# Solutions to exercises on metric and Hilbert spaces
# An invitation to functional analysis

Alexandru Ghitza*
School of Mathematics and Statistics
University of Melbourne

Version of Sat $11^{\text{th}}$ Oct, 2025 at 15:47

*(aghitza@alum.mit.edu)

# Contents

# 1. Metric and topological spaces

## Metrics

**Solution 1.1.** 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

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

**Solution 1.2.** 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 1.1.

**Solution 1.3.** 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 $x = y$.

**Solution 1.4.** 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 $(-1,1)$.

**Solution 1.5.** 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 $d(x,y) \leqslant d(x,t)+d(t,y)$.

**Solution 1.6.** 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 1.5), 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 2.3, therefore also a metric.

**Solution 1.7.** 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 $c=d(x,y)$, $a=d(x,t)$, $b=d(t,y)$.

**Solution 1.8.** Let $u \in U$, then $u \neq x$ so $r := d(u,x) > 0$. Then $x \notin \mathbf{B}_r(u)$, so $\mathbf{B}_r(u) \subseteq U$.

**Solution 1.9.** This is a variation on Example 2.8 and a generalisation of Exercise 1.8 (which is the case $r=0$).

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

I claim that $\mathbf{B}_t(y) \subseteq \big(X \setminus C\big)$: 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 $w \notin C$.

**Solution 1.10.** (a) Using the fundamental theorem of arithmetic (the existence of a unique prime factorisation of any natural number $\geqslant 2$), we have $m = p^{v_p(m)}m'$ and $n = p^{v_p(n)}n'$ with $p \nmid m'$ and $p \nmid n'$. Then

$$mn = p^{v_p(m)+v_p(n)}m'n' \qquad \text{and } p \nmid m'n',$$

so that $v_p(m)+v_p(n)$ is indeed the same as $v_p(mn)$.

(b) Write $x = \frac{m}{n}$, $y = \frac{a}{b}$, then

$$v_p(xy) = v_p\left(\frac{ma}{nb}\right) = v_p(ma) - v_p(nb) = v_p(m) + v_p(a) - v_p(n) - v_p(b) = v_p(x) + v_p(y).$$

For $v_p(x+y)$, without loss of generality assume $v := v_p(x) \leqslant v_p(y) =: u$ and write $x = p^v \frac{m'}{n'}$, $y = p^u \frac{a'}{b'}$. Then

$$x+y = p^v\,\frac{m'}{n'} + p^u\,\frac{a'}{b'} = p^v\left(\frac{m'}{n'} + p^{u-v}\,\frac{a'}{b'}\right) = p^v\left(\frac{m'b' + p^{u-v}a'n'}{n'b'}\right),$$

so that (since $p$ does not divide $n'b'$)

$$v_p(x+y) = v + v_p\big(m'b' + p^{u-v}a'n'\big).$$

Since $v_p$ of the quantity in parentheses is non-negative, we conclude that $v_p(x+y) \geqslant v = \min\big\{v_p(x), v_p(y)\big\}$.

Moreover, if $v < u$ then the quantity in parentheses has valuation zero, so that $v_p(x+y) = v = \min\big\{v_p(x), v_p(y)\big\}$.

(c) Direct from the previous part and $|x|_p = p^{-v_p(x)}$.

(d) We have

i. Clearly $v_p(y-x) = v_p(-1) + v_p(x-y) = v_p(x-y)$, so $d_p(y,x) = d_p(x,y)$.

ii. Letting $u = x - t$ and $v = t - y$, we want to prove that $|u+v|_p \leqslant |u|_p + |v|_p$. But we have already seen that

$$|u+v|_p \leqslant \max\big\{|x|_p, |y|_p\big\},$$

and the latter is clearly $\leqslant |x|_p + |y|_p$.

iii. If $x \in \mathbf{Q} \neq 0$, then $v_p(x) \in \mathbf{Z}$ so $|x|_p = p^{-v_p(x)} \in \mathbf{Q} \setminus \{0\}$. Hence $|x|_p = 0$ iff $x = 0$, which implies that $d_p(x,y) = 0$ iff $x = y$.

**Solution 1.11.** (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 Exercise 1.10, 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

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

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

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

**Solution 1.12.** Any open ball in any metric space is an open set (Example 2.8). Let's show that an arbitrary $p$-adic open ball $\mathbf{B}_r(c)$ is closed.

Let $U = \mathbf{Q} \setminus \mathbf{B}_r(c)$. 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 $|u-c|_p \geqslant r$.

## Topological spaces and continuous functions

**Solution 1.13.** 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

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

For each $i = 1, \dots, n$, $C_i$ is closed so $X \smallsetminus C_i$ is open, therefore $X \smallsetminus C$ 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 $\{C_i \colon i \in I\}$ 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

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

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

**Solution 1.14.** 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 $\{V_u \colon u \in U\}$ of open sets.

**Solution 1.15.** If $U$ is open, then it is an open neighbourhood of its elements by definition.

Conversely, suppose $U$ is a neighbourhood of every element of itself. If $x$ is an element of $U$, then $U$ contains some open neighbourhood $V_x$ of $x$. Now $U = \bigcup_{x \in U} V_x$, so $U$ is open.

**Solution 1.16.** 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.

**Solution 1.17.**

(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

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

Then $\big(Y \smallsetminus C\big) \subseteq Y$ is open and $f$ is continuous, so $f^{-1}(Y \smallsetminus C) \subseteq X$ 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

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

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

**Solution 1.18.** Suppose $f\colon X \longrightarrow Y$ is continuous. If $x$ is a point in $X$ and $N$ is a neighbourhood of $f(x)$, then $N$ contains some open neighbourhood $U$ of $f(x)$, whose inverse image $f^{-1}(U)$ is an open neighbourhood of $x$ because of continuity. Since $f^{-1}(U) \subseteq f^{-1}(N)$, it follows that $f^{-1}(N)$ is a neighbourhood of $x$.

Conversely, suppose $f\colon X \longrightarrow Y$ is continuous at every point of $X$. If $U$ be an open subset of $Y$, then $f^{-1}(U)$ is a neighbourhood of every element of itself. By Exercise 1.15, this implies $f^{-1}(U)$ is open. Hence $f$ is continuous.

**Solution 1.19. (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 $x \in U^2_x$ and $U^2_x \in \mathcal{T}_2 \subseteq \mathcal{T}_1$, we can take $U^1_x = U^2_x$ and we are done.

**(b)⇒(a):** Let $U \in \mathcal{T}_2$. We use Exercise 1.14 to prove that $U \in \mathcal{T}_1$. Let $x \in U$, then setting $U^2_x = U$ we have that $U^2_x$ is a $\mathcal{T}_2$-open neighbourhood of $x$, so by (b) there exists a $cT_1$-open neighbourhood $U^1_x$ of $x$ such that $U^1_x \subseteq U$. By Exercise 1.14 we conclude that $U$ is open in the topology $\mathcal{T}_1$.

**Solution 1.20.** 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.

**Solution 1.21.** 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 $\iota$ is continuous.

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

**Solution 1.22.** The 'only if' part follows directly from the definition of continuity.

Conversely, suppose that the inverse image of every member of $S$ is open. It follows that the final topology $\mathcal{T}'_Y$ induced by $f$ (see Tutorial Question 2.7) contains $S$, and is thus finer than $\mathcal{T}_Y$ by Tutorial Question 2.4. By part (b) of Tutorial Question 2.7, this implies that $f$ is continuous.

**Solution 1.23.** (a) We start with proving that $\mathcal{T}_X$ is a topology:

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

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

- If $f^{-1}(U_1), \dots, f^{-1}(U_n)$ are members of $\mathcal{T}_X$, 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 $X$ such that $f$ is continuous, then $f^{-1}(U) \in \mathcal{T}$ for every member $U$ of $\mathcal{T}_Y$, and thus $\mathcal{T}_X \subseteq \mathcal{T}$. Therefore, $\mathcal{T}_X$ is the coarsest 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 finer than $\mathcal{T}_X$. If $U$ is a member of $\mathcal{T}_Y$, then $f^{-1}(U) \in \mathcal{T}_X \subseteq \mathcal{T}$. Hence $f$ is continuous.

(c) Let $\mathcal{T}'_X$ be the topology on $X$ generated by the set

$$\{f^{-1}(U) : U \in S\}.$$

Since the topology $\mathcal{T}_X$ contains $f^{-1}(U)$ for every member $U$ of $S$, it follows from Tutorial Question 2.4 that $\mathcal{T}'_X \subseteq \mathcal{T}_X$. By Exercise 1.22, $f$ is continuous when the topology on $X$ is $\mathcal{T}'_X$, so part (a) implies that $\mathcal{T}_X \subseteq \mathcal{T}'_X$. Hence $\mathcal{T}'_X = \mathcal{T}_X$.

**Solution 1.24.** Let $f : X \times \{y\} \longrightarrow X$ be the map $f(x, y) = x$ and let $g : X \longrightarrow X \times \{y\}$ 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 2.18. 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 $X$.

**Solution 1.25.**

(a) We need to check that $f^{-1} : Y \longrightarrow X$ 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 2.17, 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 $\pi_2$.

**Solution 1.26.** Let $U = X \smallsetminus \{x\}$ 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 1.14 that $U$ is open, so its complement $\{x\}$ is closed.

## Interior and closure

**Solution 1.27.** Take $X = \{0, 1\}$ with the discrete metric, $x = 0$ and $\varepsilon = 1$. Then

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

**Solution 1.28.**

(a) Since $A$ and $B$ are closed in $X$ and $Y$ respectively, their complements $X \smallsetminus A$ and $Y \smallsetminus B$ are open in $X$ and $Y$ respectively, and therefore $(X \smallsetminus A) \times Y$ and $X \times (Y \smallsetminus B)$ are open in $X \times Y$. It follows that

$$(X \times Y) \smallsetminus (A \times B) = \big((X \smallsetminus A) \times Y\big) \cup \big(X \times (Y \smallsetminus B)\big)$$

is open in $X \times Y$.

(b) By part (a), $\overline{A} \times \overline{B}$ is closed in $X \times Y$. Since $A \times B \subseteq \overline{A} \times \overline{B}$, it follows that $\overline{A \times B} \subseteq \overline{A} \times \overline{B}$. It remains to prove the other inclusion.

Given an element $x$ of $A$, define $\iota_x : Y \longrightarrow X \times Y$ by $\iota_x(y) = (x, y)$. Let $\pi_X : X \times Y \longrightarrow X$ and $\pi_Y : X \times Y \longrightarrow Y$ be the projections. The composite function $\pi_X \circ \iota_x$ is the constant function sending every element of $Y$ to $x$, which is continuous by Exercise 1.16; while $\pi_Y \circ \iota_x$ is the identity function of $Y$, which is continuous by Exercise 1.21. it then follows from Tutorial Question 3.8 that $\iota_x$ is continuous.

Since $\overline{A \times B}$ is closed in $X \times Y$, it follows from Exercise 1.17 that $\iota_x^{-1}(\overline{A \times B})$ is closed. Now $B \subseteq \iota_x^{-1}(\overline{A \times B})$ implies $\overline{B} \subseteq \iota_x^{-1}(\overline{A \times B})$; in other words, $\{x\} \times \overline{B} \subseteq \overline{A \times B}$. Since $x$ is an arbitrary point in $A$, this implies $A \times \overline{B} \subseteq \overline{A \times B}$.

Following similar reasoning for points in $\overline{B}$, we can show that $\overline{A} \times \overline{B} \subseteq \overline{A \times B}$.

(c) Use (b).

**Solution 1.29.** 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) $x \longmapsto 0$.

**Solution 1.30.** Since $A^\circ \subseteq A$, we have $\big(X \smallsetminus A\big) \subseteq \big(X \smallsetminus A^\circ\big)$. But $A^\circ$ is open, so $X \smallsetminus A^\circ$ is a closed set containing $X \smallsetminus A$, hence

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

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

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

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

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

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

**Solution 1.31.** First we show that $\overline{\mathbf{Z}} = \mathbf{Z}$: letting $U = \mathbf{R} \smallsetminus \mathbf{Z}$, we have

$$U = \bigcup_{n \in \mathbf{Z}} (n - 1, n),$$

so $U$ is a union of open subsets, hence open.

Now we note that $\mathbf{Z}^\circ = \varnothing$: if $V \subseteq \mathbf{R}$ is a nonempty open subset, then $V$ contains a nonempty open interval, hence is uncountable, so it cannot be contained in $\mathbf{Z}$.

**Solution 1.32.**

(a) Let $N \subseteq X$ be nowhere dense and let $M \subseteq N$. Then $\overline{M} \subseteq \overline{N}$ by part (a) of Tutorial Question 3.1, so $\left(\overline{M}\right)^\circ \subseteq \left(\overline{N}\right)^\circ = \varnothing$ by part (a) of Tutorial Question 3.1.

(b) Suppose $N$ is nowhere dense and let $U \subseteq X$ be nonempty and open. If $U \cap \left(X \smallsetminus \overline{N}\right) = \varnothing$, then $U \subseteq \overline{N}$, so $U \subseteq \left(\overline{N}\right)^\circ = \varnothing$, contradicting the non-emptiness of $U$. So it must be that $U$ intersects $X \smallsetminus \overline{N}$ nontrivially, hence $X \smallsetminus \overline{N}$ is dense.

Conversely, suppose $X \smallsetminus \overline{N}$ is dense but $N$ is not nowhere dense, that is there exists a nonempty open $U \subseteq \overline{N}$. Then $U \cap \left(X \smallsetminus \overline{N}\right) = \varnothing$, contradicting the denseness of $X \smallsetminus \overline{N}$.

(c) It suffices to prove the case of two nowhere dense sets $M$ and $N$. Let $L = M \cup N$. Then by part (b) of Tutorial Question 3.1 we have $\overline{L} = \overline{M} \cup \overline{N}$ so $X \smallsetminus \overline{L} = \left(X \smallsetminus \overline{M}\right) \cap \left(X \smallsetminus \overline{N}\right)$. As $X \smallsetminus \overline{L}$ is the intersection of two dense open subsets, it is dense and open by Tutorial Question 3.2, hence $L$ is nowhere dense.

## Metric topologies

**Solution 1.33.**

(a) i. Put $X = \{0,1\}$ with $\mathcal{T}_X = \big\{\varnothing, \{0,1\}\big\}$, then $\mathcal{T}_X$ is not metrisable (see Tutorial Question 2.3).

Let $Y = \{1\}$ and let $f \colon X \longrightarrow Y$ be the function sending both 0 and 1 to 1. The induced topology $\mathcal{T}_Y = \mathcal{P}(Y)$ is the discrete topology and hence metrisable (see Tutorial Question 2.1).

ii. Put $X = \mathbf{R}$ with the Euclidean topology, hence metrisable.

