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 Wed 6th Aug, 2025 at 15:46

*(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 $y \in X \smallsetminus C$, then $d(x,y) > r$. Set $t = d(x,y) - r$ and consider the open ball $\mathbf{B}_t(y)$.

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

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

hence $w \notin C$.

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 $v_p(x+y) \geqslant v = \min\{v_p(x), v_p(y)\}$.

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

(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

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

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

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

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

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

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

For the second statement, let $\{C_i \colon i \in I\}$ 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

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

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

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 $f(x) \notin \big(Y \setminus S\big)$ iff $x \notin f^{-1}(Y \setminus S)$.

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

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

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

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

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

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

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 $x \in U_x^2$ and $U_x^2 \in \mathcal{T}_2 \subseteq \mathcal{T}_1$, we can take $U_x^1 = U_x^2$ and we are done.

(b)⇒(a): Let $U \in \mathcal{T}_2$. We use Exercise 1.14 to prove that $U \in \mathcal{T}_1$. Let $x \in U$, then setting $U_x^2 = U$ we have that U_x^2 is a $\mathcal{T}_2$-open neighbourhood of x, so by (b) there exists a cT_1-open neighbourhood U_x^1 of x such that $U_x^1 \subseteq U$. By Exercise 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 ι 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) Since A and B are closed in X and Y respectively, their complements $X \setminus A$ and $Y \setminus B$ are open in X and Y respectively, and therefore $(X \setminus A) \times Y$ and $X \times (Y \setminus B)$ are open in $X \times Y$. It follows that

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

is closed 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 3.4 that ι_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 $\overline{A} \times \overline{B} \subseteq \overline{A \times B}$.

Solution 1.26.

(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 π_2.

Solution 1.27. Let $U = X \setminus \{x\}$ and let $u \in U$. Then $u \neq x$, so by the Hausdorff property of X, there exist open neighbourhoods V_1 of u and V_2 of x such that $V_1 \cap V_2 = \varnothing$. In particular, $x \notin V_1$, so $V_1 \subseteq U$. As we have exhibited an open neighbourhood contained in U around every element of U, we conclude by Exercise 1.14 that U is open, so its complement $\{x\}$ is closed.

Interior and closure

Solution 1.28. 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 1.29.

(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 closed 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 $\iota_x \colon Y \longrightarrow X \times Y$ by $\iota_x(y) = (x,y)$. Let $\pi_X \colon X \times Y \longrightarrow X$ and $\pi_Y \colon X \times Y \longrightarrow Y$ 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 3.4 that ι_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 $\overline{A} \times \overline{B} \subseteq \overline{A \times B}$.

Solution 1.30. 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 1.31. Since $A^\circ \subseteq A$, we have $\big(X \smallsetminus A\big) \subseteq \big(X \smallsetminus A^\circ\big)$. But A° 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 $(X \setminus A) \subseteq \overline{X \setminus A}$, so

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

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

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

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

Solution 1.32. First we show that $\overline{\mathbf{Z}} = \mathbf{Z}$: letting $U = \mathbf{R} \setminus \mathbf{Z}$, 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 1.33.

(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 $U \cap \left(X \setminus \overline{N}\right) = \varnothing$, 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 $X \setminus \overline{N}$ nontrivially, hence $X \setminus \overline{N}$ is dense.

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

(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 $X \setminus \overline{L} = \left(X \setminus \overline{M}\right) \cap \left(X \setminus \overline{N}\right)$. As $X \setminus \overline{L}$ 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 1.34. (a) i. Put $X = \{0,1\}$, $Y = \{1\}$, $\mathcal{T}_Y = \mathcal{P}(Y)$. Let $f: X \longrightarrow Y$ be the function sending both 0 and 1 to 1. It follows that $\mathcal{T}_X = \{\varnothing, \{0,1\}\}$. The topology $\mathcal{T}_Y$ is defined by the discrete metric (see Tutorial Question 2.1), but $\mathcal{T}_X$ is not metrisable (see Tutorial Question 2.3).

ii. Put $X = \{1\}$, $Y = \{0,1\}$, $\mathcal{T}_Y = \left\{\varnothing, \{1\}, \{0,1\}\right\}$. Let $f: X \longrightarrow Y$ be the inclusion function, which sends 1 to 1. It follows that $\mathcal{T}_X = \mathcal{P}(X)$. The topology $\mathcal{T}_X$ is defined by the discrete metric (see Tutorial Question 2.1), but $\mathcal{T}_Y$ is not metrisable (see Tutorial Question 2.3).

