Solutions Manual

a supplement to

Functional Analysis by Sergei Ovchinnikov

Springer 2018

ISBN 978-3-319-91511-1

Contents

0 Using the Manual

As I stated in the Preface to the book, "... the most effective way of learning mathematics is by 'doing it'." Accordingly, I urge the student not to read the solutions in advance, but rather to make a concerted effort to find a solution to the problem in the exercise before consulting the Solution Manual to verify correctness. If the student's solution differs from the one given in the Manual, a comparison might reveal an unjustified assumption that had been made by the student or a misapplication of a theorem. Meanwhile, the instructor can use the solutions to create balanced assignments and research projects.

Solutions in the Manual are labeled in the same way as exercises in the book. For instance, item **1.10** in the Manual is a solution to the problem in Exercise 1.10.

Sergei Ovchinnikov
sergei@sfsu.edu

1 Preliminaries

1.1. Assume, without loss of generality, that $\max\{x_1,\ldots,x_n\} = x_1$. Then

$$\begin{aligned}x = \lambda_1 x_1 + \cdots + \lambda_n x_n &\le \lambda_1 x_1 + \cdots + \lambda_n x_1\\ &= (\lambda_1 + \cdots + \lambda_n)x_1 = x_1 = \max\{x_1,\ldots,x_n\}.\end{aligned}$$

A similar argument proves the other inequality.

1.2. (Necessity.) Let $P = (a, f(a))$, $Q = (b, f(b))$, and $R = (x, f(x))$, where $a \le x \le b$ (cf. Fig. 1.3). It is easy to verify that $x = \lambda a + (1-\lambda)b$, where $\lambda = \dfrac{b-x}{b-a}$. Clearly, $0 \le \lambda \le 1$. Let R' be a point on the chord PQ with the abscissa x. Elementary geometry (similar triangles) shows that

$$R' = (x, \lambda f(a) + (1-\lambda)f(b)).$$

Inasmuch as the function f is convex, we have

$$f(x) \le \lambda f(a) + (1-\lambda)f(b). \tag{1.1}$$

Hence, the point R is below the point R' or coincides with it.

(Sufficiency.) In the above notations, the condition that R lies below R' or coincides with it is equivalent to (1.1). The result follows.

1.3.(a) Hint. Use elementary algebra to show that the inequality

$$f(x) \le \frac{b-x}{b-a}f(a) + \frac{x-a}{b-a}f(b)$$

is equivalent to each of the inequalities

$$\frac{f(x)-f(a)}{x-a} \le \frac{f(b)-f(a)}{b-a}, \qquad \frac{f(b)-f(a)}{b-a} \le \frac{f(b)-f(x)}{b-x}. \tag{1.2}$$

1.3.(b) (1) Let x be an inner point of I and $G_x = \dfrac{f(x)-f(t)}{x-t}$ for $t \in I$. For three consecutive points $t_1 < t_2 < x$ we have, by the second inequality in (1.2) above,

$$\frac{f(x)-f(t_1)}{x-t_1} \le \frac{f(x)-f(t_2)}{x-t_2},$$

that is, $G_x(t_1) \le G_x(t_2)$. By applying a similar argument to points $t_1 < x < t_2$ and $x < t_1 < t_2$ we prove that $G_x(t_1) \le G_x(t_2)$ for every $t_1 < t_2$. Hence, $G_x(t)$ is an increasing function of $t \in I$. Clearly, $G_x(t)$ is bounded on the interval $(-\infty, x) \cap I$. Therefore, the $\lim_{t\to x^-} G_x(t)$ exists. This limit is the left-hand derivative $f'(x^-)$. By a similar argument, the right-hand derivative $f'(x^+)$ also exists.

(2) Let $t_1 < x_1 < t_2 < x_2$ be four points in I. By (1.2), it is not difficult to see that $G_{x_1}(t_1) \leq G_{x_2}(t_2)$. By taking limits, we obtain $f'(x_1^-) \leq f'(x_2^-)$. A similar argument shows that $f'(x_1^+) \leq f'(x_2^+)$.

(3) Inasmuch as $G_x(t_1) \leq G_x(t_2)$ for $t_1 < x < t_2$, we obtain $f'(x^-) \leq f'(x^+)$ on I.

1.3.(c) (Necessity.) If f is differentiable on I, then $f'(x) = f'(x^-) = f'(x^+)$ for $x \in I$. By **1.3.b)** part (2), f' is an increasing function on I.

(Sufficiency.) Let f be a differentiable function on I such that $f'(x)$ is increasing. For $x_1 < x < x_2$ in I, by the Mean Value Theorem, there are $c_1 \in (x_1, x)$ and $c_2 \in (x, x_2)$ for which

$$f'(c_1) = \frac{f(x) - f(x_1)}{x - x_1} \quad \text{and} \quad f'(c_2) = \frac{f(x_2) - f(x)}{x_2 - x}.$$

Because $f'(c_1) \leq f'(c_2)$, we have

$$\frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x)}{x_2 - x}.$$

As in **1.3.a)**, it can be shown that the above condition is equivalent to convexity of the function f.

If f is a concave function, then $-f$ is convex and the above argument can be used to prove the second part of Theorem 1.2.

1.4.(a) (Necessity.) We have

$$\begin{aligned}\lambda f(x) + (1-\lambda) f(y) &= \lambda p x + \lambda q + (1-\lambda) p y + (1-\lambda) q \\ &= p(\lambda x + (1-\lambda) y) + q = f(\lambda x + (1-\lambda) y).\end{aligned}$$

(Sufficiency.) Substitute $x = 1$, $y = 0$ into the equation in the Exercise 1.4.a):

$$f(\lambda) = \lambda f(1) + f(0) - \lambda f(0) = [f(1) - f(0)]\lambda + f(0).$$

For $p = f(1) - f(0)$, $q = f(0)$, we have $f(\lambda) = p\lambda + q$ for $\lambda \in \mathbf{R}$.

1.4.(b) (Necessity.) Follows from part (a).

(Sufficiency.) Suppose that f is not affine. By Exercise 1.2 (cf. **1.2.** above), there are three consecutive points P, R, Q on the graph of f such that the point R lies below the point on the chord PQ with the same abscissa x (cf. Fig. 1.3). Because the graph of $-f$ is the reflection of the graph of f across the x-axis, the function $-f$ is not convex by the same exercise, a contradiction.

1.4.(c) Let $x = \lambda a + (1-\lambda) b$, which is a pont in (a, b). Suppose that f is not affine. By Exercise 1.2, there is a point $c \in (a, b)$ such that the point $S = (c, f(c))$ lies below the pont S' on the chord PQ with the same abscissa

c, where $P = (a, f(a))$, $Q = (b, f(b))$. We let $R = (x, f(x))$. Without loss of generality, we assume that $a < c < x$. We have

$$\mathrm{slope}(SR) > \mathrm{slope}(S'R) = \mathrm{slope}(RQ),$$

which contradicts **1.3.(a)**. (To follow the above argument, it is helpful to draw these points on the coordinate plane.)

1.5. Standard properties of complex numbers.

1.6. Hint. Let $y_k = 1$ for $1 \leq k \leq n$ in the Cauchy–Schwartz inequality (Theorem 1.5).

1.7. Hint. By Theorem 1.4, partial sums of the series on the left-hand side of (1.10) are bounded by the product of the corresponding partial sums on the right-hand side. They are also nondecreasing. Use the Comparison Principle for sequences to obtain the desired inequality. Same for (1.12).

1.8. Let $p' = 1/p > 1$. The conjugate exponent is $q' = 1/(1-p)$. Apply Hölder's Inequality (cf. (1.7) on p.4) to $u_k = |x_k y_k|^p$ and $v_k = |y_k|^{-p}$ with conjugate exponents p' and q'

$$\sum_{k=1}^{n} u_k v_k \leq \left(\sum_{k=1}^{n} u_k^{p'} \right)^{1/p'} \left(\sum_{k=1}^{n} v_k^{q'} \right)^{1/q'}$$

to obtain

$$\sum_{k=1}^{n} |x_k|^p \leq \left(\sum_{k=1}^{n} |x_k y_k| \right)^p \left(\sum_{k=1}^{n} |y_k|^q \right)^{1-p},$$

where $q = p/(p-1)$ is the conjugate exponent of p. Divide both sides by $\left(\sum_{k=1}^{n} |y_k|^q \right)^{1-p}$, then raise both sides of the new inequality to the power $1/p$ to obtain the desired inequality.