Let $Y = \{0,1\}$ and let $f \colon X \longrightarrow Y$ be given by $f(\mathbf{Q}) = \{0\}$ and $f(\mathbf{R} \smallsetminus \mathbf{Q}) = \{1\}$. The induced topology $\mathcal{T}_Y$ is trivial (every nonempty open in $\mathbf{R}$ intersects both $\mathbf{Q}$ and $\mathbf{R} \smallsetminus \mathbf{Q}$, so gets mapped by $f$ onto $Y$), and hence not metrisable (see Tutorial Question 2.3).

(b) i. Let $Y = \mathbf{Z}$ with the cofinite topology, hence not Hausdorff by Example 2.19, therefore not metrisable by Example 2.27.

Let $X = \{0,1\}$ and let $f \colon X \longrightarrow Y$ be the inclusion map. The induced topology $\mathcal{T}_X$ is discrete and hence metrisable by Tutorial Question 2.1.

ii. Let $Y = \{1\}$ with the discrete topology, hence metrisable by Tutorial Question 2.1.

Let $X = \{0,1\}$ and let $f \colon X \longrightarrow Y$ be the function sending both 0 and 1 to 1. The induced topology $\mathcal{T}_X$ is trivial and hence not metrisable (see Tutorial Question 2.3).

**Solution 1.34.** 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 $f(x') \in \mathbf{B}_\varepsilon(f(x))$.

**Solution 1.35.**

(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 $\mathbf{B}_r^{d_2}(x)$ of $\mathcal{T}_2$. I claim that the open ball $\mathbf{B}_{r/M}^{d_1}(x)$ of $\mathcal{T}_1$ is contained in $\mathbf{B}_r^{d_2}(x)$: if $y \in \mathbf{B}_{r/M}^{d_1}(x)$ 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 $\mathbf{B}_r^{d_1}(x)$ of $\mathcal{T}_1$. I claim that the open ball $\mathbf{B}_{rm}^{d_2}(x)$ of $\mathcal{T}_2$ is contained in $\mathbf{B}_r^{d_1}(x)$: if $y \in \mathbf{B}_{rm}^{d_2}(x)$ 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 (1.1) 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 $y = 0$, $x = \lceil M \rceil + 1$).

**Solution 1.36.** The inequalities involving $d_1$ and $d_\infty$ 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 1.35.

**Solution 1.37.**

(a) We have

$$\begin{aligned} \mathbf{B}_r^X(y) &= \{x \in X \colon d(x,y) < r\} \\ \mathbf{B}_r^Y(y) &= \{x \in Y \colon d(x,y) < r\}, \end{aligned}$$

so that

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

(b) In one direction, suppose $A$ is open in $Y$; by Tutorial Question 3.4 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

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

which by Tutorial Question 3.4 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 $\mathbf{B}^X_r(a)$ such that $\mathbf{B}^X_r(a) \subseteq U$. Consider $\mathbf{B}^Y_r(a) = \mathbf{B}^X_r(a) \cap Y \subseteq U \cap Y = A$. So every point $a \in A$ is contained in an open (in $Y$) ball, hence $A$ is open in $Y$.

**Solution 1.38.** 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,

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

**Solution 1.39.** We need to show that $d$ induces the discrete topology on $X$. It suffices to prove that any singleton $\{x\} \subseteq X$ is an open set with respect to the metric $d$.

Fix $x \in X$. The set of distances $d(x,y)$ with $y \neq x$ is finite, so it has a minimum element, wihich is $> 0$; call it $D$, so that $d(x,y) \leqslant D$ for all $y \in X$. Then $\mathbf{B}_D(x) = \{x\}$, which is therefore an open set.

## Connectedness

**Solution 1.40.** 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.

**Solution 1.41.** It follows from Example 2.28 that $X$ is connected if it is a singleton.

Conversely, if $x_1 \neq x_2$ are elements of $X$, then $\{x_1\}$ and $X \smallsetminus \{x_1\}$ are two disjoint non-empty open subsets of $X$ such that their union is $X$, so $X$ is disconnected.

**Solution 1.42.** Let $f \colon \bigcup_{n \in \mathbf{N}} C_n \longrightarrow \{0,1\}$ 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 2.31.

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 2.31 that $f(x) = f(x') = f(x_0)$.

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

**Solution 1.43.** Let $x \in D$.

Let $f\colon X \longrightarrow \{0,1\}$ be a continuous function, where $\{0,1\}$ is given the discrete topology. Since $f$ is continuous, it follows from Exercise 1.17 that $f^{-1}\big(f(x)\big)$ is closed. By Proposition 2.31, the restriction of $f$ to $D$ is constant, so $D \subseteq f^{-1}\big(f(x)\big)$, and therefore $X = \overline{D} \subseteq f^{-1}\big(f(x)\big)$. Hence $f$ is constant, which implies that $X$ is connected.

(Alternative solution): Suppose $X$ is disconnected, so $X = U \cup V$ with $U, V$ open, non-empty, and disjoint. Then $D \subseteq U \cup V$ with $D \cap U \neq \varnothing$, $D \cap V \neq \varnothing$ (because $D$ is dense), and of course $D \cap U \cap V = \varnothing$, implying that $D$ is a disconnected subset of $X$ by Exercise 1.40.

**Solution 1.44.** Let $f\colon A \cup \bigcup_{i\in I} C_i \longrightarrow \{0,1\}$ be a continuous function, where $\{0,1\}$ is given the discrete topology. Pick an element $a$ of $A$ and consider an arbitrary element $x$ of $A \cup \bigcup_{i\in I} C_i$. If $x \in A$, then the connectedness of $A$ and Proposition 2.31 imply $f(x) = f(a)$. If $x \in C_i$ for some $i \in I$, then it follows from Tutorial Question 3.6 and Proposition 2.31 that $f(x) = f(a)$. Hence $f$ is constant, which implies $A \cup \bigcup_{i\in I} C_i$ is connected.

**Solution 1.45.** Suppose $X \times Y$ is connected. Recall from Proposition 2.18 that the projections $\pi_X\colon X \times Y \longrightarrow X$ and $\pi_Y\colon X \times Y \longrightarrow Y$ are continuous. It then follows from Proposition 2.32 that $X = \pi_X(X \times Y)$ and $Y = \pi_Y(X \times Y)$ are connected.

Conversely, suppose that both $X$ and $Y$ are connected. Let $f\colon X \times Y \longrightarrow \{0,1\}$ be a continuous function, where $\{0,1\}$ is given the discrete topology. Consider two elements $(x_1, y_1)$ and $(x_2, y_2)$ of $X \times Y$. It follows from Exercise 1.24 that $\{x_1\} \times Y$ is homeomorphic to $Y$, and is therefore connected. This implies that $f$ is constant when restricted to $\{x_1\} \times Y$. Similarly, $f$ is constant when restricted to $X \times \{y_2\}$ because $Y$ is connected. Hence

$$f(x_1, y_1) = f(x_1, y_2) = f(x_2, y_2),$$

and therefore $X \times Y$ is connected.

**Solution 1.46.** Let $X$ be a totally separated space and let $S$ be a subset of $X$ with two distinct points $x$ and $y$. It follows from total separatedness that there exists disjoint clopen neighbourhoods $U$ and $V$ of $x$ and $y$ respectively. Since $U$ is clopen and does not contain $y$, it follows that $X \setminus U$ is a clopen neighbourhood of $y$. Moreover, $S \cap U$ and $S \cap (X \setminus U)$ are two disjoint open sets in $S$ such that their union is $S$. Hence $S$ is not connected, and therefore the only connected subsets of $X$ are the singletons; in other words, $X$ is totally disconnected.

**Solution 1.47.** We will prove all of them are totally separated, which implies total disconnectedness by Exercise 1.46.

(a) Let $x$ and $y$ be two distinct rational number. Without loss of generality, we assume that $x < y$. The denseness of $\mathbf{R} \setminus \mathbf{Q}$ (see Example 2.23) implies that there exists an irrational number $z$ such that $x < z < y$. The open sets $\mathbf{Q} \cap (-\infty, z)$ and $\mathbf{Q} \cap (z, \infty)$ are open in $\mathbf{Q}$, and their intersection is empty while their union is $\mathbf{Q}$, so $\mathbf{Q} \cap (-\infty, z)$ is an clopen neighbourhood of $x$ in $\mathbf{Q}$ and $\mathbf{Q} \cap (z, \infty)$ is a clopen neighbourhood of $y$. Hence $\mathbf{Q}$ is totally disconnected when equipped with the Euclidean topology.

(b) Let $X$ be a discrete space and let $x$ and $y$ be two points in $X$. It follows from the definition of the discrete topology that $\{x\}$ and $\{y\}$ are clopen, so they are disjoint clopen neighbourhoods of $x$ and $y$ respective. Hence $X$ is totally separated.

## Compactness

**Solution 1.48.** Suppose $X$ is compact and $\{C_i : i \in I\}$ is a collection of closed sets with the finite intersection property. Suppose that

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

Then

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

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 $\{C_i : i \in I\}$.

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

**Solution 1.49.**

(a) We know that $t \longmapsto 2\pi t$, $t \longmapsto \cos(t)$ and $t \longmapsto \sin(t)$ are continuous, so by Tutorial Question 3.8 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 $f^{-1} : \mathbf{S}^1 \longrightarrow [0,1)$ 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

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

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.

**Solution 1.50.** $Y$ 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 $X \smallsetminus \{x\} \longrightarrow Z \smallsetminus \{f(x)\}$ is then also a homeomorphism, but this is impossible since $X \smallsetminus \{x\} = [-1,x) \cup (x,1]$ is disconnected, while $Z \smallsetminus \{f(x)\}$ is connected.

**Solution 1.51.**

(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.

**Solution 1.52.** TODO: maybe include a direct proof of this first direction as an alternative?

Suppose $K$ is compact as a topological space with the subspace topology from $X$. Let $\iota\colon K \longrightarrow X$ be the inclusion function, which is continuous by Exercise 1.21. It then follows from Proposition 2.39 that $\iota(K)$ is a compact subset of $X$.

Conversely, suppose $K$ is a compact subset of $X$. Let $\{U_i : i \in I\}$ be an open cover of $K$ in the subspace $K$. By the definition of the subspace topology, for every $U_i$ there exists an open subset $V_i$ of $X$ such that $U_i = K_i \cap K$. Since $\{U_i : i \in I\}$ is an open cover of $K$ in the subspace $K$, it follows that $\{V_i \in i \in I\}$ is an open cover of $K$ in $X$. The compactness of $K$ as a subset of $X$ then implies there exists a finite subset $J$ of $I$ such that $K \subseteq \bigcup_{j\in J} V_j$, and therefore

$$K = K \cap \Big(\bigcup_{j\in J} V_j\Big) = \bigcup_{j\in J}(K \cap V_j) = \bigcup_{j\in J} U_j.$$

Hence $\{U_i : i \in I\}$ has a finite sub-cover, which implies $K$ is compact as a subspace of $X$.

## Sequences

**Solution 1.53.** 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 $(x_n) \sim (z_n)$.

**Solution 1.54.** 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 $x = x'$.

**Solution 1.55.**

(a) It is clear that $\varnothing$ and $\mathbf{N}^*$ belong to $\mathcal{T}$.

Suppose $\{U_i \colon i \in I\}$ is a collection of members of $\mathcal{T}$. If $\{U_i \colon i \in I\} \subseteq \mathcal{P}(\mathbf{N})$, then $\bigcup_{i\in I} U_i \in \mathcal{P}(\mathbf{N}) \subseteq \mathcal{T}$. Otherwise, there exists a member $V$ of $\{U_i \colon i \in I\}$ such that $\infty \in V$. It then follows from

$$\mathbf{N}^* \smallsetminus \Big(\bigcup_{i\in I} U_i\Big) \subseteq \mathbf{N}^* \smallsetminus V$$

that $\mathbf{N}^* \smallsetminus \left(\bigcup_{i\in I} U_i\right)$ is finite, and therefore $\bigcup_{i\in I} U_i \in \mathcal{T}$.

For closure under finite intersection, it suffices to prove it for any two members $U$ and $V$ of $\mathcal{T}$. If at most one of $U$ and $V$ contains $\infty$, then $U \cap V \in \mathcal{P}(\mathbf{N})$. Otherwise, it then follows from

$$\mathbf{N}^* \smallsetminus (U \cap V) = (\mathbf{N}^* \smallsetminus U) \cup (\mathbf{N}^* \smallsetminus V)$$

that $\mathbf{N}^* \smallsetminus (U \cap V)$ is finite, and therefore $U \cap V \in \mathcal{T}$.

(b) Let $\{U_i \colon i \in I\}$ be an open cover of $\mathbf{N}^*$. Pick a member $V$ of the open cover such that $\infty \in V$. Since $V \in \mathcal{T}$, it follows that $\mathbf{N}^* \smallsetminus V$ is finite. For each element $x$ of $\mathbf{N}^* \smallsetminus V$, pick a member $V_x$ of the open cover such that $x \in V_x$. It follows that $\{V\} \cup \{V_x \colon x \in \mathbf{N}^* \smallsetminus V\}$ is a finite sub-cover of $\{U_i \in i \in I\}$. Hence $\mathbf{N}^*$ is compact.

(c) Suppose $f$ is continuous. It follows that for every positive real number $\varepsilon$, the inverse image $f^{-1}\big(\mathbf{B}_\varepsilon(f(\infty))\big)$ is open, and therefore $\mathbf{N}^* \smallsetminus f^{-1}\big(\mathbf{B}_\varepsilon(f(\infty))\big)$ is finite. Hence there exists a natural number $N$ such that $n \geqslant N$ implies $f(n) \in \mathbf{B}_\varepsilon(f(\infty))$.

Conversely, suppose $\big(f(n)\big)$ converges to $f(\infty)$. The space $\mathbf{N}$ is discrete as a subspace of $\mathbf{N}^*$, so $f|_{\mathbf{N}}$ is continuous; this implies $f$ is continuous at every natural number by Exercise 1.18. To apply Exercise 1.18, it remains to prove $f$ is continuous at $\infty$. Let $M$ be a neighbourhood of $\infty$ and pick a positive real number $\varepsilon$ such that $\mathbf{B}_\varepsilon(f(\infty)) \subseteq M$. Since $f(n) \longrightarrow f(\infty)$ as $n \longrightarrow \infty$, there exists a natural number $N$ such that $n \geqslant N$ implies $f(n) \in \mathbf{B}_\varepsilon(f(\infty))$. This implies

$$\mathbf{N}^* \smallsetminus f^{-1}\big(\mathbf{B}_\varepsilon(f(\infty))\big) \subseteq \{1, \dots, N\},$$

so $f^{-1}\big(\mathbf{B}_\varepsilon(f(\infty))\big)$ is open. Since $f^{-1}\big(\mathbf{B}_\varepsilon(f(\infty))\big) \subseteq f^{-1}(M)$, it follows that $f^{-1}(M)$ is a neighbourhood of $\infty$, so $f$ is continuous at $\infty$. Now apply Exercise 1.18 to $f$, we see that $f$ is continuous.

