbf{R}$ defined by

$$f_n(x) = \begin{cases} n^2 x & \text{if } 0 \leqslant x \leqslant \frac{1}{n} \\ -n^2 \left(x - \frac{2}{n}\right) & \text{if } \frac{1}{n} < x \leqslant \frac{2}{n} \\ 0 & \text{if } \frac{2}{n} < x \leqslant 2. \end{cases}$$

(You might want to graph f_1, f_2, f_3 to get a feel for what the functions look like.)

Find the pointwise limit $f(x)$ of $\big(f_n(x)\big)$ for all $x \in [0, 2]$.

Show that (f_n) does not converge to f with respect to the L^1 norm.

Solution. The pointwise limit is the constant function zero.

We have $\|f_n\|_{L^1} = 1$ for all $n \in \mathbf{N}$, so (f_n) does not converge to f with respect to the L^1 norm.

□

Exercise 3.46. (*) Given a subset $S \subseteq [0, 1]$, let $\mathbf{1}_S \colon [0, 1] \longrightarrow \mathbf{R}$ denote the *characteristic function* of S, that is

$$\mathbf{1}_S(x) = \begin{cases} 1 & \text{if } x \in S \\ 0 & \text{if } x \notin S. \end{cases}$$

Consider the sequence of functions (g_n) defined as follows: write $n \in \mathbf{N}$ in the form

$$n = 2^k + \ell, \qquad k, \ell \in \mathbf{Z}_{\geqslant 0}, 0 \leqslant \ell < 2^k,$$

then define $g_n\colon [0,1] \longrightarrow \mathbf{R}$ by

$$g_n = \mathbf{1}_{[\ell/2^k,(\ell+1)/2^k]}.$$

(You might want to graph $g_1,\dots,g_5$ to get a feel for what the functions look like.)

Show that (g_n) converges to the constant function zero with respect to the L^1 norm, but that $\big(g_n(x)\big)$ does not converge for any $x \in [0,1]$.

Solution. Note that $2^k \leqslant n < 2^{k+1}$, so that we have

$$\|g_n\|_{L^1} = \frac{1}{2^k} < \frac{2}{n} \longrightarrow 0 \qquad \text{as } n \longrightarrow \infty,$$

so $(g_n) \longrightarrow 0$ with respect to the L^1 norm.

However, for any $x \in [0,1]$ there are infinitely many values of n for which $g_n(x) = 0$ and infinitely many values of n for which $g_n(x) = 1$, which means that $\big(g_n(x)\big)$ does not converge. □

Exercise 3.47. (*) Let $f \in L(H)$ for a complex Hilbert space H. Prove that f is of finite rank if and only if there exist a finite orthonormal system $\{u_n\colon 1 \leqslant n \leqslant m\}$ and a complex matrix $C = (c_{ij}) \in M_n(\mathbf{C})$ such that

$$f(x) = \sum_{i,j=1}^{m} c_{ij}\langle x, u_j\rangle u_i \qquad \text{for all } x \in H.$$

Solution. The reverse implication is clear: if a finite orthonormal system with the given property exists, then $\operatorname{im}(f) \subseteq \operatorname{Span}\{u_1,\dots,u_m\}$ is finite-dimensional.

Conversely, let $\{u_1,\dots,u_k\}$ be an orthonormal basis of the (finite-dimensional) image W of f. For any $x \in H$, $f(x) \in W$ so we have

$$f(x) = \sum_{i=1}^{k}\langle f(x), u_i\rangle u_i = \sum_{i=1}^{k}\langle x, f^*(u_i)\rangle u_i.$$

Now apply Gram–Schmidt to the set $\{u_1,\dots,u_k, f^*(u_1),\dots,f^*(u_k)\}$ and obtain a finite orthonormal system $\{u_1,\dots,u_m\}$ for some $m \geqslant k$. In particular, for any $i = 1,\dots,k$ we have

$$f^*(u_i) = \sum_{j=1}^{m} d_{ij}u_j,$$

so that

$$f(x) = \sum_{i,j=1}^{m} c_{ij}\langle x, u_j\rangle u_i,$$

where

$$c_{ij} = \begin{cases} \overline{d}_{ij} & \text{if } i \leqslant k \\ 0 & \text{if } i > k. \end{cases}$$

□

Exercise 3.48. (*) Let $R(H)$ denote the set of all maps $f \in B(H)$ of finite rank on a complex Hilbert space H.