1.9. ***Correction.*** In the present version, the inequality does not hold. Indeed, for $x_k = 1$, $y_k = -1$, $1 \leq k \leq n$ we obtain a false inequality

$$0 \geq n^{1/p} + n^{1/p}.$$

However, if we assume that all x_k's and y_k's are nonnegative real numbers and not all y_k's (or all x_k's) are zero, the inequality holds. Here is a proof. We apply the reverse Hölder's Inequality (Exercise 1.8) to obtain

$$\begin{aligned}
\sum_{k=1}^{n}(x_k+y_k)^p &= \sum_{k=1}^{n} x_k(x_k+y_k)^{p-1} + \sum_{k=1}^{n} y_k(x_k+y_k)^{p-1} \\
&\geq \left(\sum_{k=1}^{n} x_k^p\right)^{1/p} \left(\sum_{k=1}^{n}(x_k+y_k)^{q(p-1)}\right)^{1/q} \\
&+ \left(\sum_{k=1}^{n} y_k^p\right)^{1/p} \left(\sum_{k=1}^{n}(x_k+y_k)^{q(p-1)}\right)^{1/q} \\
&= \left[\left(\sum_{k=1}^{n} x_k^p\right)^{1/p} + \left(\sum_{k=1}^{n} y_k^p\right)^{1/p}\right] \left(\sum_{k=1}^{n}(x_k+y_k)^{p}\right)^{1/q},
\end{aligned}$$

which is equivalent to the desired inequality. (Note that $q(p-1)=p$ and $1-1/q=1/p$.)

1.10. Inasmuch as the natural logarithm is a concave increasing function, we have

$$\ln\left(\frac{1}{n}\sum_{k=1}^{n} x_k\right) \geq \frac{1}{n}\sum_{k=1}^{n} \ln x_k = \sum_{k=1}^{n} \ln \sqrt[n]{x_k} = \ln \sqrt[n]{\prod_{k=1}^{n} x_k}$$

Hence, $G \leq A$. The inequality $H \leq G$ follows from $G \leq A$ (straightforward algebra).

1.11. By the convexity of f,

$$\frac{a}{a+b} f\Big(\frac{x}{a}\Big) + \frac{b}{a+b} f\Big(\frac{y}{b}\Big) \geq f\Big(\frac{ax}{(a+b)a} + \frac{by}{(a+b)b}\Big) = f\Big(\frac{x+y}{a+b}\Big)$$

which is the desired inequality.

1.12. ***Correction.*** Exercise 1.12 refers to inequalities (1.15) (Theorem 1.12, Hölder's Inequality) and (1.17) (Theorem 1.14, Minkowski's Inequality).

We prove only Theorem 1.12. We may assume that $\int_0^1 |x|^p > 0$ because otherwise the function x is zero almost everywhere on $[0,1]$ and the product xy is also zero almost everywhere. The same is true if $\int_0^1 |y|^q = 0$. Hence we set $u = (\int_0^1 |x|^p)^{1/p}$ and $u = (\int_0^1 |y|^q)^{1/q}$ and follow the steps in the proof of Theorem 1.4.

1.13. We need to show that $\sim$ is a reflexive, symmetric, and transitive binary relation on $\mathbf{N} \times \mathbf{N}$.

(Reflexivity.) Clearly, $(m,n) \sim (m,n)$ holds for any $m,n \in \mathbf{N}$.

(Symmetry.) If $m+q=p+n$, then $p+n=m+q$. Hence, $(m,n) \sim (p,q)$ implies $(p,q) \sim (m,n)$ for all $m,n,p,q \in \mathbf{N}$.

(Transitivity.) Suppose $(m, n) \sim (p, q)$ and $(p, q) \sim (r, s)$. Then we have

$$m + q = p + n \qquad \text{and} \qquad p + s = r + q.$$

Therefore,

$$(m + q) + (p + s) = (p + n) + (r + q).$$

By the associative, commutative, and cancellation laws of addition,

$$m + s = r + n,$$

that is, $(m, n) \sim (r, s)$.

The equivalence classes of the relation $\sim$ are integers.

1.14. (Reflexivity.) $(m, n) \sim (m, n)$ because $mn = mn$.

(Symmetry.) If $mn = pq$, then $pq = mn$. Hence,

$$(m, n) \sim (p, q) \quad \text{implies} \quad (p, q) \sim (m, n).$$

(Transitivity.) Suppose $(m, n) \sim (p, q)$ and $(p, q) \sim (r, s)$, so $mq = pn$ and $ps = rq$. Note that n, q, and s are assumed to be nonzero integers. We have

$$(mq)(ps) = (pn)(rq).$$

By the associative, commutative, and cancellation laws of multiplication, it follows that $ms = nr$, that is, $(m, s) \sim (r, s)$.

1.15. Let a, b be two distinct elements of X. Consider $A = X \setminus \{b\}$ and $B = X \setminus \{a\}$.

1.16. Let $u = a + b\boldsymbol{i}$, $v = c + d\boldsymbol{i}$, and $w = e + f\boldsymbol{i}$.

(Reflexivity.) $u \preceq u$ because $a \leq a$, $b \leq b$.

(Symmetry.) Let $u \preceq v$ and $v \preceq u$. Then $a \leq c$, $b \leq d$, $c \leq a$, $d \leq b$. Hence, $a = c$, $b = d$, that is, $u = v$.

(Transitivity.) Let $u \preceq v$ and $v \preceq w$. Then $a \leq c$, $b \leq d$, $c \leq e$, $d \leq f$. Hence, $a \leq e$, $b \leq f$, that is, $u \preceq w$.

1.17. (a) The disks that are tangent to at least two sides of the square. Justify!

(b) A greatest disk must be maximal. By part (a), there is no greatest disk in X.

(c), (d) If we allow for disks of zero radius (points), then they are minimal elements. Otherwise there are no minimal elements in X. Clearly, there is no least element in either case.

2 Metric Spaces

2.1. If conditions **M1**–**M3** hold, then

$$d(x,z)+d(y,z)=d(x,z)+d(z,y)\geq d(x,y), \qquad \text{for all } x,y,z\in X.$$

Hence, (2) holds. Clearly, (1) also holds.

If conditions (1) and (2) of the exercise hold, then we have, by substituting $z=x$ into (2) and applying (1), $d(x,y)\leq d(y,x)$ for all $x,y\in X$. By exchanging x and y, we obtain $d(x,y)\geq d(y,x)$. Hence, **M2** holds. **M3** follows immediately from **M2** and (2).

2.2. Suppose that the three triangle inequalities hold. The second inequality in (2.1) on page 20 is the first inequality of the three triangle inequalities. The first inequality in (2.1) is equivalent to

$$-d(x,y)\leq d(x,z)-d(y,z)\leq d(x,y).$$

The first and the second inequalities in the previous line are equivalent to the second and third triangle inequalities, respectively.

A similar argument shows that (2.1) implies the three triangle inequalities.

2.3. We need to show that

$$-d(x,u)-d(y,v)\leq d(x,y)-d(u,v)\leq d(x,u)+d(y,v). \tag{2.1}$$

We have

$$d(x,u)+d(u,v)+d(y,v)\geq d(x,v)+d(v,y)\geq d(x,y).$$

Hence the second inequality in (2.1) holds. A similar argument proves the first inequality in (2.1).

2.4. Base case, $n=3$, is property **(M3)** on page 19. Suppose (2.2) on page 20 holds for $n=k>3$. By **(M3)** and the induction hypothesis,

$$\begin{aligned} d(x_1,x_{k+1}) &\leq d(x_1,x_k)+d(x_k,x_{k+1}) \\ &\leq d(x_1,x_2)+\cdots+d(x_{k-1},x_k)+d(x_k,x_{k+1}). \end{aligned}$$

2.5.(a) Clearly, **(M1)** and **(M2)** hold. Let $\alpha=d(x,y)$, $\beta=d(y,z)$, $\gamma=d(x,z)$, so $\gamma\leq\alpha+\beta$. Inasmuch as the function $f(t)=t/(1+t)$ is increasing on $[0,\infty)$ (cf. Theorem 1.11), we have

$$\frac{\gamma}{1+\gamma}\leq\frac{\alpha+\beta}{1+(\alpha+\beta)}=\frac{\alpha}{1+(\alpha+\beta)}+\frac{\beta}{1+(\alpha+\beta)}\leq\frac{\alpha}{1+\alpha}+\frac{\beta}{1+\beta},$$

so **(M3)** holds for $\widetilde{d}$.