(d) Define a function $f \colon \mathbf{N}^* \longrightarrow X$ by

$$f(n) = \begin{cases} x_n & \text{if } n \in \mathbf{N}, \\ x & \text{otherwise.} \end{cases}$$

By part (c), $f$ is continuous, so it follows from Proposition 2.39 that

$$\{x\} \cup \{x_n \colon n \in \mathbf{N}\} = f(\mathbf{N}^*)$$

is compact.

**Solution 1.56.** First note that for any $n, m$ we have by the triangle inequality:

$$d(x_n, y_n) \leqslant d(x_n, x_m) + d(x_m, y_n) \leqslant d(x_n, x_m) + d(x_m, y_m) + d(y_m, y_n),$$

so

$$d(x_n, y_n) - d(x_m, y_m) \leqslant d(x_n, x_m) + d(y_m, y_n).$$

Similarly:

$$d(x_m, y_m) \leqslant d(x_m, x_n) + d(x_n, y_n) + d(y_n, y_m)$$

so that

$$-\big(d(x_m, x_n) + d(y_n, y_m)\big) \leqslant d(x_n, y_n) - d(x_m, y_m).$$

We can summarise this as

$$\big|d(x_n, y_n) - d(x_m, y_m)\big| \leqslant d(x_m, x_n) + d(y_n, y_m).$$

Let $\varepsilon > 0$. There exists $N_1 \in \mathbf{N}$ such that $d(x_n, x_m) < \varepsilon/2$ for all $m, n \geqslant N_1$. There exists $N_2 \in \mathbf{N}$ such that $d(y_n, y_m) < \varepsilon/2$ for all $m, n \geqslant N_2$. Let $N = \max\{N_1, N_2\}$, then for all $n, m \geqslant N$ we have:

$$\big|d(x_n, y_n) - d(x_m, y_m)\big| \leqslant d(x_n, x_m) + d(y_m, y_n) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

So $\big(d(x_n, y_n)\big)$ is a Cauchy sequence in $\mathbf{R}$.

**Solution 1.57.** It suffices to prove that $(x_n)$ being Cauchy implies $(y_n)$ is Cauchy.

Let $\varepsilon > 0$. As $(y_n) \sim (x_n)$, there exists $N_1 \in \mathbf{N}$ such that $d(y_n, x_n) < \varepsilon/3$ for all $n \geqslant N_1$. As $(x_n)$ is Cauchy, there exists $N_2 \in \mathbf{N}$ such that $d(x_n, x_m) < \varepsilon/3$ for all $n, m \geqslant N_2$. Let $N = \max\{N_1, N_2\}$, then for all $n, m \geqslant N$ we have

$$d(y_n, y_m) \leqslant d(y_n, x_n) + d(x_n, x_m) + d(x_m, y_m) < \varepsilon.$$

## Uniform continuity and completeness

**Solution 1.58.**

(a) Let $f', g' \colon \overline{D} \longrightarrow X$ denote the restrictions of $f$ and $g$ to $\overline{D}$ respectively. Since $D$ is dense in $\overline{D}$, and the functions $f'$ and $g'$ agree on $D$, property H implies that $f' = g'$. Hence the result follows.

(b) Let $X$ be a topological space, $Y$ a Hausdorff topological space, $D$ a dense subset of $Y$. Consider two continuous functions $f, g \colon Y \longrightarrow X$ that agree on $D$.

Let $(f, g) \colon Y \longrightarrow X \times X$ be the function defined by $(f, g)(y) = (f(y), g(y))$ (see Tutorial Question 3.8). Since both $f$ and $g$ are continuous, it follows from Tutorial Question 3.8 that $(f, g)$ is continuous.

Let $\Delta$ denote the diagonal function of $X$, defined in Tutorial Question 3.9. By part (c) of Tutorial Question 3.9, the Hausdorffness of $X$ implies that $\Delta(X)$ is closed in $X \times X$. It then follows from Exercise 1.17 that $(f, g)^{-1}\big(\Delta(X)\big)$ is closed.

Since $f(y) = g(y)$ for every element of $y$ of $D$, it follows that $(f, g)(D) \subseteq \Delta(X)$, and thus $D \subseteq (f, g)^{-1}\big(\Delta(X)\big)$. However, $(f, g)^{-1}\big(\Delta(X)\big)$ is closed, so it contains the closure $Y$ of $D$. This implies that $\big(f(y), g(y)\big) \in \Delta(X)$ for every element $y$ of $Y$; in other words, $f(y) = g(y)$ for every element $y$ of $Y$.

(c) In part (b), we have shown that Hausdorffness implies property H, so it suffices to prove the other direction.

Let $X$ be a topological space with property H and $\Delta$ the diagonal function of $X$, defined in Tutorial Question 3.9. Define two projections $\pi_1\colon X \times X \longrightarrow X$ and $\pi_2\colon X \times X \longrightarrow X$ by $\pi_1(x,y) = x$ and $\pi_2(x,y) = y$. It follows from Proposition 2.18 that $\pi_1$ and $\pi_2$ are both continuous. The projections $\pi_1$ and $\pi_2$ agree on $\Delta(X)$ by definition, so they agree on the closure $\overline{\Delta(X)}$ of $\Delta(X)$ by part (a). Therefore, if $(x,y) \in \overline{\Delta(X)}$, then

$$x = \pi_1(x,y) = \pi_2(x,y) = y,$$

so $(x,y) \in \Delta(X)$. It follows that $\overline{\Delta(X)} = \Delta(X)$.

(d) Follows immediately from the above.

**Solution 1.59.** (a) Let $\varepsilon$ be a positive real number. Put $\delta = \varepsilon$. If elements $x_1$ and $x_2$ of $S$ satisfy $d_S(x_1,x_2) < \delta$, then

$$d_X\big(\iota_S(x_1), \iota_S(x_2)\big) = d_S(x_1,x_2) < \varepsilon.$$

Hence $\iota_S$ is uniformly continous.

(b) If $f$ is uniformly continuous, then $\iota_S \circ f$ is uniformly continuous because of part (a) and Tutorial Question 5.2.

Conversely, suppose $\iota_S \circ f$ is uniformly continuous. Let $\varepsilon$ be a positive real number. Pick a positive real number $\delta$ such that $d_Y(y_1,y_2) < \delta$ implies $d_X\big((\iota_S \circ f)(y_1), (\iota_S \circ f)(y_2)\big) < \varepsilon$. It follows that $d_Y(y_1,y_2) < \delta$ implies

$$d_S\big(f(y_1), f(y_2)\big) = d_X\big((\iota_S \circ f)(y_1), (\iota_S \circ f)(y_2)\big) < \varepsilon.$$

Hence $f$ is uniformly continuous.

**Solution 1.60.** (a) If $\varepsilon$ is a positive real number, then $d\big((y_1,z_1),(y_2,z_2)\big) < \varepsilon$ implies

$$d_Y\big(\pi_Y(y_1,z_1), \pi_Y(y_2,z_2)\big) = d_Y(y_1,y_2) \leqslant d\big((y_1,z_1),(y_2,z_2)\big) < \varepsilon$$

and similarly $d\big(\pi_Z(y_1,z_1), \pi_Z(y_2,z_2)\big) < \varepsilon$. Hence $\pi_Y$ and $\pi_Z$ are uniformly continuous.

(b) If $f$ is uniformly continuous, then it follows from Tutorial Question 5.2 and part (a) that both $\pi_Y \circ f$ and $\pi_Z \circ f$ are uniformly continuous.

Conversely, suppose both $\pi_Y \circ f$ and $\pi_Z \circ f$ are uniformly continuous. Let $\varepsilon$ be a positive real number. It follows from the uniform continuity of $\pi_Y \circ f$ and $\pi_Z \circ f$ that there exist positive real numbers $\delta_Y$ resp. $\delta_Z$ such that $d_X(x_1,x_2) < \delta_Y$, resp. $d_X(x_1,x_2) < \delta_Z$ imply

$$d_Y\big((\pi_Y \circ f)(x_1), (\pi_Y \circ f)(x_2)\big) < \varepsilon/2 \quad \text{resp.} \quad d_Z\big((\pi_Z \circ f)(x_1), (\pi_Z \circ f)(x_2)\big) < \varepsilon/2.$$

Let $\delta = \min\{\delta_Y, \delta_Z\}$. It follows that $d_X(x_1,x_2) < \delta$ implies

$$d\big(f(x_1), f(x_2)\big) \leqslant d_Y\big((\pi_Y \circ f)(x_1), (\pi_Y \circ f)(x_2)\big) + d_Z\big((\pi_Z \circ f)(x_1), (\pi_Z \circ f)(x_2)\big) < \varepsilon,$$

so $f$ is uniformly continuous.

**Solution 1.61.** Let $g\colon Y \longrightarrow X$ denote the inverse of $f$.

(a) By Proposition 2.52, $(x_n)$ being Cauchy implies $\big(f(x_n)\big)$ is Cauchy, while $\big(f((x_n)\big)$ being Cauchy implies $(x_n) = \Big(g\big(f(x_n)\big)\Big)$ is Cauchy.

(b) Since $X$ and $Y$ are interchangeable, it suffices to prove one direction. Suppose $X$ is complete and $(y_n)$ is a Cauchy sequence in $Y$. It follows that $\big(g(y_n)\big)$ is Cauchy in $X$, and therefore converges to some point $x$ in $X$ By Theorem 2.44, $(y_n) = \Big(f\big(g(y_n)\big)\Big)$ converges to $f(x)$. Hence $Y$ is complete.

(c) Since $f$ has an inverse $\tan\colon (-\pi/2, \pi/2) \longrightarrow \mathbf{R}$, and both $f$ and tan are continuous. Hence $f$ is a homeomorphism.

Given $x_1 < x_2$, apply the Mean Value Theorem to $f(x) = \arctan(x)$ on $[x_1, x_2]$ to get some $\xi \in (x_1, x_2)$ such that

$$|f(x_2) - f(x_1)| = |f'(\xi)|\,|x_2 - x_1| = \frac{1}{1+\xi^2}\,|x_2 - x_1| \leqslant |x_2 - x_1|.$$

So for any $\varepsilon > 0$ we can take $\delta = \varepsilon$ and conclude that $f$ is uniformly continuous.

However, its inverse $\tan\colon (-\pi/2, \pi/2) \longrightarrow \mathbf{R}$ is not uniformly continuous, because $(-\pi/2, \pi/2)$ is totally bounded (since bounded in $\mathbf{R}$), but $\mathbf{R}$ is not totally bounded. (Use Proposition 2.64.)

(d) The codomain $(-\pi/2, \pi/2)$ is not complete because $(\pi/2 - 1/n)$ is Cauchy but does not converge in $(-\pi/2, \pi/2)$. However, the domain $\mathbf{R}$ is complete.

**Solution 1.62.**

(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

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

(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 $\sqrt{2} \notin \mathbf{Q}$.

**Solution 1.63.**

(a) Suppose $\big((x_n, y_n)\big)$ is a Cauchy sequence in $(X \times Y, d)$. By part (a) of Exercise 1.60, both projections $\pi_X : X \times Y \longrightarrow X$ and $\pi_Y : X \times Y \longrightarrow Y$ are uniformly continuous. Hence $(x_n) = \big(\pi_X(x_n, y_n)\big)$ and $(y_n) = \big(\pi_Y(x_n, y_n)\big)$ are Cauchy because of Proposition 2.52.

Conversely, suppose $(x_n)$ is Cauchy in $X$ and $(y_n)$ is Cauchy in $Y$. Fix $\varepsilon > 0$. Let $N_x \in \mathbf{N}$ be such that for all $m, n \geqslant N_x$ we have $d_X(x_m, x_n) < \varepsilon$. Let $N_y \in \mathbf{N}$ be such that for all $m, n \geqslant N_y$ we have $d_Y(y_m, y_n) < \varepsilon$. Let $N = \max\{N_x, N_y\}$, then for all $m, n \geqslant N$ we have

$$d\big((x_m, y_m), (x_n, y_n)\big) = \max\left\{d_X(x_m, x_n), d_Y(y_m, y_n)\right\} < \varepsilon,$$

so $\big((x_n, y_n)\big)$ is Cauchy in $X \times Y$.

(b) Let $\big((x_n, y_n)\big)$ be a Cauchy sequence in $X \times Y$. By part (a), $(x_n)$ is Cauchy in $X$ and $(y_n)$ is Cauchy in $Y$. Since $X$ and $Y$ are complete, we have $(x_n) \longrightarrow x \in X$ and $(y_n) \longrightarrow y \in Y$. By Tutorial Question 4.9, $\big((x_n, y_n)\big) \longrightarrow (x, y) \in X \times Y$.

The converse also holds: suppose $X \times Y$ is complete. Let $(x_n)$ be a Cauchy sequence in $X$, and fix some $y \in Y$. Then by (a) we have that $\big((x_n, y)\big)$ is Cauchy in $X \times Y$, so $\big((x_n, y)\big) \longrightarrow (x, y) \in X \times Y$, which by Tutorial Question 4.9 implies that $(x_n) \longrightarrow x \in X$. The same proof gives us that $Y$ is complete.

**Solution 1.64.**

(a) Let $\varepsilon > 0$. Set $\delta = \varepsilon$. If $x, x' \in X$ satisfy $d(x, x') < \delta = \varepsilon$, then

$$|f(x) - f(x')| = |d(x, y) - d(x', y)| \leqslant d(x, x') < \varepsilon.$$

(b) Let $\varepsilon > 0$. By part (a), there exists positive real numbers $\delta_1$ and $\delta_2$ such that $d(x_1, x_1') < \delta_1$ and $d(x_2, x_2') < \delta_2$ imply

$$d_{\mathbf{R}}\big(d(x_1, x_2), d(x_1', x_2)\big) < \varepsilon/2 \quad \text{and} \quad d_{\mathbf{R}}\big(d(x_1', x_2), d(x_1', x_2')\big) < \varepsilon/2.$$

Set $\delta = \min\{\delta_1, \delta_2\}$. If $(x_1, x_2), (x_1', x_2') \in X \times X$ satisfy

$$\max\{d(x_1, x_1'), d(x_2, x_2')\} = D\big((x_1, x_2), (x_1', x_2')\big) < \varepsilon$$

then

$$d_{\mathbf{R}}\big(d(x_1, x_2), d(x_1', x_2')\big) \leqslant d_{\mathbf{R}}\big(d(x_1, x_2), d(x_1', x_2)\big) + d_{\mathbf{R}}\big(d(x_1', x_2), d(x_1', x_2')\big) < \varepsilon.$$

Hence $d$ is uniformly continuous.

**Solution 1.65.** This uses the same approach as Proposition 2.48: we have

$$\big|d(x_n', y_n') - d(x_n, y_n)\big| \leqslant d(x_n', x_n) + d(y_n', y_n).$$

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.)

**Solution 1.66.** 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.

**Solution 1.67.**

(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:

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

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:

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

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

$$\big|f_a^{\circ n}(0) - f_b^{\circ n}(0)\big| \leqslant \big(1 + c + \cdots + c^{n-1}\big)\big|a - b\big|. \tag{1.1}$$

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

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

Fix $n \in \mathbf{N}$ and assume that the inequality (1.1) holds for $n$. We have

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

where in the second to last step we used the contractive property of $f$ and the inequality (1.1) for $n$.