Prove that $R(H)$ is a vector subspace of $B(H)$.

Solution. The constant zero map is certainly of finite rank.

If $f, g \in R(H)$ then $\text{im}(f)$ and $\text{im}(g)$ are finite-dimensional subspaces of H. Therefore $\text{im}(f) + \text{im}(g)$ is a finite-dimensional subspace of H, and certainly $\text{im}(f+g) \subseteq \text{im}(f) + \text{im}(g)$.

If $f \in R(H)$ and $\alpha \in \mathbf{C}$ then $\text{im}(\alpha f) \subseteq \text{im}(f)$ is finite-dimensional. ☐

Exercise 3.49. (*) Prove that if $f \in R(H)$ and $g_1, g_2 \in B(H)$ then $g_2 \circ f \circ g_1 \in R(H)$.

Solution. Clearly $\text{im}(f \circ g_1) \subseteq \text{im}(f)$ is finite-dimensional.

On the other hand, $g_2 \circ f$ has a finite-dimensional domain, hence a finite-dimensional image. ☐

Exercise 3.50. (*) Prove that if $f \in R(H)$ then $f^* \in R(H)$.

[*Hint*: Use Exercise 3.47.]

Solution. By Exercise 3.47 we have, for all $x, y \in H$:

$$\begin{aligned}\langle f(x), y\rangle &= \left\langle \sum_{i,j=1}^{m} c_{ij}\langle x, u_j\rangle u_i, y\right\rangle \\ &= \sum_{i,j=1}^{m} c_{ij}\langle x, u_j\rangle\langle u_i, y\rangle \\ &= \sum_{i,j=1}^{m} \langle x, \overline{c}_{ij}\langle y, u_i\rangle u_j\rangle \\ &= \left\langle x, \sum_{i,j=1}^{m} \overline{c}_{ij}\langle y, u_i\rangle u_j\right\rangle,\end{aligned}$$

from which we conclude that

$$f^*(y) = \sum_{i,j=1}^{m} \overline{c}_{ij}\langle y, u_i\rangle u_j \qquad \text{for all } y \in H,$$

so f^* has finite rank. ☐

Exercise 3.51. (*) Recall the right shift operator $R\colon \ell^2 \longrightarrow \ell^2$

$$R(a_1, a_2, \dots) = (0, a_1, a_2, \dots).$$

(a) Prove that R has no complex eigenvalues.

(b) Prove that $0 \in \sigma(R)$.

(c) Is R a compact map?

Solution.

(a) Suppose

$$(0, a_1, a_2, \dots) = R(a_1, a_2, \dots) = \lambda(a_1, a_2, \dots),$$

then $\lambda a_1 = 0$, so either $\lambda = 0$ implying that $a_1 = a_2 = \cdots = 0$; or $a_1 = 0$ which implies that $a_2 = 0$, and so on. In both cases the alleged eigenvector is actually the zero vector.

(b) It is clear that R is not surjective, hence not invertible, so $0 \in \sigma(R)$.

(c) No, for the same reason that id_{ℓ^2} is not compact: $R(\mathbf{D}_1(0))$ contains $\{e_2, e_3, e_4, \dots\}$, hence a sequence that has no convergent subsequences. □

Exercise 3.52. **(*)** Let H be a complex Hilbert space and let

$$\mathrm{GL}(H) = \{f \in L(H) : f \text{ is invertible}\}.$$

For $f \in \mathrm{GL}(H)$, prove that

$$\mathbf{B}_r(f) \subseteq \mathrm{GL}(H) \qquad \text{where } r = \left\| f^{-1} \right\|^{-1}.$$

[*Hint*: Given $g \in \mathbf{B}_r(f)$, consider $i := -f^{-1} \circ (g - f)$ and use Proposition 3.72 to show that $\mathrm{id}_H - i$ is invertible.]