2.5.(b) Clearly, **(M1)** and **(M2)** hold. We have

$$\begin{aligned}\ln(1+d(x,y))+\ln(1+d(y,z)) &= \ln(1+d(x,y)+d(y,z)+d(x,y)d(y,z)) \\ &\geq \ln(1+d(x,y)+d(y,z)) \geq \ln(1+d(x,z)),\end{aligned}$$

so **(M3)** holds for $\widetilde{d}$.

2.5.(c) Clearly, **(M1)** and **(M2)** hold. To prove **(M3)**, let $d(x,y)=s$, $d(y,z)=r$, and $d(x,z)=t$, so $\widetilde{d}(x,y)=s^\alpha$, $\widetilde{d}(y,z)=r^\alpha$, $\widetilde{d}(x,z)=t^\alpha$, where $x,y,z \in X$. We may assume that $s>0$, (why ?). Let $u=r/s$. The function

$$f(u)=1+u^\alpha-(1+u)^\alpha$$

is increasing on $(0,\infty)$. (Use the 1st Derivative Test and note that $\alpha-1<0$.) Hence, $1+u^\alpha \geq (1+u)^\alpha$, which is equivalent to $s^\alpha+r^\alpha \geq (s+r)^\alpha$. Because $s+r \geq t$, we obtain $s^\alpha+r^\alpha \geq t^\alpha$—the desired triangle inequality for $\widetilde{d}$.

2.6. Note that pseudometric is a nonnegative function on $X \times X$:

$$2d(x,y)=d(x,y)+d(y,x) \geq d(x,x)=0, \qquad \text{for all } x,y \in X.$$

Clearly, $\sim$ is reflexive and symmetric (by **(PM2)**). If $x \sim y$ and $y \sim z$, then $d(x,y)=0$ and $d(y,z)=0$. By **(PM3)**, $d(x,z)=0$, so $x \sim z$. Hence, $\sim$ is transitive.

A straightforward verification shows that $X/\sim$ is a metric space.

2.7. (a) By substituting $x=0$ into $p(\alpha x)=|\alpha|p(x)$, we obtain $p(0)=0$.
(b) For $\alpha=-1$ we obtain $p(-x)=p(x)$ from $p(\alpha x)=|\alpha|p(x)$.
(c) For $y=-x$ we obtain $p(x)\geq 0$ from $p(x+y)\leq p(x)+p(y)$ and (b).

A straightforward verification shows that $d(x,y)=p(x-y)$ is a pseudometric.

2.8. Only the case of metrics is considered. The other case is similar. We verify the defining properties of a metric.

(M1) If $x=y$, then, clearly, $d(x,y)=0$. Suppose now that $d(x,y)=0$. We may assume that $\lambda_1>0$. Then $\lambda_1 d(x,y)=0$. Hence, $x=y$.

(M2) Trivial.

(M3)

$$d(x,y)+d(y,z)=\sum_{k=1}^{n}\lambda_k(d_k(x,y)+d_k(y,z)) \geq \sum_{k=1}^{n}\lambda_k d_k(x,z)=d(x,z).$$

2.9. Only the case of metrics is considered. The other case is similar.

The series in Exercise 2.9 converges by the Comparison Test, because

$$\frac{1}{2^k}\frac{d_k}{1+d_k}<\frac{1}{2^k}.$$

We verify the defining properties of a metric.

(M1) $\widetilde{d}(x,x) = 0$ because $d_k(x,x) = 0$ for all $1 \leq k \leq n$. On the other hand, if $\widetilde{d}(x,y) = 0$, then $d_1(x,y) = 0$, so $x = y$.

(M2) Trivial.

(M3) By Exercise 2.5 (a), the functions $D_k = \dfrac{d_k}{1+d_k}$, $k \in \mathbf{N}$, are metrics. By properties of convergent series, we have

$$\widetilde{d}(x,y) + \widetilde{d}(y,z) = \sum_{k=1}^{\infty} \frac{1}{2^k}\big(D_k(x,y) + D_k(y,z)\big) \geq \sum_{k=1}^{\infty} \frac{1}{2^k} D_k(x,z) = \widetilde{d}(x,z).$$

2.10. Let $r' = r - d(x,y) > 0$, so $d(x,y) = r - r'$. For $z \in B(y,r')$, we have

$$d(x,z) \leq d(x,y) + d(y,z) < r - r' + r' = r.$$

Hence, $z \in B(x,r)$. It follows that $B(y,r') \subseteq B(x,r)$.

Because $B(x,r)$ is the union of the balls $B(y,r')$ for $y \in B(x,r)$ and corresponding r''s, $B(x,r)$ is an open set.

2.11. Let $X = [0,1] \cup \{2\}$ with the metric inherited from $\mathbf{R}$, and $x_1 = 1$, $r_1 = 1.1$, $x_2 = 0$, $r_2 = 1.5$. We have $B(1,1.1) = [0,1] \cup \{2\}$, $B(0,1.5) = [0,1]$, so $B(0,1.5)$ is a proper subset of $B(1,1.1)$. Note that $1.5 > 1.1$.

2.12. ***Correction.*** The statement is trivial in this form. There should be "$B(x_1,r_1)$ is a proper subset of $B(x_2,r_2)$". Solution:

Take $z \in B(x_2,r_2) \setminus B(x_1,r_1)$. Then, by the first inequality in (2.1) on page 20,

$$r_2 > d(x_1,x_2) \geq |d(x_1,z) - d(x_2,z)| > |r_1 - r_2|,$$

because $d(x_1,z) > r_1$, $d(x_2,z) < r_2$. The inequality $|r_1 - r_2| < r_2$ is equivalent to $-r_2 < r_1 - r_2 < r_2$. Hence, $r_1 < 2r_2$. (Compare this with Exercise 2.11.)

2.13. (a) Suppose there is $x \in B(x_1,r_1) \cap B(x_2,r_2)$. Then $d(x,x_1) < r_1$ and $d(x,x_2) < r_2$ which implies $d(x_1,x_2) \leq d(x_1,x) + d(x,x_2) < r_1 + r_2$, a contradiction.

(b) For $x \in B(x_2,r_2)$ we have

$$d(x_1,x) \leq d(x_1,x_2) + d(x_2,x) < (r_1 - r_2) + r_2 = r_1.$$

Therefore, $x \in B(x_1,r_1)$. Hence, the result.

Counterexamples. (a) X is a discrete space, not a singleton, and open balls are $B(x_1,1) = \{x_1\}$, $B(x_2,1) = \{x_2\}$ where $x_1 \neq x_2$.

(b) Solution to Exercise 2.11.

2.14. Let $z \in B(x,r_1) \cap B(y,r_2)$. By Exercise 2.10, there are open balls $B(z,r')$ and $B(z,r'')$ that are contained in the balls $B(x,r_1)$ and $B(y,r_2)$, respectively. The ball $B(z,r)$ where $r = \min\{r',r''\}$ is contained in $B(x,r_1) \cap B(y,r_2)$.

2.15. Clearly, X is a closed set. By Theorem 2.2, $\varnothing = X^c$ is open.

2.16. ***Correction.*** $\mathbf{Q}$ in the exercise is considered as a subset of $\mathbf{R}$. Hint: Every nonempty interval in $\mathbf{R}$ contains an irrational point. Note that any metric space is interior of itself.

2.17. Let a, b be two distinct points in a metric space X. The sets $\{a\}$ and $\{b\}$ are closed. Hence, $X \setminus \{a\}$ and $X \setminus \{b\}$ are distinct open sets that are, clearly, different from X and $\varnothing$.

2.18. Let $y \in \overline{B(x,r)}$. We may assume that y is a limit point of $B(x,r)$ that does not belong to this ball. For every $\varepsilon > 0$, the ball $B(y,\varepsilon)$ contains a point $z \in B(x,r)$. We have

$$d(x,y) \leq d(x,z) + d(z,y) < r + \varepsilon,$$

which implies $d(x,y) \leq r$. Hence, $y \in \overline{B}(x,r)$.