Finally, we have

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

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

**Solution 1.68.** 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 $\mathbf{R}$.

**Solution 1.69.** 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 $f(x_0) = x_0$.

## Boundedness and total boundedness

**Solution 1.70.** If $S \subseteq \mathbf{D}_r(x)$ then $\operatorname{diam}(S) \leqslant \operatorname{diam}\big(\mathbf{D}_r(x)\big) = 2r$ 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 $S \subseteq \mathbf{D}_r(x)$.

**Solution 1.71.** 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

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

**Solution 1.72.**

(a) Let $a$ and $b$ be real numbers and let $r = \max\{|a|, |b|\}$. Since $[a, b] \subseteq \mathbf{B}_r(0)$, it follows that $[a, b]$ is bounded, and therefore totally bounded by Example 2.62. As a closed subset of the complete space $\mathbf{R}$, the closed interval $[a, b]$ is also complete. Hence $[a, b]$ is compact by the Heine–Borel theorem (Theorem 2.66).

(b) Let $r$ be a positive real number and let $v$ be an element of $\mathbf{R}^n$.

We start with proving $\mathbf{D}_r(0)$ is compact. Since $[-r, r]$ is compact, it follows from Theorem 2.41 that $[-r, r]^n$ is compact. Since $\mathbf{D}_r(0)$ is a closed subset of $[-r, r]^n$, it follows from Proposition 2.38 that $\mathbf{D}_r(0)$ is compact.

Let $R_v : \mathbf{R}^n \longrightarrow \mathbf{R}^n$ be the continuous function defined by $R_v(w) = v + w$ (see Proposition B.9). Since

$$\mathbf{D}_r(v) = R_v\big(\mathbf{D}_r(0)\big),$$

it follows from Proposition 2.39 that $\mathbf{D}_r(v)$ is compact.

(c) Suppose $K$ is a compact subset of $\mathbf{R}^n$. It follows from Proposition 2.37 that $K$ is closed and it follows from the Heine–Borel theorem (Theorem 2.66) that $K$ is totally bounded, which implies $K$ is bounded by Tutorial Question 7.1.

Conversely, suppose $K$ is a bounded closed subset of $\mathbf{R}^n$. It follows from Exercise 1.70 that $K$ is contained in some closed ball $\mathbf{D}_r(v)$, which is compact by part (b). Hence $K$ is compact since it is a closed subset of a compact set (see Proposition 2.38).

(d) Let $S$ be a bounded subset of $\mathbf{R}^n$ and suppose $S$ is contained in some closed ball $\mathbf{D}_r(v)$ (see Exercise 1.70), which is compact by part (b) and therefore totally bounded by the Heine–Borel theorem (Theorem 2.66). It now follows from part (c) of Tutorial Question 7.4 that $S$ is totally bounded.

**Solution 1.73.** 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

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

**Solution 1.74.** 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 1.70 equivalent to $f(X)$ being a bounded subset of $Y$.

**Solution 1.75.** Let $\varepsilon > 0$ and let

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

be corresponding covers of $S$, respectively $T$.

Then

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

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

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

where $\mathbf{B}^{d_\infty}$ denotes an open ball with respect to the $d_\infty$ 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 1.36.

**Solution 1.76.** 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 1.75 so is their product, hence so is its subset $Z$.

**Solution 1.77.** 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.

**Solution 1.78.**

(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 $n \neq m$.

**Solution 1.79.** As you know from Exercise 1.64, the distance function $d\colon X \times X \longrightarrow \mathbf{R}$ is continuous. By Theorem 2.41, $C \times C$ is compact, so by Proposition 2.63 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 $C$.

**Solution 1.80.**

(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 1.70 that $B$ is contained in some closed ball $\mathbf{D}_r(v)$, which is compact by part (b) of Exercise 1.72. Hence $f\big(\mathbf{D}_r(v)\big)$ is compact by Proposition 2.39, and therefore bounded by part (c) of Exercise 1.72. Since $f(B) \subseteq f\big(\mathbf{D}_r(v)\big)$, it follows that

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

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 $\mathrm{id}_{\mathbf{N}} : X \longrightarrow Y$ is uniformly continuous. Indeed, for every positive real number $\varepsilon$, put $\delta = 1$. If $d_1(x,y) < 1$, then $x = y$, and therefore $d_2\big(\mathrm{id}_X(x), \mathrm{id}_X(y)\big) = 0 < \varepsilon$.

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

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

**Solution 1.81.**

(a) The subset $S$ of $\mathbf{R}^n$ is bounded because of Tutorial Question 7.1, and therefore $f(S)$ is bounded by part (a) of Exercise 1.80. It then follows from part (d) of Exercise 1.72 that $f(S)$ is totally bounded.

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

## Function spaces

**Solution 1.82.** Given $\varepsilon > 0$, let $\delta = \varepsilon$. If $f, g \in B(X,Y)$ satisfy

$$d_\infty(f,g) < \delta = \varepsilon,$$

then

$$d_Y\big(f(x), g(x)\big) \leqslant \sup_{t \in X} d_Y\big(f(t), g(t)\big) = d_\infty(f,g) < \varepsilon.$$

**Solution 1.83.** 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.

**Solution 1.84.** 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

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

**Solution 1.85.**

(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.

**Solution 1.86.** 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 $2/(n+1)$.

**Solution 1.87.** 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

$$\left|p_n(t) - |t|\right| < \frac{\varepsilon}{a} \qquad \text{for all } t \in [-1,1].$$

Then for all $n \geqslant N$

$$\left|q_n(x) - |x|\right| = \left|a\,p_n(x/a) - a\,|x/a|\right| = a\left|p_n(x/a) - |x/a|\right| < \varepsilon \qquad \text{for all } x \in [-a,a].$$

**Solution 1.88.** Suppose $\mathcal{A}$ separates points of $X$ and is non-vanishing on $X$ and let $(x_1, y_1)$, $(x_1, y_2) \in X \times \mathbf{R}$ with $x_1 \neq x_2$. Then there exist elements $g, h_1, h_2 \in \mathcal{A}$ such that

$$g(x_1) \neq g(x_2), \qquad h_1(x_1) \neq 0, \qquad h_2(x_2) \neq 0.$$

Define $k_1, k_2 \in \mathcal{A}$ by

$$\begin{aligned} k_1(t) &= \big(g(t) - g(x_1)\big)\,h_2(t) \\ k_2(t) &= \big(g(t) - g(x_2)\big)\,h_1(t), \end{aligned}$$

then

$$k_1(x_1) = 0, \qquad k_1(x_2) \neq 0, \qquad k_2(x_1) \neq 0, \qquad k_2(x_2) = 0.$$

Finally let

$$f(t) = \frac{y_1}{k_2(x_1)}\,k_2(t) + \frac{y_2}{k_1(x_2)}\,k_1(t).$$

Conversely, let $x_1 \neq x_2$ with $x_1, x_2 \in X$ and consider the pair of points $(x_1, 0)$, $(x_2, 1) \in X \times \mathbf{R}$. Then there exists $f \in \mathcal{A}$ such that

$$f(x_1) = 0 \neq 1 = f(x_2),$$

therefore $\mathcal{A}$ separates points of $X$.

Now let $x_1 \in X$. Choose $x_2 \in X$ such that $x_2 \neq x_1$ and consider the pair of points $(x_1, 1)$, $(x_2, 0) \in X \times \mathbf{R}$, then there exists $f \in \mathcal{A}$ such that $f(x_1) = 1 \neq 0$. Therefore $\mathcal{A}$ is non-vanishing on $X$.

**Solution 1.89.**

(a) Immediate from

$$\begin{aligned} \operatorname{Re}(f) &= \frac{f + \overline{f}}{2}, \\ \operatorname{Im}(f) &= \frac{f - \overline{f}}{2i}, \end{aligned}$$

as $\mathcal{C}$ is a $\mathbf{C}$-algebra, so it is closed under taking $\mathbf{C}$-linear combinations.

(b) It suffices to prove that $\mathcal{C}_{\mathbf{R}}$ is non-vanishing and separates points, then use Stone–Weierstrass.

Let $x \in X$. Since $\mathcal{C}$ is non-vanishing, there exists $f \in \mathcal{C}$ such that $f(x) \neq 0$. Then at least one of $\operatorname{Re}(f)(x)$ and $\operatorname{Im}(f)(x)$ is non-zero, so we can conclude by part (a).

Let $x \neq y \in X$. Since $\mathcal{C}$ separates points, there exists $g \in \mathcal{C}$ such that $g(x) \neq g(y)$. If $\operatorname{Re}(g)(x) \neq \operatorname{Re}(g)(y)$, we are done. Otherwise we must have $\operatorname{Im}(g)(x) \neq \operatorname{Im}(g)(y)$, and we are done.

(c) Given a function $f$ in $C_0(X, \mathbf{C})$, $\operatorname{Re}(f)$ and $\operatorname{Im}(f)$ both belong to $C_0(X, \mathbf{R})$, so they are both in $\overline{\mathcal{C}}$.

As $\overline{\mathcal{C}}$ is a $\mathbf{C}$-vector space, we conclude that $f = \operatorname{Re}(f) + i \operatorname{Im}(f) \in \overline{\mathcal{C}}$. Hence $C_0(X, \mathbf{C}) = \overline{\mathcal{C}}$.

**Solution 1.90.** Let $\mathcal{A} = \mathbf{C}[x, y]$, that is, the algebra of complex polynomial functions in two variables $x$ and $y$ (which can be thought of as coordinate projection maps $X \longrightarrow \mathbf{R}$) mapping $X \longrightarrow \mathbf{C}$.

By Exercise 1.89, it suffices to prove $\mathcal{A}$ is closed under complex conjugation, is non-vanishing, and separates points. The first property holds as the coefficients of the polynomials can be anything in $\mathbf{C}$, which is closed under complex conjugation. Non-vanishing is immediate by taking $f(x, y) = 1 \in \mathcal{A}$.

To see $\mathcal{A}$ separates points, let $(x_1, y_1), (x_2, y_2) \in X$ be any to distinct points. If $x_1 \neq x_2$, then take $f(x, y) = x \in \mathcal{A}$ so that $f(x_1, y_1) = x_1 \neq x_2 = f(x_2, y_2)$. Otherwise if $y_1 \neq y_2$, then take $f(x, y) = y \in \mathcal{A}$ so that $f(x_1, y_1) = y_1 \neq y_2 = f(x_2, y_2)$.

**Solution 1.91.** Let $\mathcal{A} = \mathbf{C}[x, y]$, that is, the algebra of complex polynomial functions in two variables $x$ and $y$ (which can be thought of as coordinate projection maps $X \longrightarrow \mathbf{R}$) mapping $X \longrightarrow \mathbf{C}$. The proof then follows exactly that of Exercise 1.90.

Alternatively, one may take $\mathcal{A} = \mathbf{C}[q, q^{-1}]$ where $q(x, y) = \exp(i \operatorname{Arg}(x + iy))$ (i.e. complex Laurent polynomials in one variable). We use Exercise 1.89: non-vanishing is trivial, and $\mathcal{A}$ is closed under complex conjugation since $\overline{q} = q^{-1}$. Finally, $\mathcal{A}$ separates points since $q \colon \mathbf{S}^1 \longrightarrow \mathbf{C}$ is an injective map to the unit circle in $\mathbf{C}$.

## Compactness in function spaces

**Solution 1.92.**

(a) Obvious from the definitions.

(b) Given $\varepsilon > 0$, let $\delta = \varepsilon$. For any $f \in \mathcal{F}$ there exists $C_f \in [0, 1)$ such that

$$d_Y(f(x_1), f(x_2)) \leqslant C_f d_X(x_1, x_2) < d_X(x_1, x_2) < \delta = \varepsilon.$$

**Solution 1.93.** Since $Y$ is complete, so is $C_0(X, Y)$ by Propositions 2.70 and 2.71. Therefore it suffices to show that the sequence $(f_n)$ is Cauchy in $C_0(X, Y)$.

Let $\varepsilon > 0$. Since $(f_n)$ is equicontinuous, there exists $\delta > 0$ such that for all $n \in \mathbf{N}$ and all $x_1, x_2 \in X$ with $d_X(x_1, x_2) < \delta$ we have $d(f_n(x_1), f_n(x_2)) < \varepsilon/4$.

Let

$$X \subseteq \mathbf{B}_{\delta/2}(x_1) \cup \cdots \cup \mathbf{B}_{\delta/2}(x_k)$$

be a finite open cover of $X$ by open balls of radius $\delta/2$. Since $Z$ is dense in $X$, for each $i = 1, \ldots, k$ there exists $z_i \in Z \cap \mathbf{B}_{\delta/2}(x_i)$, so that $\mathbf{B}_{\delta/2}(x_i) \subseteq \mathbf{B}_\delta(z_i)$ and

$$X \subseteq \mathbf{B}_\delta(z_1) \cup \cdots \cup \mathbf{B}_\delta(z_k).$$

The sequences $(f_n(z_1)), \dots, (f_n(z_k))$ are convergent, hence Cauchy, so there exists $N \in \mathbf{N}$ such that for all $n, m \geqslant N$ we have

$$d(f_n(z_i), f_m(z_i)) < \frac{\varepsilon}{4} \qquad \text{for } i = 1, \dots, k.$$

Given $x \in X$, there exists $i = 1, \dots, k$ such that $x \in \mathbf{B}_\delta(z_i)$. For all $n, m \geqslant N$ we have

$$d(f_n(x), f_m(x)) \leqslant d(f_n(x), f_n(z_i)) + d(f_n(z_i), f_m(z_i)) + d(f_m(z_i), f_m(x)) < \frac{3\varepsilon}{4}.$$

Therefore

$$d_\infty(f_n, f_m) = \sup_{x \in X}\{d(f_n(x), f_m(x))\} \leqslant \frac{3\varepsilon}{4} < \varepsilon,$$

so the sequence $(f_n)$ is indeed Cauchy.

**Solution 1.94.** Enumerate $Z = \{z_1, z_2, \dots\}$.

The sequence $(f_n(z_1))$ is bounded in $\mathbf{R}^m$, hence has a convergent subsequence $(f_{n_k^1}(z_1))$.

The sequence $(f_{n_k^1}(z_2))$ is bounded in $\mathbf{R}^m$, hence has a convergent subsequence $(f_{n_k^2}(z_2))$.