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

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

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

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

Solution 1.35. 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 1.36.

(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 1.37. The inequalities involving d_1 and d_∞ follow simply from

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

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

The inclusions between open balls now follow by the same reasoning as in part (a) of Exercise 1.36.

Solution 1.38.

(a) We have

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

so that

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

(b) In one direction, suppose A is open in Y; by Tutorial Question 3.6 we have some indexing set I such that

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

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

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

which by Tutorial Question 3.6 is an open in X. It is clear that $A = U \cap Y$ by part (a).

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

Solution 1.39. 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,

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

2. Normed and Hilbert spaces

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 π 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 π 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 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 $f \nsim g$.

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

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

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

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 $W_n = \operatorname{Span}\{b_1, \ldots, b_n\}$. 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 $n \in \mathbf{N}$, $a_1, \ldots, a_n \in \mathbf{Q}$ and $k_1, \ldots, k_n \in \mathbf{N}$ such that

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

Let $k = \max\{k_1, \ldots, k_n\}$, 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 $A' = P^{-1}BP$ where P is a (unitary) matrix whose columns are orthonormal eigenvectors of A, and $B = \operatorname{diag}(b_1, \ldots, b_n)$ is a diagonal matrix with the corresponding eigenvalues. If $\mathbf{r}_i$ denotes the i^{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 Ω has at least two distinct elements x_1 and x_2, then $\{x_1\}$ and $\{x_2\}$ are not comparable under $\subseteq$, so the latter is not a total order.

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

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

Let $J = \{1, \ldots, n\}$. For each $j \in J$, there exists $S_j \in C$ such that $u_j \in S_j$. As C is totally ordered, there exists $i \in J$ such that $S_j \subseteq S_i$ for all $j \in J$. But this means that $u_1, \ldots, u_n \in S_i$, so that the linear relation of Equation (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 $v \in V \smallsetminus \operatorname{Span}(B)$, then $B' \coloneqq B \cup \{v\}$ 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 $A \coloneqq \bigcup_{i \in I} A_i$, 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 $y_0 \in Y \smallsetminus M$. 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 $S = \{e_1, e_2, \ldots\}$ and $W = \operatorname{Span}(S)$.

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

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

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 $m \in \mathbf{N}$, $a_1, \ldots, a_m \in \mathbf{R}$ and $k_1, \ldots, k_m \in \mathbf{N}$ such that

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

Set $n = \max\{k_1, \ldots, k_m\}$, 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 $v = (1, 1, 1, \ldots) \notin W$. So S does not span $\mathbf{R}^{\mathbf{N}}$.

Solution B.6. This is a straightforward rewriting of the definition of algebraic: α is algebraic if and only if it satisfies a polynomial equation with coefficients in $\mathbf{Q}$, which is equivalent to a nontrivial linear relation between the powers of α, which exists if and only if T is linearly dependent.

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 $v_1, \ldots, v_n$:

$$T^\vee(v_i^\vee)(v_j) = (v_i^\vee \circ T)(v_j) = v_i^\vee\big(T(v_j)\big) = v_i^\vee\big(a_{1j}v_1 + a_{2j}v_2 + \cdots + a_{nj}v_n\big) = a_{ij}.$$

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 Γ is injective. Let $W = \operatorname{Span}\{v_1, \ldots, v_n\}$. 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 Γ.

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 Γ is injective.

(c) Suppose $\Gamma \colon V^\vee \longrightarrow \mathbf{F}^n$ is surjective. Let

$$a_1 v_1 + \cdots + a_n v_n = 0$$

be a linear relation.

Let $i \in \{1, \dots, n\}$. Since Γ is surjective, given the standard basis vector $e_i \in \mathbf{F}^n$ (1 in the i-th entry), there exists $\varphi_i \in V^\vee$ such that $\Gamma(\varphi_i) = e_i$. If we apply φ_i on both sides of the linear relation, we get

$$a_i = 0.$$

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

Conversely, suppose $\{v_1, \dots, v_n\}$ 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 Γ 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^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, $\operatorname{rank}(A) = n = \dim(V)$. Then $\operatorname{rank}(A^\mathsf{T}) = n = \dim(V^\vee)$, so A^T has full-rank and thus $T^\vee$ is surjective.

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

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

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 $W^\vee$.