Example. Let $X = (0,1) \cup \{2\} \cup \{4\}$ with the metric $d(x,y) = |x-y|$. The open ball $B(2,2) = (0,1) \cup \{2\}$ is a closed set which is a proper subset of the closed ball $\overline{B}(2,2) = (0,1) \cup \{2\} \cup \{4\}$.

2.19. We prove only the first statement of the theorem. Let O be an open set contained in E. Every point in O is a point in $\operatorname{int} E$. Therefore, $\operatorname{int} E$ contains every open set contained in E and is the union of the family of these sets. Hence, $\operatorname{int} E$ is the largest open set contained in E.

2.20. (a) It suffices to note that every limit point of A is a limit point of B.

(b) Because $\overline{A \cup B}$ is the smallest closed set containing $A \cup B$ (cf. Theorem 2.4), we have $\overline{A \cup B} \subseteq \overline{A} \cup \overline{B}$. Because A and B are subsets of $A \cup B$, by part (a), $\overline{A} \cup \overline{B} \subseteq \overline{A \cup B}$.

(c) Similar to part (b). Counterexample. Let $A = (0,1)$, $B = (1,2)$ in $\mathbf{R}$. Then $\overline{A \cap B} = \varnothing$, but $\overline{A} \cap \overline{B} = \{1\}$.

2.21. (Necessity.) Suppose that Z is closed in X. Then $X \setminus Z$ is open in X. Hence there is an open set O in the space Y such that $X \setminus Z = O \cap X$. We have

$$Z = X \setminus (X \setminus Z) = X \setminus (O \cap X) = X \setminus O = (Y \setminus O) \cap X.$$

We set $F = Y \setminus O$ which is a closed subset of Y.

(Sufficiency.) Let $Z = F \cap X$, where F is a closed subset of Y. Then $O = Y \setminus F$ is open in Y. We have

$$X \setminus Z = X \setminus (F \cap X) = X \setminus F = (Y \setminus F) \cap X = O \cap X,$$

so $X \setminus Z$ is open in X. It follows that Z is closed in X.

2.22. We denote $B^d(x,r)$ and $B^{\widetilde{d}}(x,s)$ open balls in metric spaces (X,d) and $(X,\widetilde{d})$, respectively. We have $B^d(x,r) \subseteq B^{\widetilde{d}}(x,r)$ because

$$d(x,y) < r \qquad \text{implies} \qquad \widetilde{d}(x,y) = \frac{d(x,y)}{1+d(x,y)} \leq d(x,y) < r.$$

On the other hand, $B^{\widetilde{d}}(x,s) \subseteq B^d(x,r)$, where $s = r/(1+r)$. Indeed,

$$\widetilde{d}(x,y) < s \qquad \text{is equivalent to} \qquad \frac{d(x,y)}{1+d(x,y)} < \frac{r}{1+r},$$

which implies $d(x,y) < r$ because the function $f(t) = t/(1+t)$ is strictly increasing on $[0,\infty)$.

To complete the solution it suffices to note that every nonempty open set in a metric space is the union of open balls that are contained in it.

2.23. Hint. Show that the set $E = \{x + y\boldsymbol{i} : x, y \in \mathbf{Q}\}$ is dense in $\mathbf{C}$.

2.24. (Necessity.) Clearly a countable discrete space is dense in itself and hence is separable.

(Sufficiency.) Suppose to the contrary that an uncountable discrete space X contains a countable dense subset E, and let $a \in X \setminus E$. Every ball $B(a,r)$ with $r \leq 1$ has an empty intersection with E. Hence the result.

2.25. Hint. A point x is a limit point of the set A if and only if it is a limit point of the set $A \setminus \{x\}$.

2.26. Let Y be a subspace of a separable space X and $A = \{x_1, x_2, \ldots, x_n, \ldots\}$ be a dense subset of X. For every $y \in Y$ and $m \in \mathbf{N}$, there is x_n such that $d(y,x_n) < 1/m$, so $B(x_n,1/m) \cap Y \neq \varnothing$. We consider only pairs (n,m) for which the intersection $B(x_n,1/m) \cap Y$ is not empty, denote this set K, and select $y_{n,m} \in Y$ such that $d(x_n,y_{n,m}) < 1/m$. For a given $\varepsilon > 0$ we assume that $1/m \leq \varepsilon/2$. For every $n \in \mathbf{N}$ we choose m such that $d(x_n,y_{n,m}) < 1/m < \varepsilon/2$. Now choose n such that $d(y,x_n) < 1/m$. Finally, we have

$$d(y,y_{n,m}) \leq d(y,x_n) + d(x_n,y_{n,m}) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

It follows that the countable set $\{y_{n,m} : (n,m) \in K\}$ is dense in Y.

2.27. Because the set $A = \{x_n : n \in \mathbf{N}\}$ is finite, there is $a \in A$ and a sequence $n_1 < n_2 < \cdots < n_k < \cdots$ such that $x_{n_k} = a$ for all $k \in \mathbf{N}$. Hence the subsequence (x_{n_k}) converges.

2.28. (Necessity.) Trivial.

(Sufficiency.) Suppose that every neighborhood of x contains infinitely many terms of (x_n). We can choose n_1 such that $d(x_{n_1},x) < 1$, then choose $n_2 > n_1$

such that $d(x_{n_2}, x) < 1/2$. By continuing this process, we construct a sequence $n_1 < n_2 < \cdots < n_k < \cdots$ satisfying $d(x_{n_k}, x) < 1/k$ for all $k \in \mathbf{N}$. Then $x_{n_k} \to x$.

2.29. Hint. Apply the definition of limit.

2.30. Let $\varepsilon > 0$. Because (x_n) is Cauchy, there is N_1 such that $d(x_n, x_m) < \varepsilon/2$ for all $n, m > N_1$. Because $x_{n_k} \to x$, there is N_2 such that $d(x_{n_k}, x) < \varepsilon/2$ for all $k > N_2$. Let $k > N_2$. For all $n > n_k > N_1$, we have

$$d(x_n, x) \leq d(x_n, x_{n_k}) + d(x_{n_k}, x) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Therefore, $x_n \to x$.

2.31. Let (x_n) be a Cauchy sequence. There is N such that $d(x_n, x_m) < 1$ for all $n, m > N$. Let $m_0 > N$. Hence, $x_n \in B(x_{m_0}, 1)$ for all $n > N$. For

$$r = \max\{1, d(x_1, x_{m_0}), \ldots, d(x_N, x_{m_0})\}$$

all points of the sequence (x_n) belong to the ball $B(x_{m_0}, r)$.

2.32. There is N_1 such that $d(x_n, x_m) < 1/2$ for all $n, m > N_1$. We set $n_1 = N_1 + 1$. Note that $d(x_{n_1}, x_m) < 1/2$ for all $m > n_1$.

There is $N_2 > N_1$ such that $d(x_n, x_m) < 1/4$ for all $n, m > N_2$. We set $n_2 = N_2 + 1$. Note that $n_2 > n_1$ and $d(x_{n_2}, m) < 1/4$ for all $m > n_2$.

By continuing this process we obtain the required subsequence.

2.33. Clearly, every sequence in a singleton is convergent. In the other direction, suppose to the contrary that the space X is not a singleton. Let a, b be two distinct points in X. The sequence (x_n) defined by $x_{2k-1} = a$, $x_{2k} = b$ for $k \in \mathbf{N}$ is bounded but not convergent.

2.34. Let $x_n \to x$ in a discrete space X. There is $N \in \mathbf{N}$ such that $d(x_n, x) < 1$ for $n > N$, which implies $x_n = x$ for $n > N$. The converse is obvious.

2.35. (a) Suppose that $1/n \to a \in (0, 1)$. Then for $\varepsilon = a/2$ there is $N \in \mathbf{N}$ such that

$$\left|\frac{1}{n} - a\right| < \frac{a}{2}, \qquad \text{for } n > N,$$

which implies $a < 1/n$ for $n > N$, a contradiction.

(b) Hint. For $m > n$, $\dfrac{1}{n} - \dfrac{1}{m} < \varepsilon$ if $n > \dfrac{1}{\varepsilon}$.