We continue in this manner. At the $j$-th step, we get a subsequence $(f_{n_k^j})$ of $(f_{n_k^{j-1}})$ such that $(f_{n_k^j}(z_i))$ converges for $i = 1, 2, \dots, j$:

$$\begin{array}{lccccccll}
(f_n): & f_1, & f_2, & f_3, & f_4, & \dots & & \text{s.t.} & (f_n) \subseteq C_0(X, \mathbf{R}^m) \\
\hline
(f_{n_k^1}): & f_{n_1^1}, & f_{n_2^1}, & f_{n_3^1}, & f_{n_4^1}, & \dots & & \text{s.t.} & (f_{n_k^1}(z_1)) \text{ converges} \\
(f_{n_k^2}): & f_{n_1^2}, & f_{n_2^2}, & f_{n_3^2}, & f_{n_4^2}, & \dots & & \text{s.t.} & (f_{n_k^2}(z_2)) \text{ converges} \\
(f_{n_k^3}): & f_{n_1^3}, & f_{n_2^3}, & f_{n_3^3}, & f_{n_4^3}, & \dots & & \text{s.t.} & (f_{n_k^3}(z_3)) \text{ converges} \\
\vdots & & & & & & & & \vdots \\
(f_{n_k^j}): & f_{n_1^j}, & f_{n_2^j}, & f_{n_3^j}, & f_{n_4^j}, & \dots & f_{n_j^j} & \text{s.t.} & (f_{n_k^j}(z_j)) \text{ converges} \\
\vdots & & & & & & & & \vdots
\end{array}$$

We turn these nested subsequences into the subsequence desired in the statement by the diagonal argument we used in Proposition 2.65: let $f_{n_1}$ be the first term of the sequence $(f_{n_k^1})$, let $f_{n_2}$ be the second term of the sequence $(f_{n_k^2})$, etc.

Given $j \in \mathbf{N}$, $(f_{n_k}(z_j))$ converges, since after ignoring the first $j$ terms, $(f_{n_k})$ is a subsequence of $(f_{n_k^j})$. Since this holds for all $j$, we get that $(f_{n_k}(z))$ converges for every $z \in Z$.

**Solution 1.95.** Let $(f_n)$ be a sequence in $K$, then $(f_n)$ is bounded and equicontinuous. Since $X$ is totally bounded, it is separable by Tutorial Question 7.5; let $Z$ be a countable dense subset. By Exercise 1.94, $(f_n)$ has a subsequence $(f_{n_k})$ that converges at every $z \in Z$. By Exercise 1.93, $(f_{n_k})$ converges in $C_0(X, \mathbf{R}^m)$. Since $K$ is closed, $(f_{n_k})$ converges to an element of $K$.

By Theorem 2.66, $K$ is compact.

**Solution 1.96.** We know that $K$ is bounded (since every compact subset is totally bounded, hence bounded by Tutorial Question 7.1) and that $K$ is closed by Proposition 2.37.

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 1.84 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 $\varepsilon$ apart.

# 2. Normed and Hilbert spaces

## Norms and inner products

**Solution 2.1.** 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.

**Solution 2.2.** Suppose the norms are equivalent, so there exist $m, M > 0$ such that

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

Then, for any $v \in V$ and any $\varepsilon > 0$, the open balls with respect to the two norms satisfy:

$$\mathbf{B}^{(1)}_{\varepsilon/M}(v) \subseteq \mathbf{B}^{(2)}_{\varepsilon}(v) \subseteq \mathbf{B}^{(1)}_{\varepsilon/m}(v).$$

Since open balls generate the two topologies, these inclusions show that the topologies are equal.

Conversely, suppose that the two topologies are equal. The set $\mathbf{B}^{(2)}_1(0)$ is an open neighbourhood of 0, so there exists $\varepsilon > 0$ such that $\mathbf{B}^{(1)}_{\varepsilon}(0) \subseteq \mathbf{B}^{(2)}_1(0)$. In other words, if $\|v\|_1 < \varepsilon$, then $\|v\|_2 < 1$.

Let $M = 2/\varepsilon$. For any $v \in V$, $v \neq 0$, let $w = \frac{1}{M\|v\|_1}\, v$. Then

$$\|w\|_1 = \frac{1}{M\|v\|_1}\, \|v\|_1 = \frac{1}{M} = \frac{\varepsilon}{2} < \varepsilon.$$

Therefore

$$\|w\|_2 < 1 \qquad \Rightarrow \qquad \frac{1}{M\|v\|_1}\, \|v\|_2 < 1 \qquad \Rightarrow \qquad \|v\|_2 \leqslant M\|v\|_1.$$

Starting over with the roles of $\|\cdot\|_1$ and $\|\cdot\|_2$ interchanged, we obtain some $M' > 0$ such that

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

so letting $m = 1/M'$ we conclude that

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

**Solution 2.3.** (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 $\alpha S$, 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 $\alpha S$.

**Solution 2.4.** 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

$$\begin{aligned}
\big|\langle x_n, y_n\rangle - \langle x, y\rangle\big| &= \big|\langle x_n, y_n\rangle - \langle x, y_n\rangle + \langle x, y_n\rangle - \langle x, y\rangle\big| \\
&= \big|\langle x_n - x, y_n\rangle + \langle x, y_n - y\rangle\big| \\
&\leqslant \big|\langle x_n - x, y_n\rangle\big| + \big|\langle x, y_n - y\rangle\big| \\
&\leqslant \|x_n - x\|\,\|y_n\| + \|x\|\,\|y_n - y\| \\
&\leqslant C\|x_n - x\| + \|x\|\,\|y_n - y\| \\
&< \varepsilon.
\end{aligned}$$

We conclude that $\big(\langle x_n, y_n\rangle\big) \longrightarrow \langle x, y\rangle$.

**Solution 2.5.** In the proof of the Parallelogram Law (Proposition 3.10) we added the two equalities

$$\begin{aligned}\|v+w\|^2 &= \|v\|^2 + 2\operatorname{Re}\langle v,w\rangle + \|w\|^2\\ \|v-w\|^2 &= \|v\|^2 - 2\operatorname{Re}\langle v,w\rangle + \|w\|^2.\end{aligned}$$

Subtracting them instead gives us

$$4\operatorname{Re}\langle v,w\rangle = \|v+w\|^2 - \|v-w\|^2.$$

If $\mathbf{F} = \mathbf{R}$, we are done.

If $\mathbf{F} = \mathbf{C}$, note that

$$\operatorname{Im}\langle v,w\rangle = \operatorname{Re}\langle v,iw\rangle,$$

and we are done.

**Solution 2.6.**

(a) Equation (2.1) is obvious from the definition. For Equation (2.2), we have

$$\begin{aligned}4[iv,iw] &= \|iv+iw\|^2 - \|iv-iw\|^2 = \|i(v+w)\|^2 - \|i(v-w)\|^2\\ &= |i|^2\|v+w\|^2 - |i|^2\|v-w\|^2 = 4[v,w].\end{aligned}$$

For Equation (2.3):

$$\begin{aligned}4[v,iw] &= \|v+iw\|^2 - \|v-iw\|^2 = \|i(w-iv)\|^2 - \|-i(w+iv)\|^2\\ &= |i|^2\|w-iv\|^2 - |-i|^2\|w+iv\|^2 = -4[w,iv].\end{aligned}$$

(b) Equation (2.4) follows from the calculation

$$\begin{aligned}4[2u,w] + 4[2v,w] &= \|2u+w\|^2 - \|2u-w\|^2 + \|2v+w\|^2 - \|2v-w\|^2\\ &= \Big(\big\|(u+v+w)+(u-v)\big\|^2 + \big\|(u+v+w)-(u-v)\big\|^2\Big)\\ &\quad - \Big(\big\|(u+v-w)+(u-v)\big\|^2 + \big\|(u+v-w)-(u-v)\big\|^2\Big)\\ &= 2\big(\|u+v+w\|^2 + \|u-v\|^2\big) - 2\big(\|u+v-w\|^2 + \|u-v\|^2\big)\\ &= 8[u+v,w].\end{aligned}$$

In particular, setting $u = 0$ we have $[2u,w] = 0$ (from the definition of $[\cdot,\cdot]$), hence Equation (2.5).

Using this on the LHS of Equation (2.4) we get Equation (2.6).

(c) We already have Equation (2.7) for $n = 0,1,2$. Clearly repeated application of part (b) gives us

$$[nv,w] = n[v,w] \qquad \text{for all } n \in \mathbf{N}.$$

For $(-1)$ we have

$$4[-v,w] = \|-v+w\|^2 - \|-v-w\|^2 = \|v-w\|^2 - \|v+w\|^2 = -4[v,w],$$

hence

$$[nv,w] = n[v,w] \qquad \text{for all } n \in \mathbf{Z}.$$

For any $q \in \mathbf{Q}$, write $q = m/n$ with $\gcd(m,n) = 1$:

$$n[qv,w] = [nqv,w] = [mv,w] = m[v,w],$$

therefore

$$[qv, w] = q[v, w] \qquad \text{for all } q \in \mathbf{Q}.$$

Finally, for any $x \in \mathbf{R}$ choose a rational sequence $(q_n) \longrightarrow x$:

$$\begin{aligned}[xv, w] &= \left[\left(\lim_{n\longrightarrow\infty} q_n\right) v, w\right] = \left[\lim_{n\longrightarrow\infty} (q_n v), w\right] = \lim_{n\longrightarrow\infty} [q_n v, w] \\ &= \lim_{n\longrightarrow\infty} (q_n [v, w]) = \left(\lim_{n\longrightarrow\infty} q_n\right) [v, w] = x[v, w].\end{aligned}$$

Somewhere in the middle we used the fact that $[\cdot,\cdot]$ is continuous in the first variable (which follows easily from the definition of $[\cdot,\cdot]$ and the fact that the norm is continuous).

(d) Straightforward, as

$$4[v, v] = 4\|v\|^2.$$

(e) We check the inner product properties:

$$\langle w, v\rangle = [w, v] + i[w, iv] = [v, w] - i[v, iw] = \overline{\langle v, w\rangle}.$$

$$\begin{aligned}\langle u + v, w\rangle &= [u + v, w] + i[u + v, iw] \\ &= [u, w] + [v, w] + i[u, iw] + i[v, iw] \\ &= \langle u, w\rangle + \langle v, w\rangle.\end{aligned}$$

Writing $\alpha = x + iy \in \mathbf{C}$, we have

$$\begin{aligned}\langle \alpha v, w\rangle &= [\alpha v, w] + i[\alpha v, iw] \\ &= [xv + iyv, w] + i[xv + iyv, iw] \\ &= [xv, w] + [iyv, w] + i[xv, iw] + i[iyv, iw] \\ &= x[v, w] + y[iv, w] + ix[v, iw] + iy[iv, iw] \\ &= (x + iy)[v, w] + i(x + iy)[v, iw] \\ &= \alpha\langle v, w\rangle.\end{aligned}$$

Here we used in order Equations (6), (10), (2), and (3).

Finally, note that for all $v \in V$ we can apply (3) with $w = v$ and then (1) with $w = iv$ to see that

$$[v, iv] = -[v, iv] \quad \text{so} \quad [v, iv] = 0.$$

Therefore

$$\langle v, v\rangle = [v, v] + i[v, iv] = [v, v],$$

so positive-definiteness follows from part (d). Moreover, we have seen in part (d) that $[v, v] = \|v\|^2$, so $\langle v, v\rangle = [v, v] = \|v\|^2$, which implies that the norm defined by $\langle\cdot,\cdot\rangle$ is $\|\cdot\|$.

**Solution 2.7.** We know that $\mathrm{Span}(S)$ is a subspace of $V$, and by Corollary 3.5 that $\overline{\mathrm{Span}(S)}$ is a closed subspace of $V$.

Let $W \subseteq V$ be some closed subspace of $V$ that contains $S$. Then $\mathrm{Span}(S) \subseteq W$, and so $\overline{\mathrm{Span}(S)} \subseteq \overline{W} = W$, whence the minimality property.

**Solution 2.8.** Let $f\colon \mathbf{F} \longrightarrow \mathbf{F}v$ be the bijection defined by $f(\alpha) = \frac{\alpha}{\|v\|}v$. If $\alpha$ and $\beta$ are elements of $\mathbf{F}$, then

$$d_{\mathbf{F}}(\alpha, \beta) = |\alpha - \beta| = \frac{|\alpha - \beta|}{\|v\|}v = \left\|\frac{\alpha}{\|v\|}v - \frac{\beta}{\|v\|}v\right\| = d_V\big(f(\alpha), f(\beta)\big).$$

Hence $f$ is a bijective isometry.

**Solution 2.9.** If $(w_n)$ is a sequence in $W$ and it converges to $v$ in $V$, then $(w_n)$ is Cauchy by Proposition 2.45. It follows from Proposition 3.11 that $W$ is complete, so $v \in W$. Hence $W$ is closed by part (c) of Proposition 2.42.

**Solution 2.10.** The relation is reflexive because if $\|\cdot\|$ is a norm on a vector space $V$, then

$$1 \cdot \|v\| \leqslant \|v\| \leqslant 1 \cdot \|v\| \qquad \text{for all } v \in V.$$

The relation is symmetric because if $\|\cdot\|_1$ and $\|\cdot\|_2$ are norms on a vector space $V$, and if there exists positive real numbers $m$ and $M$ such that

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

then

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

The relation is transitive because if $\|\cdot\|_1$, $\|\cdot\|_2$, and $\|\cdot\|_3$ are norms on a vector space $V$, and if there exists positive real numbers $m$, $k$, $M$, and $K$ such that

$$m\|v\|_1 \leqslant \|v\|_2 \leqslant M\|v\|_1 \quad \text{and} \quad k\|v\|_2 \leqslant \|v\|_3 \leqslant K\|v\|_2 \qquad \text{for all } v \in V,$$

then

$$mk\|v\|_1 \leqslant \|v\|_2 \leqslant MK\|v\|_1 \qquad \text{for all } v \in V.$$

**Solution 2.11.** (a) Let $m$ and $M$ be positive real numbers such that

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

and let $d_1$ and $d_2$ be the metrics defined by $\|\cdot\|_1$ and $\|\cdot\|_2$ respectively (see Proposition 3.1 for the definition). If $\varepsilon$ is a positive real number, then $d_1(v,w) < \varepsilon/M$ implies

$$d_2\big(\mathrm{id}_V(v), \mathrm{id}_V(w)\big) = d_2(v,w) = \|v-w\|_2 \leqslant M\|v-w\|_1 = Md_1(v,w) < \varepsilon.$$

Hence $\mathrm{id}_V : (V, \|\cdot\|_1) \longrightarrow (V, \|\cdot\|_2)$ is uniformly continuous.

(b) By Exercise 2.10 and part (a), the equivalence between $\|\cdot\|_1$ and $\|\cdot\|_2$ implies the identity functions $\mathrm{id}_V : (V, \|\cdot\|_1) \longrightarrow (V, \|\cdot\|_2)$ and $\mathrm{id}_V : (V, \|\cdot\|_2) \longrightarrow (V, \|\cdot\|_1)$ are both uniformly continuous, so $\mathrm{id}_V : (V, \|\cdot\|_1) \longrightarrow (V, \|\cdot\|_2)$ is a uniform homeomorphism (see Exercise 1.61). It then follows from part (b) of Exercise 1.61 that $(V, \|\cdot\|_1)$ is complete if and only if $(V, \|\cdot\|_2)$ is complete.