Conclude that $\mathrm{GL}(H)$ is an open subset of $L(H)$.

Solution. Take $g \in \mathbf{B}_r(f)$, then $\|g - f\| < r$. Let $i = -f^{-1} \circ (g - f)$, then

$$\|i\| = \|f^{-1} \circ (g - f)\| \leqslant \|f^{-1}\| \, \|g - f\| < \|f^{-1}\| r = 1,$$

so by Proposition 3.72 we get that $\mathrm{id}_H - i$ is invertible. But then

$$f \circ \big(\mathrm{id}_H - i\big) = f \circ \big(\mathrm{id}_H + f^{-1} \circ (g - f)\big) = f + g - f = g,$$

so g is the composition of two invertible maps, hence is itself invertible. □

Exercise 3.53. **(*)** Prove that the spectrum of any $f \in L(H)$ is a compact set.

[*Hint*: Use Exercise 3.52 to show that the resolvent $\rho(f)$ is an open subset of $\mathbf{C}$, then use Corollary 3.73.]

Solution. Consider the map $F_f : \mathbf{C} \longrightarrow L(H)$ given by

$$F_f(\lambda) = f - \lambda \, \mathrm{id}_H .$$

This is a continuous function (check this!), and $\rho(f) = F_f^{-1}\big(\mathrm{GL}(H)\big)$ is an open subset of $\mathbf{C}$, hence $\sigma(f)$ is a closed subset of $\mathbf{C}$. But by Corollary 3.73 $\sigma(f)$ is a subset of the compact disc (sic) $\mathbf{D}_{\|f\|}(0)$, so it is compact. □

Exercise 3.54. **(*)** Let $f \in L(H)$ be a self-adjoint map on a complex Hilbert space H and let $a + ib \in \mathbf{C}$. Prove that

$$\big\| \big(f - (a + ib)\, \mathrm{id}_H\big)(x) \big\| \geqslant |b| \, \|x\| \qquad \text{for all } x \in H.$$

[*Hint*: Expand $\|(f - (a + ib)\, \mathrm{id}_H)(x)\|^2$ using the inner product, take advantage of $f^* = f$, and manipulate until you get a sum of two squares, one of which is $b^2\|x\|^2$.]

Solution. We follow the hint:

$$\begin{aligned}\big\|\big(f-(a+ib)\operatorname{id}_H\big)\big\|^2 &= \big\langle\big(f-(a+ib)\operatorname{id}_H\big)(x),\big(f-(a+ib)\operatorname{id}_H\big)(x)\big\rangle\\ &= \big\langle\big(f-(a+ib)\operatorname{id}_H\big)(x),\big(f-(a-ib)\operatorname{id}_H\big)^*(x)\big\rangle\\ &= \big\langle\big(f-(a-ib)\operatorname{id}_H\big)\big(f-(a+ib)\operatorname{id}_H\big)(x),x\big\rangle\\ &= \big\langle\big((f-a\operatorname{id}_H)^2+b^2\operatorname{id}_H\big)(x),x\big\rangle\\ &= \big\langle\big(f-a\operatorname{id}_H\big)^2(x),x\big\rangle+b^2\|x\|^2\\ &= \big\|(f-a\operatorname{id}_H)^2(x)\big\|^2+b^2\|x\|^2\\ &\geqslant b^2\|x\|^2.\end{aligned}$$

□

Exercise 3.55. (*) Let X be a metric space that has at least two elements and let $\mathcal{A}$ be a subalgebra of $C_0(X,\mathbf{R})$.