2.36. $\sum_{k=1}^{\infty} (-1)^{k-1} c_k$ is an alternating series. For $n \in \mathbf{N}$, let

$$t_n = r - \sum_{k=1}^{n} (-1)^{k-1} c_k$$

be the sum of its remainder. We have

$$0 < t_{2m} = (c_{2m+1} - c_{2m+2}) + \cdots = c_{2m+1} - (c_{2m+2} - c_{2m+3}) - \cdots < c_{2m+1}.$$

This proves the inequalities in the exercise for $n = 2m$. A similar argument establishes the desired inequalities for $n = 2m - 1$.

2.37. Let $\varepsilon = \min\{d(x,y) : x, y \in X,\ x \neq y\}$. If (x_n) is a Cauchy sequence in X, then there is $N \in \mathbf{N}$ such that $d(x_n, x_m) < \varepsilon$ for all $n, m > N$. It follows that $x_n = x_m$ for $n, m > N$, that is, the sequence is "eventually constant" and therefore convergent.

2.38. For $\varepsilon > 0$ and points $a, b \in \overline{B(x,r)}$, let y, z be points in $B(x,r)$ such that $d(a,y) < \varepsilon$ and $d(b,z) < \varepsilon$. We have

$$\begin{aligned} d(a,b) &\leq d(a,y) + d(y,z) + d(z,b) \leq 2\varepsilon + d(y,z) \\ &\leq 2\varepsilon + d(y,x) + d(x,z) < 2\varepsilon + 2r. \end{aligned}$$

Because 2ε is an arbitrary positive number, $d(a,b) \leq 2r$ for all $a, b \in \overline{B(x,r)}$.

2.39. Because $A \subseteq \overline{A}$, $\operatorname{diam}(A) \leq \operatorname{diam}(\overline{A})$. For the reverse inequality, apply steps from Solution **2.38**.

2.40. In a countable discrete space $X = \{a_1, a_2, \ldots\}$ let $A_n = \{a_n, a_{n+1}, \ldots\}$ for $n = 1, 2, \ldots$. Clearly, $\bigcap_{n=1}^{\infty} A_n = \varnothing$.

2.41. Let (x_n) be a Cauchy sequence in X. Define $E_n = \{x_k : k \geq n\}$ and $F_n = \overline{E}_n$ for $n = 1, 2, \ldots$. The sequence (E_n) is descending, so is the sequence (F_n). Inasmuch as for every $\varepsilon > 0$ there is N such that $d(x_n, x_m) < \varepsilon$ for $n, m > N$, $\operatorname{diam}(F_n) = \operatorname{diam}(E_n) < \varepsilon$ (cf. Exercise 2.39) for all $n > N$. Hence, $\operatorname{diam}(F_n) \to 0$, so (F_n) is a contracting sequence. Let x be a point in X for which $\{x\} = \bigcap_{n=1}^{\infty} F_n$. For each $n \in \mathbf{N}$ the point x belongs to F_n. Therefore any ball centered at x has a nonempty intersection with E_n. Hence we may inductively select a strictly increasing sequence on natural numbers (n_k) such that for each k, $d(x, x_{n_k}) < 1/k$. Thus, $x_{n_k} \to x$. By Exercise 2.30, $x_n \to x$. It follows that X is compete. (This proof is found in [3, p. 195].)

2.42. Hint: Apply Theorem 2.11.

2.43. Let $\delta = \varepsilon$. By (2,1) on page 20, for $d(x,y) < \delta$,

$$|f(x) - f(y)| = |d(a,x) - d(a,y)| \leq d(x,y) < \delta = \varepsilon.$$

2.44. Let $\varepsilon > 0$ and $x, y \in X$. Without loss of generality, we may assume that $f(x) < f(y)$. There is $y_\varepsilon \in A$ such that $f(x) > d(x, y_\varepsilon) - \varepsilon$. Clearly, $d(y, y_\varepsilon) \geq f(y)$. We have

$$|f(y) - f(x)| < |d(y, y_\varepsilon) - d(x, y_\varepsilon) + \varepsilon| \leq |d(y, y_\varepsilon) - d(x, y_\varepsilon)| + \varepsilon \leq d(x,y) + \varepsilon.$$

Because ε is an arbitrary positive number, $|f(x) - f(y| \leq d(x, y)$, that is, f is uniformly continuous on X (cf. **2.43**).

2.45. (Necessity.) Let U be an open set and $f(x) = \mathrm{dist}(x, U^c)$ ($U^c = X \setminus U$). Clearly, $f(x) = 0$ on U^c. For $x \in U$ there is an open ball $B(x, r) \subseteq U$. For any $y \in U^c$, $d(x, y) \geq r$. Therefore, $f(x) = \inf\{d(x, y) : y \in U^c\} \geq r > 0$.

(Sufficiency.) Let $f : X \to \mathbf{R}$ be a continuous function X. Consider the set

$$U = \{x \in X : f(x) > 0\}$$

Let $\varepsilon = (1/2)f(a) > 0$ for $a \in U$. Inasmuch as f is continuous, there is $\delta > 0$ such that $|f(x) - f(a)| < (1/2)f(a)$ for $x \in B(a, \delta)$, which implies $0 < (1/2)f(a) < f(x)$. Hence, $B(a, \delta) \subseteq U$, so U is an open set.

2.46. Let (y_n) be a Cauchy sequence in $T(X)$. Inasmuch as T is an isometry, $(T^{-1}y_n)$ is a Cauchy sequence in X and therefore converges to some $x \in X$. Because T is continuous, the sequence (y_n) converges to $Tx \in T(X)$. Hence, $T(X)$ is a complete subspace of Y.

2.47. Hint. Every Cauchy sequence in a discrete space is eventually constant. Therefore a discrete space is complete.

2.48. Let T be an isometry from X onto Y. Note that T^{-1} is an isometry from Y onto X. Let (y_n) be a Cauchy sequence in Y. Define $x_n = T^{-1}y_n$. By **2.46**, (x_n) is Cauchy in X and therefore converges to some x. It is easy to verify that $(y_n) = (Tx_n)$ converges to Tx.

2.49. It suffices to show that a sequence (x_n) is Cauchy in (X, d) if and only if it is Cauchy in $(X, \widetilde{d})$. This follows from

$$d(x_m, x_n) < \varepsilon \qquad \text{if and only if} \qquad \widetilde{d}(x_m, x_n) < \widetilde{\varepsilon} = \frac{\varepsilon}{1 + \varepsilon}.$$

2.50. The mapping $f : \mathbf{R} \to (-\pi/2, \pi/2)$ given by $f(x) = \tan^{-1} x$ is an isometry. (We assume a given metric on $\mathbf{R}$ and the standard metric on the interval $(-\pi/2, \pi/2)$). The completion of $(-\pi/2, \pi/2)$ is $[-\pi/2, \pi/2]$. Hence, the completion of $\mathbf{R}$ with the "inverse tangent" metric is isometric to the closed interval $[-\pi/2, \pi/2]$.

2.51. The space X is isometric to the half-plane $\{(x, y) \in \mathbf{R}^2 : x < y\}$ with the distance function on $\mathbf{R}^2$ given by

$$d((x_1, y_1), (x_2, y_2)) = |x_1 - x_2| + |y_1 - y_2|.$$

The half-plane is an open proper subset of $\mathbf{R}^2$, hence incomplete. The completion is the closure of this half-plane $\{(x, y) \in \mathbf{R}^2 : x \leq y\}$.

2.52. Enumerate the rational numbers: $\mathbf{Q} = \{r_1, r_2 \ldots\}$. Consider $X = (\mathbf{Q}, d)$ with $d(x,y) = |x-y|$, and $X_n = \mathbf{Q} \setminus \{r_n\}$, $n \in \mathbf{N}$.