**Solution 2.12.** If $T = 0$, then $\langle Tv, v\rangle = \langle 0, v\rangle = 0$ for every vector $v$ in $V$.

Conversely, suppose $\langle Tv, v\rangle = 0$ for every vector $v$ in $V$. If $v$ and $w$ are two vectors in $V$, then

$$\begin{aligned} 0 &= \big\langle T(v+w), v+w\big\rangle \\ &= \langle Tv, v\rangle + \langle Tw, v\rangle + \langle Tv, w\rangle + \langle Tw, w\rangle \\ &= \langle Tv, w\rangle + \langle Tw, v\rangle. \end{aligned}$$

Substituting $v$ by $iv$ gives

$$0 = \big\langle T(iv), w\big\rangle + \langle Tw, iv\rangle = i\langle Tv, w\rangle - i\langle Tw, v\rangle,$$

and it follows that

$$0 = \langle Tv, w\rangle - \langle Tw, v\rangle.$$

Hence $\langle Tv, w\rangle = \langle Tw, w\rangle = 0$. Since the inner product is non-degenerate and $w$ is an arbitrary vector in $V$, it follows that $Tv = 0$ for every vector $v$ in $V$, and therefore $T = 0$.

This statement does not hold for real vector spaces. Let $V = \mathbf{R}^2$ with the inner product defined by

$$\big\langle (x_1, y_1), (x_2, y_2)\big\rangle = x_1x_2 + y_1y_2$$

and let $T$ be the linear operator defined by $T(x, y) = (y, -x)$. For every vector $(x, y)$ in $V$, we have

$$\big\langle T(x, y), (x, y)\big\rangle = \big\langle (y, -x), (x, y)\big\rangle = 0,$$

but $T \neq 0$ because $T(1, 0) = (0, -1)$.

**Solution 2.13.** In $\mathbf{R}$, consider

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n}.$$

Taking absolute values we get the harmonic series, which does not converge.

The original series has alternating signs, and its terms in absolute value form a decreasing sequence $(1/n)$ that converges to zero, hence the series converges by the alternating series test.

## Bounded linear functions

**Solution 2.14.** Recall from Tutorial Question 9.1 that

$$\|u\|_W = \sup_{\|w\|_W=1} \big|\langle u, w\rangle_W\big| \qquad \text{for all } u \in W.$$

Setting $u = f(v)$ for some $v \in V$, we get

$$\|f(v)\|_W = \sup_{\|w\|_W=1} \big|\langle f(v), w\rangle_W\big| \qquad \text{for all } v \in V.$$

Therefore

$$\|f\| = \sup_{\|v\|_V=1} \|f(v)\|_W = \sup_{\|v\|_V=\|w\|_W=1} \big|\langle f(v), w\rangle_W\big|.$$

**Solution 2.15.** 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 $S = \max\big\{a^2, b^2\big\}$ and so $\|C\| = \sqrt{S} = \max\big\{|a|, |b|\big\}$.

**Solution 2.16.** 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 $w \in \operatorname{im}(f)$.

**Solution 2.17.** Let $v_1, \ldots, v_m$ be a basis of $V$ and let $w_1, \ldots, w_n$ be a basis of $W$. By Theorem 3.7, the norms on $V$ and on $W$ are equivalent to the norms $\|\cdot\|_V$ and $\|\cdot\|_W$ defined by

$$\|\alpha_1 v_1 + \cdots + \alpha_m v_m\|_V = |\alpha_1| + \cdots + |\alpha_m| \quad \text{and} \quad \|\beta_1 w_1 + \cdots + \beta_m w_n\|_V = |\beta_1| + \cdots + |\beta_n|.$$

Since continuity of $f$ is determined by the topologies on $V$ and $W$, and since equivalent norms give rise to the same topology (see Exercise 2.2), we can assume without loss of generality that $V$ and $W$ are equipped with the norms $\|\cdot\|_V$ and $\|\cdot\|_W$ respectively.

Put $M = \max\{\|f(v_1)\|_W, \ldots, \|f(v_n)\|_W\}$. If $v = \alpha_1 v_1 + \cdots \alpha_m v_m$, then

$$\begin{aligned}
\|f(v)\|_W &= \|f(\alpha_1 v_1 + \cdots + \alpha_m v_m)\|_W \\
&= \|\alpha_1 f(v_1) + \cdots + \alpha_m f(v_m)\|_W \\
&\leqslant \|\alpha_1 f(v_1)\|_W + \cdots + \|\alpha_m f(v_m)\|_W \\
&= |\alpha_1|\|f(v_1)\|_W + \cdots + |\alpha_m|\|f(v_m)\|_W \\
&\leqslant \alpha_1 M + \cdots + \alpha_m M \leqslant M\|v\|_V.
\end{aligned}$$

Hence $f$ is bounded, and therefore continuous by Proposition 3.16.

**Solution 2.18.** We start with proving that $f$ is linear. If $v_1$ and $v_2$ are vectors in $V$, then

$$\begin{aligned}
f(v_1 + v_2) &= \big(f_1(v_1 + v_2), f_2(v_1 + v_2)\big) \\
&= \big(f_1(v_1) + f_1(v_2), f_2(v_1) + f_2(v_2)\big) \\
&= \big(f_1(v_1), f_2(v_1)\big) + \big(f_1(v_2), f_2(v_2)\big) = f(v_1) + f(v_2).
\end{aligned}$$

If $\alpha$ is a scalar and $v$ is a vector in $V$, then

$$f(\alpha v) = \big(f_1(\alpha v), f_2(\alpha v)\big) = \big(\alpha f_1(v), \alpha f_2(v)\big) = \alpha f(v).$$

Hence $f$ is linear.

It remains to prove that $f$ is continuous. Let $\pi_1 \colon W_1 \times W_2 \longrightarrow V_1$ and $\pi_2 \colon W_1 \times W_2 \longrightarrow W_2$ be the projections. Since $\pi_1 \circ f = f_1$ and $\pi_2 \circ f = f_2$, it follows from Tutorial Question 3.8 that $f$ is continuous.

**Solution 2.19.** Let
$$x_m = \sum_{n=1}^{m} \alpha_n v_n, \qquad x = \sum_{n=1}^{\infty} \alpha_n v_n.$$
We know that $(x_m) \longrightarrow x$ in $V$.

Since $f \in L(V, W)$ is continuous, we have that $\big(f(x_m)\big) \longrightarrow f(x)$ in $W$.

But $f$ is also linear, so
$$f(x_m) = \sum_{n=1}^{m} \alpha_n f(v_n).$$
Hence
$$\left(\sum_{n=1}^{m} \alpha_n f(v_n)\right) \longrightarrow f(x),$$
so that the series
$$\sum_{n=1}^{\infty} \alpha_n f(v_n) \quad \text{converges to } f(x).$$

# Convexity

**Solution 2.20.** Let $I \subseteq \mathbf{R}$ be an interval and let $v, w \in I$, $a, b \in \mathbf{R}_{\geqslant 0}$ such that $a + b = 1$.

Without loss of generality, $v \leqslant w$. Then
$$av + bw - v = (a-1)v + bw = b(w - v) \geqslant 0 \Rightarrow v \leqslant av + bw$$
and
$$av + bw - w = av + (b-1)w = a(v - w) \leqslant 0 \Rightarrow av + bw \leqslant w.$$
Therefore $v \leqslant av + bw \leqslant w$, hence $av + bw \in I$ by the definition of an interval.

**Solution 2.21.** Suppose $v, w \in S$ and $a, b \in \mathbf{R}_{\geqslant 0}$ such that $a + b = 1$. Then
$$f(av + bw) = \|av + bw\| \leqslant \|av\| + \|bw\| = |a|\,\|v\| + |b|\,\|w\| = a\|v\| + b\|w\| = af(v) + bf(w).$$

**Solution 2.22.** 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 $I^\circ$, therefore $f''(x) \geqslant 0$ on $I^\circ$.

**Solution 2.23.**

(a) We have $\exp''(x) = e^x \geqslant 0$ for all $x \in \mathbf{R}$, now use Exercise 2.22.

(b) If $x = 0$ or $y = 0$, the inequality is trivial, so we may assume $x, y > 0$. Setting $x = e^s$, $y = e^t$, we are trying to prove that

$$e^{as+bt} \leqslant ae^s + be^t,$$

which is the same as $e^x$ being a convex function.

**Solution 2.24.** If $y = 0$, the inequality is obvious, so we may assume $y > 0$. Setting $t = x/y$, we are trying to show that

$$t^p + 1 \leqslant (t+1)^p \qquad \text{for all } t \geqslant 0.$$

Let $f\colon \mathbf{R}_{\geqslant 0} \longrightarrow \mathbf{R}$ be given by $f(t) = t^p + 1$, and $g(t)\colon \mathbf{R}_{\geqslant 0} \longrightarrow \mathbf{R}$ be given by $g(t) = (t+1)^p$. We have $f(0) = g(0) = 1$. Also

$$f'(t) = pt^{p-1} \leqslant p(t+1)^{p-1} = g'(t) \qquad \text{for all } t > 0,$$

therefore $f(t) \leqslant g(t)$ for all $t \geqslant 0$, as desired. (There's an appeal to the Mean Value Theorem hiding in here, if you want to write out all the details.)

**Solution 2.25.** Without loss of generality $x \leqslant y$ so $\min\{x, y\} = x$ and $\max\{x, y\} = y$.

(a) $x \leqslant y$ so $x^{-1} \geqslant y^{-1}$ so $x^{-q} \geqslant y^{-q}$ so $bx^{-q} \geqslant by^{-q}$ so $ax^{-q} + bx^{-q} \geqslant ax^{-q} + by^{-q}$ so

$$\min\{x, y\} = x = \left(ax^{-q} + bx^{-q}\right)^{-1/q} \leqslant \left(ax^{-q} + by^{-q}\right)^{-1/q}.$$

(b) Let $X = x^{-q}$, $Y = y^{-q}$, then by Exercise 2.23 we have

$$\begin{aligned} X^a Y^b \leqslant aX + bY &\Rightarrow x^{-aq} y^{-bq} \leqslant ax^{-q} + by^{-q} \\ &\Rightarrow x^{aq} y^{bq} \geqslant \left(ax^{-q} + by^{-q}\right)^{-1} \\ &\Rightarrow \left(ax^{-q} + by^{-q}\right)^{-1/q} \leqslant x^a y^b. \end{aligned}$$

(c) Similar to (b), use Exercise 2.23 with $X = x^{1/p}$, $Y = y^{1/p}$.

(d) Use Lemma 3.22 with $X = x^{1/p}$, $Y = y^{1/p}$.

(e) Precisely Lemma 3.22.

(f) Similar to (a).

**Solution 2.26.** (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.

**Solution 2.27.**

(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 $\alpha$ 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 $\alpha$ 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 $au + bv \in f^{-1}(S)$.

# Sequence spaces

**Solution 2.28.** Consider $f\colon \mathbf{F}^n \longrightarrow \ell^2$ given by

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

We have

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

so $f(a) \in \ell^2$, and $f$ is an isometry.

Linearity is straightforward.

**Solution 2.29.**

(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 $v_n = \big(1, 1/2, \dots, 1/n, 0, 0, \dots\big) \in \mathbf{F}^{\mathbf{N}}$. Since $v_n$ has only finitely many nonzero terms, it is in $\ell^1$. 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

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

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 $\ell^1$.

**Solution 2.30.** We will verify that none of the norms in the question satisfy the Parallelogram Law (Proposition 3.10), so they cannot be defined by inner products. Consider $e_1 = (1, 0, 0, 0, 0, \dots)$ and $e_2 = (0, 1, 0, 0, 0, \dots)$. Then

$$\|e_1 + e_2\|_{\ell^\infty}^2 + \|e_1 - e_2\|_{\ell^\infty}^2 = 2 \neq 4 = 2\big(\|e_1\|_{\ell^\infty}^2 + \|e_2\|_{\ell^\infty}^2\big)$$

and

$$\|e_1 + e_2\|_{\ell^p}^2 + \|e_1 - e_2\|_{\ell^p}^2 = 2 \times 2^{2/p} \neq 4 = 2\big(\|e_1\|_{\ell^p}^2 + \|e_2\|_{\ell^p}^2\big).$$

**Solution 2.31.** Let $(x^{(n)})$ be a Cauchy sequence in $\ell^\infty$. 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 $\ell^\infty$.

As $(x^{(n)})$ is a Cauchy sequence in $\ell^\infty$, it is bounded in $\ell^\infty$, 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.

**Solution 2.32.**

(a) It's pretty clear that $c_0$ is a subspace of $\mathbf{F}^{\mathbf{N}}$, and hence of $\ell^\infty$. To show that $c_0$ is closed in $\ell^\infty$, let $(x_n) \longrightarrow x \in \ell^\infty$ with $x_n \in c_0$ for all $n \in \mathbf{N}$. We want to prove that $x \in c_0$.

Write $x_n = \big(a_{nm}\big) = \big(a_{n1}, a_{n2}, a_{n3}, \dots\big)$ and $x = (a_m) = (a_1, a_2, a_3, \dots)$. Let $\varepsilon > 0$. There exists $N \in \mathbf{N}$ such that for all $n \geqslant N$ we have

$$\sup_m |a_m - a_{nm}| = \|x - x_n\|_{\ell^\infty} < \frac{\varepsilon}{2}.$$

Consider the sequence $x_N = (a_{Nm}) \in c_0$. It converges to 0, so that there exists $M \in \mathbf{N}$ such that for any $m \geqslant M$ we have

$$|a_{Nm}| < \frac{\varepsilon}{2}.$$

Therefore, for $m \geqslant M$, we get

$$|a_m| = |a_m - a_{Nm} + a_{Nm}| \leqslant |a_m - a_{Nm}| + |a_{Nm}| < \varepsilon.$$

Hence $x = (a_m) \longrightarrow 0$.

(b) Since $c_0$ is closed and $\ell^\infty$ is Banach, $c_0$ is Banach.

(c) I claim that $c_0$ has the same Schauder basis at the one given in Proposition 3.26 for $\ell^p$: $\{e_1, e_2, \dots\}$ where $e_n = (0, \dots, 0, 1, 0 \dots)$ with the 1 in the $n$-th spot.

Take $v = (v_n) \in c_0$, then $(v_n) \longrightarrow 0$. I claim that the series

$$\sum_{n=1}^{\infty} v_n e_n$$

converges to $v$ with respect to the norm on $c_0$, which is the $\ell^\infty$-norm:

$$\left\| v - \sum_{n=1}^{m} v_n e_n \right\|_{\ell^\infty} = \|(0, \dots, 0, v_{m+1}, v_{m+2}, v_{m+3}, \dots)\|_{\ell^\infty} = \sup_{n \geqslant m+1} |v_n|,$$

and the latter converges to 0 as $m \longrightarrow \infty$, since $(v_n) \longrightarrow 0$. The uniqueness of the coefficients follows in precisely the same way as for Proposition 3.26.