We say that $\mathcal{A}$ *separates points of* X if for any $x,x'\in X$ with $x\neq x'$ there exists $f\in\mathcal{A}$ such that $f(x)\neq f(x')$.

We say that $\mathcal{A}$ *is non-vanishing on* X if for any $x\in X$ there exists $f\in\mathcal{A}$ such that $f(x)\neq 0$.

Prove that $\mathcal{A}$ interpolates pairs of points on X if and only if $\mathcal{A}$ separates points of X and is non-vanishing on X.

[*Hint*: For the "if" direction, given $(x,y),(x',y')\in X\times\mathbf{R}$ with $x\neq x'$, find elements $k,k'\in\mathcal{A}$ such that $k(x)=0$, $k(x')\neq 0$, $k'(x)\neq 0$, $k'(x')=0$.]

Solution. Suppose $\mathcal{A}$ interpolates pairs of points on X. Let $x\neq x'$ with $x,x'\in X$ and consider the pair of points $(x,0),(x',1)\in X\times\mathbf{R}$. Then there exists $f\in\mathcal{A}$ such that

$$f(x)=0\neq 1=f(x'),$$

therefore $\mathcal{A}$ separates points of X.

Now let $x\in X$. Choose $x'\in X$ such that $x'\neq x$ and consider the pair of points $(x,1),(x',0)\in X\times\mathbf{R}$, then there exists $f\in\mathcal{A}$ such that $f(x)=1\neq 0$. Therefore $\mathcal{A}$ is non-vanishing on X.

Conversely, suppose $\mathcal{A}$ separates points of X and is non-vanishing on X and let $(x,y),(x',y')\in X\times\mathbf{R}$ with $x\neq x'$. Then there exist elements $g,h,h'\in\mathcal{A}$ such that

$$g(x)\neq g(x'),\qquad h(x)\neq 0,\qquad h'(x')\neq 0.$$

Define $k,k'\in\mathcal{A}$ by

$$\begin{aligned}k(t)&=\big(g(t)-g(x)\big)\,h'(t)\\ k'(t)&=\big(g(t)-g(x')\big)\,h(t),\end{aligned}$$

then

$$k(x)=0,\qquad k(x')\neq 0,\qquad k'(x)\neq 0,\qquad k'(x')=0.$$

Finally let

$$f(t)=\frac{y}{k'(x)}\,k'(t)+\frac{y'}{k(x')}\,k(t).$$

□

A. Appendix

Exercise A.1. Let V be a vector space over $\mathbf{F}$. Prove that $\mathrm{End}(V) := \mathrm{Hom}(V, V)$ is an associative unital $\mathbf{F}$-algebra under composition of functions.

Solution. TODO □

Exercise A.2. Let V, W be vector spaces over $\mathbf{F}$ and let B be a basis of V. Suppose $g\colon B \longrightarrow W$ is a function, and let $f\colon V \longrightarrow W$ be its extension to V by linearity.

Prove that

(a) f is injective if and only if $g(B)$ is linearly independent in W;

(b) f is surjective if and only if $g(B)$ spans W;

(c) f is bijective if and only if $g(B)$ is a basis for W.

Solution. TODO □

Exercise A.3. If S and T are subspaces of a vector space V with field of scalars $\mathbf{F}$, then so are $S + T$ and αS for any $\alpha \in \mathbf{F}$.

Solution. TODO □

Exercise A.4. Let $V = \mathbf{F}[x]$ be the vector space of polynomials in one variable with coefficients in $\mathbf{F}$. Given a scalar $\alpha \in \mathbf{F}$, consider the function $\mathrm{ev}_\alpha\colon V \longrightarrow \mathbf{F}$ given by evaluation at α:

$$\mathrm{ev}_\alpha(f) = f(\alpha).$$

Prove that $\mathrm{ev}_\alpha \in V^\vee$.

Solution. We have to prove that $\mathrm{ev}_\alpha\colon V \longrightarrow \mathbf{F}$ is linear.