2.53. (Cf. [3, pp. 211–212].) Let $x_0 \in X$ and $r_0 > 0$. We must show that $B(x_0, r_0)$ contains a point of $\bigcap_{n=1}^{\infty} X_n$. The open set $B(x_0,r_0) \cap X_1$ is nonempty because X_1 is dense in X. For $x_1 \in B(x_0,r_0) \cap X_1$ there is $0 < r_1 < 1$ such that

$$\overline{B}_1 \subseteq B(x_0, r_0) \cap X_1, \tag{2.2}$$

where $B_1 = B(x_1, r_1)$. Suppose that for $n \in \mathbf{N}$ a descending collection of open balls $\{B_k\}_{k=1}^n$ has been chosen with the property that for $1 \leq k \leq n$, B_k has a radius less that $1/k$ and $\overline{B}_k \subseteq X_k$. The set $B_n \cap X_{n+1}$ is nonempty because X_{n+1} is dense in X. Choose $0 < r_{n+1} < 1/(n+1)$, for which, if we define $B_{n+1} = B(x_{n+1}, r_{n+1})$, $\overline{B}_{n+1} \subseteq B_n \cap X_{n+1}$. This inductively defines a contracting sequence of closed sets $(\overline{B}_n)$ such that $\overline{B}_n \subseteq X_n$ for each $n \in \mathbf{N}$. By the Cantor Intersection Property (Theorem 2.9), $\bigcap_{n=1}^{\infty} \overline{B}_n \neq \varnothing$. Let x belong to this intersection. Then $x \in \bigcap_{n=1}^{\infty} X_n$. On the other hand, by (2.2), $x \in B(x_0, r_0)$, which is the desired result.

2.54. Let $[a,b]$ be a closed interval in U. For each $n \in \mathbf{N}$ the set

$$f^{-1}(-\infty, n] \cap [a,b]$$

is a closed subset of the complete space $[a,b]$ and $[a,b]$ is the union of these closed sets. By Theorem 2.17, at least one of these sets contains an open interval. Claerly, the function f is bounded above over this interval.

The family $\mathcal{C} = \{f^{-1}(\alpha, \infty)\}_{\alpha \in \mathbf{R}}$ is an open covering of the interval $[a,b]$. Inasmuch as $[a,b]$ is a compact set, $\mathcal{C}$ contains a finite subcovering (cf. Theorem 2.19) of $[a,b]$. Hence the function f is bounded below on $[a,b]$.

2.55. Suppose that $([0,1] \setminus \mathbf{Q}) = \bigcup_{n=1}^{\infty} F_n$, where all sets F_n are closed. Because $\mathbf{Q}$ is countable, the interval $[0,1]$ is a countable union of singletons of rational numbers and sets F_n. By Theorem 2.17, at least one of the F_n's contains an open interval—a clear contradiction.

2.56. Follows immediately from the definition on p. 24.

2.57 (Necessity.) Suppose that $(X, \mathcal{T})$ is a T_1-space and let $y \in X$. For every $x \in X \setminus \{y\}$ there is $U \in \mathcal{T}$ containing x. Hence, $X \setminus \{y\}$ is an open set, so $\{y\}$ is a closed one.

(Sufficiency.) For $x \neq y$ in X, the set $X \setminus \{y\}$ is open and contains x.

2.58. It suffices to note that every open ball centered at x contains y.

2.59. Let $\{V_i\}_{i \in J}$ be a family of sets in $\mathcal{T}'$. There is a family $\{U_i\}_{i \in J}$ such that $V_i = U_i \cap X'$ for every $i \in J$. Inasmuch as $\bigcup_{i \in J} U_i \in \mathcal{T}$, we have

$$\bigcup_{i \in J} V_i = \bigcup_{i \in J} (U_i \cap X') = \Big(\bigcup_{i \in J} U_i\Big) \cap X' \in \mathcal{T}'.$$

A similar argument shows that finite intersections of sets in $\mathcal{T}'$ are sets in $\mathcal{T}'$. Clearly, $\varnothing \in \mathcal{T}'$ and $X' \in \mathcal{T}'$.

2.60. (Necessity.) Let $\mathcal{B}$ be a base of a topology on X. (a) Because X is an open set, X is the union of a subfamily of the family $\mathcal{B}$ and therefore is the union of the entire $\mathcal{B}$. (b) Because $x \in B_1 \cap B_2$, this intersection is a nonempty open set and hence is the union of a subfamily of $\mathcal{B}$. Clearly, there is an element B of the base $\mathcal{B}$ in this subfamily such that $x \in B \subseteq B_1 \cap B_2$.

(Sufficiency.) Let $\mathcal{T}$ be the set of unions of subfamilies of $\mathcal{B}$ together with the empty set $\varnothing$. We show that $\mathcal{T}$ is a topology on X. Clearly, $X \in \mathcal{T}$ and $\varnothing \in \mathcal{T}$. It is also clear that a union of a subfamily of $\mathcal{T}$ is a also union of subfamily of $\mathcal{B}$ and therefore belongs to $\mathcal{T}$. Suppose that $O_1, O_2 \in \mathcal{T}$. We need to show that $O_1 \cap O_2 \in \mathcal{T}$. We may assume that this intersection is not empty. Let $x \in O_1 \cap O_2$. Then there are B_1 and B_2 in $\mathcal{B}$ such that $x \in B_1 \subseteq O_1$ and $x \in B_2 \subseteq O_2$. By condition (b) of the theorem, there is $B_x \in \mathcal{B}$ with $x \in B_x \subseteq B_1 \cap B_2$. Then $O_1 \cap O_2 = \bigcup_{x \in O_1 \cap O_2} B_x$, the union of a subfamily of the family $\mathcal{B}$. Hence, $O_1 \cap O_2 \in \mathcal{T}$.

2.61. The spaces $X = \mathbf{R}$ and $Y = (-\pi/2, \pi/2)$ both with the usual metric, are homeomorphic. Indeed, the continuous function $f(x) = \tan^{-1} x$ is a homeomorphism $X \to Y$. The space Y is a bounded set, whereas X is not. Hence, the spaces are not isometric.

2.62. We need to verify conditions (a) and (b) of Theorem 2.22. Because X_1 and X_2 are open sets in $\mathcal{T}_1$ and $\mathcal{T}_2$, respectively, condition (a) holds. For condition (b) it suffices to note that

$$(U_1 \times U_2) \cap (V_1 \times V_2) = (U_1 \cap V_1) \times (U_2 \cap V_2),$$

where $U_1, V_1 \in \mathcal{T}_1$, $U_2, V_2 \in \mathcal{T}_2$.

2.63. (Necessity.) Let $T : X \to Y$ be a mapping. The inverse image of an open set in Y is a subset of X which is open. By Theorem 2.23, T is continuous.

(Sufficiency.) We may assume that the set X is not a singleton. Let Y be a discrete space on 2-point set $\{a, b\}$. Suppose that, for $E \subseteq X$, a mapping $T : X \to Y$ defined by

$$Tx = \begin{cases} a, & \text{if } x \in E, \\ b, & \text{otherwise} \end{cases}$$

is continuous. It follows that $E = T^{-1}\{a\}$ is an open set and the result follows.

3 Special Spaces

3.1. The operation of addition on a vector space X is a function of two variables $T : X \times X \to X$ defined by $T(x, y) = x + y$ for $x, y \in X$. If (X, d) is a metric space, then the product $X \times X$ can be made into a metric space by defining

$$\widehat{d}((x_1, y_1), (x_2, y_2)) = d(x_1, y_1) + d(x_2, y_2), \qquad x_1, x_2, y_1, y_2 \in X.$$

(The reader should verify that $\widehat{d}$ is, indeed, a metric on $X \times X$. This metric is known as ℓ_1 or "taxicab" metric on the Cartesian product $X \times X$. Compare it with the ℓ_1–norm introduced in Section 3.2.) Suppose that T is a continuous function in the sense of Definition 3.1 and let $(x_0, y_0) \in X \times X$. Then for every $\varepsilon > 0$ there is a $\delta > 0$ such that

$$B(x_0, \delta) + B(y_0, \delta) \subseteq B(x_0 + y_0, \varepsilon).$$

Equivalently,

$$d(x, x_0) < \delta \quad \text{and} \quad d(y, y_0) < \delta \quad \text{imply} \quad d(x + y, x_0 + y_0) < \varepsilon,$$

for all $(x, y) \in X \times X$. Because $\widehat{d}((x, y), (x_0, y_0)) = d(x, x_0) + d(y, y_0)$, it follows that

$$\widehat{d}((x, y), (x_0, y_0)) < \delta \quad \text{implies} \quad d(T(x, y), T(x_0, y_0)) < \varepsilon,$$

Therefore, T is continuous in the sense of Definition 2.9. The desired result follows from Theorem 2.10.