**Solution 2.33.**

(a) Suppose $S$ is countable and enumerate its elements:

$$\begin{aligned} a_1 &= \big(a_{11}, a_{12}, a_{13}, \dots\big) \\ a_2 &= \big(a_{21}, a_{22}, a_{23}, \dots\big) \\ a_3 &= \big(a_{31}, a_{32}, a_{33}, \dots\big) \\ &\vdots \end{aligned}$$

Go down the diagonal of this infinite grid of 0's and 1's, and define $b_n = 1 - a_{nn}$ for all $n \in \mathbf{N}$. Then $b = (b_n) \in S$, but $b \neq a_m$ for any $m \in \mathbf{N}$, contradiction.

(b) If $a = (a_n), b = (b_n) \in S$ with $a \neq b$ then

$$\|a - b\| = \sup_n |a_n - b_n| = 1,$$

so $\mathbf{B}_{1/2}(a) \cap \mathbf{B}_{1/2}(b) = \varnothing$.

Therefore we can take

$$T = \big\{\mathbf{B}_{1/2}(s) : s \in S\big\}.$$

(c) Any dense subset $D$ of $\ell^\infty$ must contain at least one point (in fact, must be dense) in each open ball in the set $T$. Since $T$ is uncountable, $D$ must also be uncountable, so $\ell^\infty$ is not separable.

**Solution 2.34.** (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

$$\left| \lim_{n \longrightarrow \infty} a_n \right| \leqslant \sup_{n \in \mathbf{N}} |a_n| = \|a\|_{\ell^\infty},$$

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 Exercise 2.32 that $c_0$ is closed in $\ell^\infty$, so $c$ must also be closed in $\ell^\infty$ as it is homeomorphic to $c_0$. But $\ell^\infty$ 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 $v$.

**Solution 2.35.**

(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 $H_{\text{odd}}$.

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 $H_{\text{odd}}$.

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 $H_{\text{odd}}$.

(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 an isometry.

To prove surjectivity of $f$, we show that the restriction of the function $g$ from part (b) maps to $\ell^1$: 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):

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

which shows that $f$ is continuous.

It also indicates that $f$ is not an isometry: 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 $\ell^\infty$: 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.

## Dual spaces

**Solution 2.36.** Suppose $\beta$ is continuous at $(0,0)$ but not bounded. 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 $\beta$.

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 $\beta$ is bounded; 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

$$\begin{aligned}
\|\beta(u', v') - \beta(u, v)\|_W &= \|\beta(u', v') - \beta(u', v) + \beta(u', v) - \beta(u, v)\|_W \\
&= \|\beta(u', v' - v) + \beta(u' - u, v)\|_W \\
&\leqslant \|\beta(u', v' - v)\|_W + \|\beta(u' - u, v)\|_W \\
&\leqslant c\,\|u'\|_U\,\|v' - v\|_V + c\,\|u' - u\|_U\,\|v\|_V \\
&\leqslant c\left(\|u\|_U + \|u' - u\|_U\right)\|v' - v\|_V + c\,\|u' - u\|_U\,\|v\|_V \\
&< c\,(\|u\|_U + 1)\delta + c\delta\|v\|_V \\
&\leqslant c(\|u\|_U + 1)\frac{\varepsilon}{2c(\|u\|_U + 1)} + c\|v\|_V\frac{\varepsilon}{2c(\|v\|_V + 1)} \\
&< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \\
&= \varepsilon.
\end{aligned}$$

Therefore, $\mathbf{B}_\delta(u, v) \subseteq \beta^{-1}(\mathbf{B}_\varepsilon(\beta(u, v)))$ and thus $\beta$ is continuous.

Obviously, if $\beta$ is continuous on $U \times V$ then it is continuous at $(0, 0)$, closing the cycle of equivalences.

**Solution 2.37.** Since $U, V, W$ are nonzero and $\beta$ 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

$$\left\|(0, bv) - (au, bv)\right\|_{U\times V} = \left\|(-au, 0)\right\|_{U\times V} = a\|u\|_U = \frac{\delta}{2} < \delta,$$

but

$$\left\|\beta(0, bv) - \beta(au, bv)\right\|_W = \left\|\beta(-au, bv)\right\|_W = ab\left\|\beta(u, v)\right\|_W = \frac{3}{2} > 1 = \varepsilon.$$

Therefore, $\beta$ is not uniformly continuous.

(In fact, the proof shows that $\beta$ is not even uniformly continuous on the subspace $\mathbf{F}u \times \mathbf{F}v \subseteq U \times V$.)

**Solution 2.38.** Since $\beta$ is continuous and bilinear, it is bounded: 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 $\beta$, 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 $\beta$ means that $\beta(0,v) = 0$ and $\beta(u,0) = 0$ for all $u \in U$, $v \in V$, so we conclude that $\beta = 0$.

**Solution 2.39.**

(a) **First approach (direct):** Let $v \in V$. We prove that $f_u : V \longrightarrow W$ is continuous at $v$. (Note that, crucially, $u$ remains fixed.)

Let $\varepsilon > 0$; as $\beta$ 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 boundedness):** Let $\varepsilon > 0$; as $\beta$ is continuous, it is bounded, 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 bounded and thus continuous.

(b) Let $\varepsilon > 0$; as $\beta$ is continuous, it is bounded, 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, $\beta_U$ is bounded and thus continuous.

**Solution 2.40.** By Exercise 2.36, $\beta_U \colon \ell^\infty \longrightarrow \left(\ell^1\right)^\vee$ 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 $\ell^1$ 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 bounded, 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 $\ell^\infty$ is isometric to $\left(\ell^1\right)^\vee$ and all dual spaces as Banach, $\ell^\infty$ is Banach.

**Solution 2.41.**

(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 2.36, $\beta_U$ is linear and continuous. In our notation, this is the function $u \longmapsto u^\vee \colon \ell^1 \longrightarrow \left(c_0\right)^\vee$, where

$$u^\vee(v) = \beta(u,v) = \sum_{n=1}^{\infty} u_n v_n.$$

For surjectivity, we need to show that each $\varphi \in \left(c_0\right)^\vee$ is of the form $\varphi = u^\vee$ for some $u \in \ell^1$. Take such $\varphi$. Recall that $c_0$ has Schauder basis $\{e_1, e_2, \ldots\}$, 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 an isometry.

Putting it all together, we have a linear bijective isometry $\ell^1 \longrightarrow \left(c_0\right)^\vee$.

(b) We know that duals of normed spaces are complete, so $\left(c_0\right)^\vee$ is complete, so $\ell^1$, 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 isometry property.

# Orthogonality

**Solution 2.42.**

(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 $s \in \left(S^\perp\right)^\perp$.

(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

$$\langle x, w\rangle = \langle x, \lim v_n\rangle = \lim\langle x, v_n\rangle = \lim 0 = 0.$$

**Solution 2.43.** We know that the composition of continuous linear maps is continuous linear, so this is true for $\varphi \circ \psi$. To conclude that it is a projection, we need to compute its square:

$$\big(\varphi \circ \psi\big) \circ \big(\varphi \circ \psi\big) = \big(\varphi \circ \varphi\big) \circ \big(\psi \circ \psi\big) = \varphi \circ \psi,$$

where it was crucial that $\varphi$ and $\psi$ commute.

For the statement about the image, note that $w \in \operatorname{im}\big(\varphi \circ \psi\big)$ if and only if there exists $v \in V$ such that

$$w = \varphi\big(\psi(v)\big) = \psi\big(\varphi(v)\big),$$

which implies that $w \in \operatorname{im}\varphi \cap \operatorname{im}\psi$. So $\operatorname{im}\big(\varphi \circ \psi\big) \subseteq \operatorname{im}\varphi \cap \operatorname{im}\psi$.

Conversely, suppose $w \in \operatorname{im}\varphi \cap \operatorname{im}\psi$, then there exists $v \in V$ such that $w = \psi(v)$. But $w \in \operatorname{im}\varphi$ and $\varphi$ is a projection, so that

$$w = \varphi(w) = \varphi\big(\psi(v)\big) \in \operatorname{im}\big(\varphi \circ \psi\big).$$

**Solution 2.44.**

(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 $\ell^{\infty}$-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

$$\begin{aligned} d(y,x) &= \|y - x\| = \left\|az_1 + (1-a)z_2 - x\right\| = \left\|az_1 - ax + (1-a)z_2 - (1-a)x\right\| \\ &= \left\|a(z_1 - x) + (1-a)(z_2 - x)\right\| \leqslant \left\|a(z_1 - x)\right\| + \left\|(1-a)(z_2 - x)\right\| \\ &= a\|z_1 - x\| + (1-a)\|z_2 - x\| = aD + (1-a)D = D. \end{aligned}$$

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 $y \in Z$.

# A. Appendix: Prerequisites

## Equivalence relations

**Solution A.1.**

- Given $x \in A$, we have $f(x) = f(x)$ so $x \sim x$.
- If $x \sim y$, then $f(x) = f(y)$, so $f(y) = f(x)$, that is $y \sim x$.
- If $x \sim y$ and $y \sim z$ then $f(x) = f(y)$ and $f(y) = f(z)$, so that $f(x) = f(z)$, that is $x \sim z$.

**Solution A.2.** Suppose $\pi$ is bijective. I claim that the only way $x \sim y$ can happen is if $x = y$: if $x \sim y$ then $\pi(x) = \pi(y)$, but $\pi$ is bijective so $x = y$.

We conclude that the equivalence relation on $A$ must be given by: $x \sim y$ if and only if $x = y$.

**Solution A.3.**

(a) We check the equivalence relation conditions:

- Given $(a,b) \in \mathbf{N} \times \mathbf{N}$, we have $a + b = b + a$ so $(a,b) \sim (a,b)$.
- If $(a,b) \sim (c,d)$ then $a + d = b + c$, so $c + b = d + a$, that is $(c,d) \sim (a,b)$.
- If $(a,b) \sim (c,d)$ and $(c,d) \sim (x,y)$ then $a + d = b + c$ and $c + y = d + x$. Adding these two equalities gives $a + d + c + y = b + c + d + x$, and cancelling out $c + d$ on both sides we get $a + y = b + x$, that is $(a,b) \sim (x,y)$.

(b) Define $g\colon (A/\sim) \longrightarrow \mathbf{Z}$ by $g\big([(a,b)]\big) = b - a$. We first need to make sure that this is a well-defined function, in other words that the value does not depend on the chosen representative $(a,b)$ of $[(a,b)]$: suppose $(a',b') \in [(a,b)]$, then $(a',b') \sim (a,b)$ so $a' + b = b' + a$, hence $a' - b' = a - b$.

Let's show that $g$ is injective: if $g([(a,b)]) = g([(c,d)])$ then $a - b = c - d$, so $a + d = b + c$, so $(a,b) \sim (c,d)$, so $[(a,b)] = [(c,d)]$.

Finally, to see that $g$ is surjective, let $n \in \mathbf{Z}$. If $n \geqslant 0$ then $n = g([(n+1,1)])$; if $n < 0$ then $n = g([(1,1-n)])$.

**Solution A.4.**

(a) Let $f\colon V \longrightarrow V$. Clearly $\mathrm{id}_V$ is unipotent and $f = \mathrm{id}_V \circ f$, so $f \sim f$.

(b) Suppose $f \sim g$ so that $f = u \circ g$, where $(u - \mathrm{id}_V)^k = 0$. Pick $m \in \mathbf{Z}_{\geqslant 1}$ such that $p^m > k$, and observe that

$$0 = (u - \mathrm{id}_V)^{p^m} = u^{p^m} - \mathrm{id}_V$$

as $\mathrm{End}(V)$ has characteristic $p$. Thus $u^{p^m-1} \circ f = g$, and $u^{p^m-1}$ is unipotent as

$$(u^{p^m-1} - \mathrm{id}_V)^{p^m} = u^{p^m(p^m-1)} - \mathrm{id}_V = 0.$$

(c) Define $f, g \in \mathrm{End}(V)$ by

$$f(s)(1) = s(1) + s(2), \quad f(s)(j) = s(j)\ \forall\ j \neq 1$$
$$g(s)(2) = s(1) + s(2), \quad g(s)(j) = s(j)\ \forall\ j \neq 2.$$

We have $(f - \mathrm{id}_V)^2 = (g - \mathrm{id}_V)^2 = 0$, so $f$ and $g$ are unipotent and thus $f \sim \mathrm{id}_V$ and $g \sim \mathrm{id}_V$. But $g$ is invertible, and

$$(f \circ g^{-1} - \mathrm{id}_V)^3(s)(j) = \begin{cases} s(j) & \text{if } j = 1, 2 \\ 0 & \text{otherwise,} \end{cases}$$

so $(f \circ g^{-1} - \mathrm{id}_V)^{3m} \neq 0$ for all $m \geqslant 1$, and thus $f \circ g^{-1}$ cannot be unipotent, meaning $f \not\sim g$.

**Solution A.5.** For part (a), reflexiveness follows as $v - v = 0 \in W$, symmetry follows as $v - v' \in W$ implies $-1 \times (v - v') = v' - v \in W$, and transitivity follows as $v - v', v' - v'' \in W$ imply $v - v' + v' - v'' = v - v'' \in W$.

For part (b), if $[v] = [u]$ (and hence $v - u \in W$), then $[v] + [v'] = [v + v'] = [u + v'] = [u] + [v']$ where the middle equality follows as $v + v' - (u + v') = v - u \in W$. This shows addition is well-defined. Similarly, $\lambda[v] = [\lambda v] = [\lambda u] = \lambda[u]$ where the middle equality follows as $\lambda v - \lambda u = \lambda(v - u) \in W$. This shows scalar multiplication is well-defined.

For part (c), define $g([v]) \coloneqq f(v)$, which clearly satisfies $f = g \circ \pi$. To show this is well-defined, suppose $[v] = [v']$ so that $v' - v \in W$. Then

$$g([v]) = f(v) = f(v) + 0 = f(v) + f(v' - v) = f(v + v' - v) = f(v') = g([v']).$$

Also $g$ is linear as $g([v + v']) = f(v + v') = f(v) + f(v') = g([v]) + g([v'])$ and $g(\lambda[v]) = g([\lambda v]) = f(\lambda v) = \lambda f(v) = \lambda g([v])$. Finally to show it is unique, suppose $g_1, g_2 \colon V/W \longrightarrow U$ both satisfy $f = g_1 \circ \pi = g_2 \circ \pi$. Then subtracting these equations gives $0 = (g_1 - g_2) \circ \pi$, which implies $g_1 = g_2 = 0$ as $\pi$ is surjective.

## (Un)countability

**Solution A.6.**

- Given $S \in X$, the identity function $\mathrm{id}_S \colon S \longrightarrow S$ is bijective, so $S \sim S$.
- If $S \sim T$ then there is a bijective function $f \colon S \longrightarrow T$, so there's a bijective inverse function $f^{-1} \colon T \longrightarrow S$, that is $T \sim S$.
- If $S \sim T$ and $T \sim W$, then there are bijective functions $f \colon S \longrightarrow T$ and $g \colon T \longrightarrow W$. The composition $g \circ f \colon S \longrightarrow W$ is bijective, so $S \sim W$.