If $f_1, f_2 \in \mathbf{F}[x]$, then

$$\mathrm{ev}_\alpha(f_1 + f_2) = (f_1 + f_2)(\alpha) = f_1(\alpha) + f_2(\alpha) = \mathrm{ev}_\alpha(f_1) + \mathrm{ev}_\alpha(f_2).$$

If $f \in \mathbf{F}[x]$ and $\lambda \in \mathbf{F}$, then

$$\mathrm{ev}_\alpha(\lambda f) = (\lambda f)(\alpha) = \lambda f(\alpha) = \lambda\, \mathrm{ev}_\alpha(f).$$ □

Exercise A.5. In the setup of Proposition A.4, suppose $W = V$ so that $T\colon V \longrightarrow V$ and $T^\vee\colon V^\vee \longrightarrow V^\vee$.

Let M be the matrix representation of T with respect to an ordered basis B of V, and let $M^\vee$ be the matrix representation of $T^\vee$ with respect to the dual basis $B^\vee$.

Express $M^\vee$ in terms of M.

Solution. As in Proposition A.2, we have $B = (v_1, \ldots, v_n)$ and $B^\vee = (v_1^\vee, \ldots, v_n^\vee)$. Write (a_{ij}) for the entries of the matrix M. For future reference, the i-th row of M is

$$\begin{bmatrix} a_{i1} & a_{i2} & \ldots & a_{in} \end{bmatrix}.$$

By the definition of matrix representations, we have

$$\begin{aligned} T(v_1) &= a_{11}v_1 + a_{21}v_2 + \cdots + a_{n1}v_n \\ T(v_2) &= a_{12}v_1 + a_{22}v_2 + \cdots + a_{n2}v_n \\ &\vdots \\ T(v_n) &= a_{1n}v_1 + a_{2n}v_2 + \cdots + a_{nn}v_n. \end{aligned}$$

The i-th column of $M^\vee$ is given by the $B^\vee$-coordinates of the vector $T^\vee(v_i^\vee) = v_i^\vee \circ T$. To determine these, we apply $v_i^\vee \circ T$ to the basis vectors $v_1, \ldots, v_n$:

$$T^\vee(v_i^\vee)(v_j) = (v_i^\vee \circ T)(v_j) = v_i^\vee(T(v_j)) = v_i^\vee(a_{1j}v_1 + a_{2j}v_2 + \cdots + a_{nj}v_n) = a_{ij}.$$

This means that

$$T^\vee(v_i^\vee) = a_{i1}v_1^\vee + a_{i2}v_2^\vee + \cdots + a_{in}v_n^\vee$$

and the i-th column of $M^\vee$ is

$$\begin{bmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{bmatrix},$$

precisely the i-th row of M.

We conclude that $M^\vee = M^T$, the transpose of the matrix M. □

Exercise A.6. Let $v_1, \ldots, v_n \in V$. Define $\Gamma\colon V^\vee \longrightarrow \mathbf{F}^n$ by

$$\Gamma(\varphi) = \begin{bmatrix} \varphi(v_1) \\ \vdots \\ \varphi(v_n) \end{bmatrix}.$$

(a) Prove that Γ is a linear transformation.

(b) Prove that Γ is injective if and only if $\{v_1, \ldots, v_n\}$ spans V.

(c) Prove that Γ is surjective if and only if $\{v_1, \ldots, v_n\}$ is linearly independent.

Solution.