The scalar multiplication operation on a vector space X is also a function of two variables $S : \mathbf{F} \times X \to X$ defined by $S(\lambda, x) = \lambda x$ for $\lambda \in \mathbf{F}$, $x \in X$. For a metric space (X, d), the product $\mathbf{F} \times X$ is a metric space with the distance defined by

$$\widetilde{d}((\lambda_1, x_1), (\lambda_2, x_2)) = |\lambda_1 - \lambda_2| + d(x_1, x_2), \qquad \lambda_1, \lambda_2 \in \mathbf{F},\ x_1, x_2 \in X.$$

The reader is invited to apply the argument that we used above to obtain the required result. (Recall that $\mathbf{F}$ is $\mathbf{R}$ or $\mathbf{C}$.)

3.2. a) By substituting $x = 0$ into $\|\lambda x\| = |\lambda|\|x\|$, we obtain $\|0\| = 0$. In particular, $\|x\| = 0$ if and only if $x = 0$.

b) We have $0 = \|x - x\| \leq 2\|x\|$ for all $x \in X$. By part a), $\|x\| > 0$ for $x \neq 0$.

3.3. Because the scalar multiplication is continuous on a metric vector space, the mapping $T : \mathbf{F} \to X$ defined by $T(\lambda) = \lambda a$, where a is a nonzero vector in the space X, is continuous. By assuming that the metric on X is discrete, we obtain a contradiction because $T^{-1}(\{a\} = \{1\}$ which is not an open subset of the field $\mathbf{F}$.

3.4. Let $p = p_1 + p_2$ where p_1 and p_2 are seminorms. We have

$$p(\lambda x) = p_1(\lambda x) + p_2(\lambda x) = |\lambda|(p_1(x) + p_2(x)) = |\lambda| p(x),$$

and

$$\begin{aligned} p(x+y) = p_1(x+y) + p_2(x+y) &\le p_1(x) + p_1(y) + p_2(x) + p_2(y) \\ &= p(x) + p(y). \end{aligned}$$

3.5. It suffices to verify the triangle inequality. Let

$$x = \sum_{i\in J'} \lambda_i x_i \qquad \text{and} \qquad y = \sum_{i\in J''} \mu_i x_i,$$

where J' and J'' are finite subsets of J. We denote $J_1 = J' \setminus J''$, $J_2 = J'' \setminus J'$, and $J_3 = J' \cap J''$. Then

$$\begin{aligned} \|x+y\| &= \sum_{i\in J_1} |\lambda_i| + \sum_{i\in J_2} |\mu_i| + \sum_{i\in J_3} |\lambda_i + \mu_i| \\ &\le \sum_{i\in J_1} |\lambda_i| + \sum_{i\in J_2} |\mu_i| + \sum_{i\in J_3} |\lambda_i| + \sum_{i\in J_3} |\mu_i| = \sum_{i\in J'} |\lambda_i| + \sum_{i\in J''} |\mu_i| \\ &= \|x\| + \|y\|. \end{aligned}$$

3.6. By definition, $x_n \to x$ in X if $p(x_n - x) \to 0$. It follows from the triangle inequality for p that $|p(x_n) - p(x)| \le p(x_n - x)$. Hence, if $x_n \to x$, then $p(x_n) \to p(x)$. Every norm is a seminorm and therefore a continuous function.

3.7. For $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ we have

$$|x_i + y_i| \le |x_i| + |y_i| \le \|x\|_\infty + \|y\|_\infty, \qquad \text{for } 1 \le i \le n.$$

Therefore, $\|x+y\|_\infty = \max\{|x_i + y_i| : 1 \le i \le n\} \le \|x\|_\infty + \|y\|_\infty$.

3.8. Without loss of generality, we may assume that $\|x\|_\infty = |x_1|$. Then

$$\|x\|_\infty^p = |x_1|^p \le \sum_{k=1}^n |x_k|^p.$$

Hence, $\|x\|_\infty \le \|x\|_p$. On the other hand,

$$\|x\|_p = \Big(\sum_{k=1}^n |x_k|^p\Big)^{1/p} \le \big(n|x_1|^p\big)^{1/p} = n^{1/p}\|x\|_\infty.$$

In summary, $\|x\|_\infty \le \|x\|_p \le n^{1/p}\|x\|_\infty$. Because $1/p \to 0$ as $p \to \infty$, $n^{1/p} \to 1$. The result follows from the Squeeze Theorem.

3.9. Word-for-word repetition of the case $p = 2$. Proofs are found in many texts on functional analysis. See, for instance, [2, Section 1.5].

3.10. (a) (Necessity.) Let $T : X \to X$ be a homeomorphism from (X, d) onto (X, d') (cf. Definition 2.15). For a point $a \in X$, let $b = T(a)$. Let B_1 be an open ball centered at b in the space (X, d'). By continuity of T, there is an open ball B_2 centered at a in the space (X, d) such that $T(B_2) \subseteq B_1$. Inasmuch as T^{-1} is continuous, there is an open ball B_3 centered at b in the space (X, d') such that $T^{-1}(B_3) \subseteq B_2$. Thus $B_3 \subseteq T(B_2) \subseteq B_1$. By symmetry, we obtain the desired result.

(Sufficiency.) It suffices to note that, in this case, the identity mapping is a homeomorphism.

(b) Hint. Because the addition operation on a metric vector space is continuous, the translation mapping $T_a : X \to X$ defined by $x \mapsto x + a$ is a homeomorphism of (X, d) onto itself. Hence, by part a), every open ball in X contains a translation of an open ball centered at zero and vice versa.

3.11. Let $x \in A + B$, so $x = a + b$ for some $a \in A$, $b \in B$. Then $a \leq \sup A$, $b \leq \sup B$ and $x = a + b \leq \sup A + \sup B$. Hence, $\sup(A + B) \leq \sup A + \sup B$.

For $\varepsilon > 0$, there is $a \in A$ such that $\sup A - \varepsilon/2 < a$ and there is $b \in B$ such that $\sup B - \varepsilon/2 < b$. Therefore,

$$\sup A + \sup B < a + b + \varepsilon \leq \sup(A + B) + \varepsilon, \qquad \text{for every } \varepsilon > 0.$$

It follows that $\sup A + \sup B \leq \sup(A + B)$. Hence the desired equality.

3.12. It suffices to verify the triangle inequality. By the triangle inequality for $|\cdot|$ and Exercise 3.11,

$$\begin{aligned} \|x + y\|_\infty &= \sup\{|x_i + y_i| : i \in \mathbf{N}\} \leq \sup\{|x_i| + |y_i| : i \in \mathbf{N}\} \\ &= \sup\{|x_i| : i \in \mathbf{N}\} + \sup\{|y_i| : i \in \mathbf{N}\} = \|x\|_\infty + \|y\|_\infty. \end{aligned}$$

3.13 and 3.14. For ℓ_p^n with $p \in (0, \infty]$, see Exercise 3.9. For the spaces ℓ_p with $0 < p < 1$, the proofs are word-for-word taken from Theorems 3.5 and 3,7.

3.15. ***Correction.*** $0 < p < q$. Let $J = \{k \in \mathbf{N} : |x_k| \geq 1\}$. This set is finite because the series $\sum_{k=1}^\infty |x_k|^p$ converges. We have

$$\begin{aligned} \sum_{k \in \mathbf{N}} |x_k|^q &= \sum_{k \notin J} |x_k|^q + \sum_{k \in J} |x_k|^q \leq \sum_{k \notin J} |x_k|^p + \sum_{k \in J} |x_k|^q \\ &\leq \sum_{k \in \mathbf{N}} |x_k|^p + \sum_{k \in J} |x_k|^q < \infty. \end{aligned}$$

This proves the inclusion $\ell_p \subseteq \ell_q$. Note that $(n^{-1/p}) \in \ell_q$, whereas $(n^{-1/p}) \notin \ell_p$. Hence the inclusion is proper.

Clearly, $\ell_p \subseteq \ell_\infty$ for every $p > 0$. It suffices to consider a constant nonzero sequence to prove that the inclusion is proper.

Note that the paragraph on page 56 in the book, before Theorem 3.8, mistakingly refers to Exercise 3.22.