**Solution A.7.** 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 $\mathrm{id}_Y$: 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 by Proposition A.6 that $Y$ is finite or countable.

**Solution A.8.** 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 $\varphi_n : \mathbf{N} \longrightarrow S_n$. (As Chengjing rightfully points out to me, this uses the Axiom of Countable Choice.) Define a function $\psi : \mathbf{N} \times \mathbf{N} \longrightarrow S$ 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.

**Solution A.9.** Since $B$ is countable we can enumerate it as $B = \{b_n : n \in \mathbf{N}\}$. For each $n \in \mathbf{N}$, let $W_n = \operatorname{Span}\{b_1, \dots, b_n\}$. 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 $w \in W = \operatorname{Span}(B)$, 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

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

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 A.8.

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

## Linear algebra

**Solution A.10.** TODO

**Solution A.11.** TODO

**Solution A.12.** TODO

**Solution A.13.** We can write $f(\mathbf{x}) = \mathbf{x}^\mathsf{T} A \mathbf{x}$ where $A = (a_{ij})$ is a real matrix. Observe that we can replace $A$ by $A' := \frac{1}{2}(A + \overline{A}^\mathsf{T})$ and the equation holds true. Since $A'$ is a Hermitian matrix, by the Spectral theorem we can write $A' = P^{-1} B P$ where $P$ is a (unitary) matrix whose columns are orthonormal eigenvectors of $A$, and $B = \operatorname{diag}(b_1, \dots, b_n)$ is a diagonal matrix with the corresponding eigenvalues. If $\mathbf{r}_i$ denotes the $i^{\text{th}}$ row of $P^{-1}$, then setting $g_i(\mathbf{x}) := \mathbf{r}_i \mathbf{x}$ gives the desired result.

*Comment.* The result still holds true if we allow $a_{ij} \in \mathbf{C}$, but the above proof does not apply; research Takagi factorisation.

## Uniform continuity and uniform convergence

**Solution A.14.** (a) Take (for example) $f_n(x) = e^{x-n}$, which converges pointwise to $f(x) = 0$.

(b) Suppose for the sake of contradiction that $f$ is uniformly continuous. Let $\varepsilon > 0$ be given. By uniform convergence, there exists $N > 0$ such that $|f_N(x) - f(x)| < \varepsilon/3$ for all $x \in \mathbf{R}$. Also by the uniform continuity of $f$, there exists $\delta > 0$ such that $|f(x) - f(x')| < \varepsilon/3$ whenever $|x - x'| < \delta$. Then for all $x, x' \in \mathbf{R}$ with $|x - x'| < \delta$, we have

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

which contradicts the fact that $f_N$ is not uniformly continuous.

**Solution A.15.** (a) Take for example $f_n(x) = e^{-x^{2n}}$, which converges pointwise to $f(x) = \mathbf{1}_{\{0\}}(x)$.

(b) Let $\varepsilon > 0$ be given. By uniform convergence, there exists some $f_n$ such that $|f_n(x) - f(x)| < \varepsilon/3$ for all $x \in \mathbf{R}$. By the uniform continuity of $f_n$, there exists some $\delta > 0$ such that $|f_n(x) - f_n(x')| < \varepsilon/3$ whenever $|x - x'| < \delta$. Then by the triangle inequality, for all $x, x' \in \mathbf{R}$ satisfying $|x - x'| < \delta$, we have

$$\begin{aligned} |f(x) - f(x')| &\leqslant |f(x) - f_n(x)| + |f_n(x) - f_n(x')| + |f_n(x') - f(x')| \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 \\ &= \varepsilon. \end{aligned}$$

# B. Appendix: Miscellaneous

## Zorn's Lemma

**Solution B.1.** The fact that $\subseteq$ is a partial order follows directly from known properties of set inclusion.

If $\Omega$ 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.

**Solution B.2.** 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.

**Solution B.3.** 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 $n \in \mathbf{N}$, $a_1, \dots, a_n \in \mathbf{F}$, and $u_1, \dots, u_n \in U$ such that

$$a_1 u_1 + \cdots + a_n u_n = 0. \tag{B.1}$$

Let $J = \{1, \dots, n\}$. 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 $u_1, \dots, u_n \in S_i$, so that the linear relation of Equation (B.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 $B$.

**Solution B.4.** (a) Clearly $(A, s_A) \leqslant (A, s_A)$. Now if $(A, s_A) \leqslant (B, s_B)$ and $(B, s_B) \leqslant (A, s_A)$, then $A \subseteq B \subseteq A \implies A = B$, and thus $s_A\big|_B = s_A = s_B = s_B\big|_A$. For the last condition, if $(A, s_A) \leqslant (B, s_B)$ and $(B, s_B) \leqslant (C, s_C)$, then clearly $A \subseteq C$, and $s_C\big|_A = s_C\big|_B\big|_C = s_B\big|_A = s_A$.

(b) Let $\mathcal{C} = \{(A_i, s_{A_i})\}_{i\in I}$ be a nonempty chain in $P(f)$. Define $A := \bigcup_{i\in I} A_i$, and $s_A(y) = s_{A_i}(y)$ if $y \in A_i$. This is well-defined as if $y \in A_i \cap A_j$, then without loss of generality $A_i \leqslant A_j$, and so $s_{A_i}(y) = s_{A_j}\big|_{A_i}(y) = s_{A_j}(y)$. Observe that $A_i \subseteq A$ and $s_A\big|_{A_i} = s_{A_i}$ for all $i \in I$, so we have constructed the desired upper bound.

(c) We deduce from the previous part and Zorn's lemma that there exists a maximal element $(M, s_M) \in P(f)$. Suppose that $M \neq Y$; then there exists $y_0 \in Y \setminus M$. By the surjectivity of $f$, there exists $x_0 \in X$ such that $f(x_0) = y_0$. Then we can define $M' = M \cup \{y_0\}$ and $s_{M'}$ by $s_{M'}\big|_M = s_M$ and $s_{M'}(y_0) = x_0$ so that $f \circ s_{M'} = \mathrm{id}_{M'}$. But this contradicts the maximality of $(M, s_M)$, so $M = Y$ and we obtain the desired map $s = s_M$.

## Linear algebra

**Solution B.5.** Let $S = \{e_1, e_2, \dots\}$ and $W = \operatorname{Span}(S)$.

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

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

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

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

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 $\mathbf{R}^{\mathbf{N}}$.

**Solution B.6.** This is a straightforward rewriting of the definition of algebraic: $\alpha$ 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 $\alpha$, which exists if and only if $T$ is linearly dependent.

**Solution B.7.** 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

$$\mathrm{ev}_\alpha(\lambda f) = (\lambda f)(\alpha) = \lambda f(\alpha) = \lambda\, \mathrm{ev}_\alpha(f).$$

**Solution B.8.** As in Proposition B.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 $M$.

**Solution B.9.**

(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 $\Gamma$ is injective. Let $W = \mathrm{Span}\{v_1, \ldots, v_n\}$. 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

$$v_m^\vee(a_1w_1 + \cdots + a_mw_m) = a_m.$$

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 $\Gamma$.

We conclude that $W = V$, in other words $\{v_1, \dots, v_n\}$ spans $V$.

**Conversely,** suppose $\{v_1, \dots, v_n\}$ 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 $\{v_1, \dots, v_n\}$ 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 $\Gamma$ 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 $i \in \{1, \dots, n\}$. Since $\Gamma$ 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 $\varphi_i$ on both sides of the linear relation, we get

$$a_i = 0.$$

Since this holds for all $i$, the relation is trivial.

**Conversely,** suppose $\{v_1, \dots, v_n\}$ is linearly independent. This set can be enlarged to a basis $B = \{v_1, \dots, v_n, v_{n+1}, \dots, v_m\}$ of $V$, with dual basis $v_1^\vee, \dots, v_m^\vee$.

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 $\Gamma$ is surjective.

**Solution B.10.** (a) Suppose $T^\vee(\ell) = 0$, that is $\ell \circ T$ is the zero map. But since $T$ is surjective, this implies $\ell = 0 \in W^\vee$.

(b) Let $A$ be the matrix representation of $T$ with respect to some basis $B = (b_1, \dots, b_n)$; recall that $A^\mathsf{T}$ is the matrix representation of $T^\vee$ with respect to the basis $B^\vee = (b_1^\vee, \dots, b_n^\vee)$. Since $T$ is injective, $\mathrm{rank}(A) = n = \dim(V)$. Then $\mathrm{rank}(A^\mathsf{T}) = n = \dim(V^\vee)$, so $A^\mathsf{T}$ has full-rank and thus $T^\vee$ is surjective.

(c) Let $V$ be the vector space of finitely supported real sequences, that is

$$V = \left\{(x_1, x_2, \dots) \in \mathbf{R}^\mathbf{N} : \text{finitely many } x_i \neq 0\right\},$$

and let $W = \mathbf{R}^\mathbf{N}$ be the space of all real sequences. Clearly $V \hookrightarrow W$ is injective. But the induced map $W^\vee \longrightarrow V^\vee$ is not surjective; the functional $(x_1, x_2, \dots) \longmapsto x_1 + x_2 + \dots$ in $V^\vee$ does not extend to a functional in $W^\vee$.

## Topological groups

**Solution B.11.**

(a) By Exercise 1.26, 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 $\ker(f) = f^{-1}(e)$.

**Solution B.12.** 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 B.9).

**(a)⇒(b):** This follows from Exercise 1.18.

**(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 Exercise 1.18.

**Solution B.13.**

(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 B.12 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 $b, c \in B \cap \left(\mathbf{R} \setminus \mathbf{Q}\right)$. Consider the bijective function $\sigma\colon B \longrightarrow B$ given by

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

Since $B$ is a $\mathbf{Q}$-basis of $\mathbf{R}$, $\sigma$ 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 $b \neq c$.

**Solution B.14.**

(a) Suppose $H$ is open. If $g$ is an element of $G$, then $gH$ is open because $gH = L_{g^{-1}}^{-1}(H)$ and $L_{g^{-1}}$ is continuous by Proposition B.9. Now the result follows from the equation

$$G \setminus H = \bigcup_{g \notin H} gH.$$

The converse does not hold. If $G = \mathbf{R}$, which is given the Euclidean topology, and if $H = \{0\}$, then $H$ is a closed subgroup of $G$ but it is not open.

(b) Suppose $H$ is closed. If $g$ is an element of $G$, then $L_{g^{-1}}$ is continuous by Proposition B.9, so $gH = L_{g^{-1}}^{-1}(H)$ is closed because of Exercise 1.17. Since $H$ is of finite index, it has are only finitely many cosets $H, g_1H, \ldots, g_nH$. It follows that

$$G \setminus H = \bigcup_{n=1}^{n} gH = G,$$

which is closed because it is a finite union of closed sets. Hence $H$ is open.

The converse does not hold. Let $G = \mathbf{R}$ but endow it with the discrete topology, and let $H = \mathbf{Z}$. Then $H$ is open in $G$ but it is not of finite index (because if it is, then $\mathbf{R}$ is a finite union of countable sets, and is thus countable by Exercise A.8).

(c) Arguing as in part (a), we have

$$G = \bigcup_{g \in G} gH,$$

so $\{gH : g \in G\}$ is an open cover of $G$. Since $G$ is compact, this open cover admits a finite sub-cover, which implies that $H$ has finite index.

(d) Yes. Let $G$ be any infinite group with the discrete topology, and let $H = \{e\}$, then $H$ is open in $G$ but it does not have finite index.

**Solution B.15.**

(a) If $g$ is an element of $S$, then $gT$ is open because $gT = L_{g^{-1}}^{-1}(T)$ and $L_{g^{-1}}$ is continuous by Proposition B.9. It then follows from

$$ST = \bigcup_{s \in S} sT$$

that $ST$ is open.

(b) If $S$ or $T$ is empty, then $ST \neq \varnothing$, so it is connected. Otherwise, the product $S \times T$ is connected by Exercise 1.45, so $ST = m(S \times T)$ is connected by Proposition 2.32.

(c) The product $S \times T$ is compact by Theorem 2.41, so $ST = m(S \times T)$ is compact by Proposition 2.39.

(d) Since inversion is a homeomorphism, it follows from Proposition 2.39 that $S^{-1}$ is compact. The inclusion $j\colon S^{-1} \times G \longrightarrow G \times G$ is continuous by Exercise 1.21. Since $T$ is closed, it follows from Exercise 1.17 that $m^{-1}(T) \subseteq G \times G$ is closed and then $j^{-1}\big(m^{-1}(T)\big) \subseteq S^{-1} \times G$ is closed.

We now claim that

$$ST = \pi_2\Big(j^{-1}\big(m^{-1}(T)\big)\Big);$$

and this implies $ST$ is closed because $\pi_2$ is closed by Theorem 2.41 (here we crucially need $S^{-1}$ to be compact). To prove this equation, we start with an element $g$ of $ST$. Since $g \in ST$, there exists an element $s$ of $S$ and an element $t$ of $T$ such that $g = st$. It follows that $(s^{-1}, g) \in j^{-1}\big(m^{-1}(T)\big)$, so

$$g \in \pi_2\Big(j^{-1}\big(m^{-1}(T)\big)\Big).$$

For the other inclusion, suppose $(s', g) \in j^{-1}\big(m^{-1}(T)\big)$, It follows that $s'g \in T$, so $g \in s'^{-1}T$, which implies $g \in ST$ because $s' \in S^{-1}$. Hence the equation holds.

(e) Since $\mathbf{Z} + \pi\mathbf{Z} = \bigcup_{n \in \mathbf{Z}}(n + \pi\mathbf{Z})$, it follows from Exercise A.8 that $\mathbf{Z} + \pi\mathbf{Z} \neq \mathbf{R}$; but we know it is dense in $\mathbf{R}$, so it cannot be closed. Hence $\mathbf{Z}$ and $\pi\mathbf{Z}$ are closed in $\mathbf{R}$, but $\mathbf{Z} + \pi\mathbf{Z}$ is not closed.

**Solution B.16.** By Proposition 3.3, the addition $a\colon V \times V \longrightarrow V$ and the scalar multiplication $s\colon \mathbf{F} \times V \longrightarrow V$ are both continuous. It is straightforward to verify that the inversion $i\colon V \longrightarrow V$ defined by $i(v) = -v$ is equal to the composition $s' \circ f$, where $s'\colon \{-1\} \times V \longrightarrow V$ is the restriction of $s$ to $\{-1\} \times V$ and $f\colon V \longrightarrow \{-1\} \times V$ defined by $f(v) = (-1, v)$. The function $s'$ is continuous because it is the restriction of the continuous function $s$ to $\{-1\} \times V$, while $f'$ is continuous by Exercise 1.24, so the inversion $i$ is continuous by Tutorial Question 2.9. Hence $V$ is a topological group.