(a) Given $\varphi_1, \varphi_2 \in V^\vee$, we have

$$\begin{aligned} \Gamma(\varphi_1 + \varphi_2) &= ((\varphi_1 + \varphi_2)(v_1), \ldots, (\varphi_1 + \varphi_2)(v_n)) \\ &= (\varphi_1(v_1), \ldots, \varphi_1(v_n)) + (\varphi_2(v_1), \ldots, \varphi_2(v_n)) \\ &= \Gamma(\varphi_1) + \Gamma(\varphi_2). \end{aligned}$$

Given $\varphi \in V^\vee$ and $\lambda \in \mathbf{F}$, we have

$$\begin{aligned}\Gamma(\lambda\varphi) &= ((\lambda\varphi)(v_1), \ldots, (\lambda\varphi)(v_n)) \\ &= (\lambda\varphi(v_1), \ldots, \lambda\varphi(v_n)) \\ &= \lambda\Gamma(\varphi).\end{aligned}$$

(b) Suppose Γ is injective. Let $W = \text{Span}\{v_1, \ldots, v_n\}$. We want to prove that $W = V$.

Suppose $W \neq V$. Let $C = \{w_1, \ldots, w_k\}$ be a basis of W and extend it to a basis $B = \{w_1, \ldots, w_k, w_{k+1}, \ldots, w_m\}$ of V.

Let $B^\vee$ be the dual basis to B and consider its last element $v_m^\vee$ given by

$$v_m^\vee(a_1 w_1 + \cdots + a_m w_m) = a_m.$$

Then $v_m^\vee \neq 0$ (since $v_m^\vee(w_m) = 1$, for instance) but $v_m^\vee(w) = 0$ for all $w \in W$. In particular, $v_m^\vee(v_1) = \cdots = v_m^\vee(v_n) = 0$, so $\Gamma(v_m^\vee) = 0$, contradicting the injectivity of Γ.

We conclude that $W = V$, in other words $\{v_1, \ldots, v_n\}$ spans V.

Conversely, suppose $\{v_1, \ldots, v_n\}$ spans V. If $\varphi_1, \varphi_2 \in V^\vee$ are such that $\Gamma(\varphi_1) = \Gamma(\varphi_2)$, then $\Gamma(\varphi_1 - \varphi_2) = 0$, so setting $\varphi = \varphi_1 - \varphi_2$, we want to show that $\varphi = 0$, the constant zero function.

If $\varphi \neq 0$, then there exists $v \in V - \{0\}$ such that $\varphi(v) \neq 0$. Since $\{v_1, \ldots, v_n\}$ spans V, then we can write v as

$$v = b_1 v_1 + \cdots + b_n v_n.$$

But $\Gamma(\varphi) = 0$, so

$$0 \neq \varphi(v) = b_1\varphi(v_1) + \cdots + b_n\varphi(v_n) = 0,$$

which is a contradiction. So we must have $\varphi = 0$, that is $\varphi_1 = \varphi_2$. We conclude that Γ is injective.

(c) Suppose $\Gamma\colon V^\vee \longrightarrow \mathbf{F}^n$ is surjective. Let

$$a_1 v_1 + \cdots + a_n v_n = 0$$

be a linear relation.

Let $i \in \{1, \ldots, n\}$. Since Γ is surjective, given the standard basis vector $e_i \in \mathbf{F}^n$ (1 in the i-th entry), there exists $\varphi_i \in V^\vee$ such that $\Gamma(\varphi_i) = e_i$. If we apply φ_i on both sides of the linear relation, we get

$$a_i = 0.$$

Since this holds for all i, the relation is trivial.

Conversely, suppose $\{v_1, \ldots, v_n\}$ is linearly independent. This set can be enlarged to a basis $B = \{v_1, \ldots, v_n, v_{n+1}, \ldots, v_m\}$ of V, with dual basis $v_1^\vee, \ldots, v_m^\vee$.

Now take an arbitrary vector in $\mathbf{F}^n$:

$$w = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

Let

$$\varphi = a_1 v_1^\vee + \cdots + a_n v_n^\vee,$$

then

$$\Gamma(\varphi) = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = w.$$

We conclude that Γ is surjective. $\square$

Bibliography

[1] Alexander Nagel. Convexity and the sign of the second derivative. https://people.math.wisc.edu/~ajnagel/convexity.pdf, University of Wisconsin lecture notes, 2010. [Online; accessed 25 August 2023].

Index