3.16. By Theorem 3.8, we may assume that $0 < p < q$. Let $0 < r < p$. For the sequence $x = (n^{-1/r})$, we have $\|x\|_q < \|x\|_p$. Therefore, ℓ_p is not a normed subspace of ℓ_q.

3.17. Let $x^{(n)}$ be a sequence of vectors in c that converges to $x = (x_n) \in \ell_\infty$. It suffices to show that $x \in c$ (cf. proof of Theorem 3.9).

For $\varepsilon > 0$, there is $N \in \mathbf{N}$ such that

$$\|x^{(n)} - x\|_\infty = \sup_{i \in \mathbf{N}} \big|x_i^{(n)} - x_i\big| < \varepsilon/3, \qquad \text{for all } n \geq N.$$

It follows that

$$\big|x_i^{(N)} - x_i\big| < \varepsilon/3, \qquad \text{for all } i \in \mathbf{N}.$$

Because $x^{(N)}$ is a vector in c, the sequence $\big(x_i^{(N)}\big)$ converges and therefore is Cauchy. Hence there is $N_1 \in \mathbf{N}$ such that

$$\big|x_i^{(N)} - x_j^{(N)}\big| < \varepsilon/3, \qquad \text{for all } i, j \geq N_1.$$

By the triangle inequality,

$$|x_i - x_j| \leq \big|x_i - x_j^{(N)}\big| + \big|x_j^{(N)} - x_i^{(N)}\big| + \big|x_i^{(N)} - x_j\big| < \varepsilon, \quad \text{for all } i, j \geq N_1.$$

It follows that the sequence (x_n) is Cauchy (in $\mathbf{F}$) and therefore converges. Hence, $x \in c$.

3.18. We consider the case of c_0 and assume that $\mathbf{F} = \mathbf{R}$.

Let $M = \{(x_n) \in c_0 : x_n \in \mathbf{Q}, \text{ for all } n \in \mathbf{N}\}$. This set is a countable subset of c_0. Let $\varepsilon > 0$ and $x \in c_0$. We need to show that there is $y \in M$ such that $\|x - y\|_\infty < \varepsilon$. Because $x \in c_0$, there is $N \in \mathbf{N}$ such that

$$\sup_{k > N} |x_k| < \varepsilon.$$

Because $\mathbf{Q}$ is dense in $\mathbf{R}$, we can choose rational numbers y_k for $1 \leq k \leq N$ such that

$$\sup_{1 \leq k \leq N} |x_k - y_k| < \varepsilon.$$

Then for $y = (y_1, \ldots, y_N, 0, 0, \ldots)$, we have

$$\|x - y\|_\infty = \sup_{k \in \mathbf{N}} |x_k - y_k| = \sup\Big\{\sup_{1 \leq k \leq N} |x_k - y_k|, \sup_{k > N} |x_k|\Big\} < \varepsilon.$$

Hint. For the space c consider the set of convergent sequences with rational terms.

3.19. ***Correction.*** The norm in $B(S)$ is defined by $\|x\| = \sup\{|x(t)| : t \in S\}$.

Let (x_n) be a Cauchy sequence in $B(S)$ and $\varepsilon > 0$. There exists $N \in \mathbf{N}$ such that

$$\|x_m - x_n\| = \sup_{t \in S}\{|x_m(t) - x_n(t)|\} < \varepsilon, \qquad \text{for all } m, n > N.$$

Hence for every $t \in S$,

$$|x_m(t) - x_n(t)| < \varepsilon, \qquad \text{for all } m, n > N. \tag{3.1}$$

Thus for every $t \in S$, the sequence $(x_n(t))$ is Cauchy in $\mathbf{F}$ and therefore convergent. Denote its limit by $x(t)$, which defines a function on S.

By taking $n \to \infty$ in (3.1), we obtain

$$|x_m(t) - x(t)| \leq \varepsilon, \qquad \text{for all } m > N,\, t \in S. \tag{3.2}$$

Because the function x_{N+1} is bounded, (3.2) implies $x \in B(S)$. By (3.2), $\|x_m - x\| \leq \varepsilon$ for $m > N$. Hence, $x_n \to x$ in $B(S)$, which is the desired result.

3.20. (a) Let $x = (x_n)$ be an element of the set bv. For $m < n$ we have

$$|x_m - x_n| = \left| \sum_{k=m}^{n-1} (x_k - x_{k+1}) \right| \leq \sum_{k=m}^{n-1} |x_k - x_{k+1}|$$

Because the series $\sum_{k=1}^{\infty} |x_k - x_{k+1}|$ converges, the sequence (x_n) is Cauchy and therefore converges. Clearly, $\lambda x \in bv$ for every $\lambda \in \mathbf{F}$ and $x \in bv$. For $x = (x_n)$ and $y = (y_n)$ in bv, we have

$$\begin{aligned} \|x + y\|_{bv} &= |x_1 + y_1| + \sum_{k=1}^{\infty} |x_{k+1} - x_k + y_{k+1} - y_k| \qquad (3.3) \\ &\leq |x_1| + |y_1| + \sum_{k=1}^{\infty} |x_{k+1} - x_k| + \sum_{k=1}^{\infty} |y_{k+1} - y_k| = \|x\|_{bv} + \|y\|_{bv}. \end{aligned}$$

It follows that bv is closed under scalar multiplication and addition. Properties defining a vector space are verified directly. The triangle inequality for $\|x\|_{bv}$ is established in (3.3). The other properties of the norm are clearly satisfied.

(b) Same steps as in part (a).

3.21. Clearly, the inclusions $c_{00} \subseteq \ell_p$ and $c_0 \subseteq c \subseteq \ell_\infty$ are proper and $\ell_p \subseteq c_0$. By Theorem 3.8, for $p < q$, the space ℓ_p is a proper subspace of ℓ_q. Hence it is a proper subspace of c_0.

(a) Let $(a_n) \in c_0$, so $a_n \to 0$. For $\varepsilon > 0$ there is $N \in \mathbf{N}$ such that $|a_n| < \varepsilon/2$ for all $n > N$. We define $x = (x_n) \in c_{00}$ by

$$x_n = \begin{cases} a_n, & \text{for } n \leq N, \\ 0, & \text{for } n > N. \end{cases}$$

Then

$$\|x - a\|_\infty = \sup\{|x_n - a_n| : n \in \mathbf{N}\} = \sup\{|a_n| : n \in \mathbf{N}\} < \varepsilon.$$

Therefore, c_{00} is dense in c_0.

(b) Let $(a_n) \in \ell_p$. Then $a_n \to 0$ and the above argument applies.

(c) The constant sequence $a = (1, 1, \dots)$ is in c. It is easy to verify that the open ball $B(a, 1/2)$ (in the sup-metric) does not contain vectors from c_{00}. Hence, c_{00} is not dense in c.

(d) Same example as in (c).

3.22. ***Correction.*** Should be "isometric to a metric subspace of ℓ_∞".

Let $A = \{x_1, x_2, \dots\}$ be a countable dense subset of the metric space (M, d), and let $T : x \mapsto (s_n^x)$ be a mapping from M into ℓ_∞ defined by

$$s_n^x = d(x, x_n) - d(x_n, x_1), \qquad \text{for } n \in \mathbf{N}.$$

We show that T is an isometry.

Let $a, b \in M$. The function $f_{a,b} : M \to \mathbf{R}$ given by

$$f_{a,b}(x) = |d(a, x) - d(b, x)|, \qquad x \in M,$$

is continuous on (M, d) (cf. Exercise 2.43) with $f_{a,b}(b) = d(a, b)$. Therefore for $\varepsilon > 0$ there is $\delta > 0$ such that

$$|d(a, x) - d(b, x)| > d(a, b) - \varepsilon, \quad \text{if } d(b, x) < \delta,$$

because, by the triangle inequality (cf. Exercise 2.2), $|d(a, x) - d(b, x)| \leq d(a, b)$. Since the set A is dense in (M, d), there is $x_n \in A$ such that $d(b, x_n) < \delta$. Then

$$d(a, b) - \varepsilon < |d(a, x_n) - d(b, x_n)| \leq d(a, b).$$

Because $|s_n^a - s_n^b| = |d(a, x_n) - d(b, x_n)|$, we have

$$d_\infty((s_n^a), (s_n^b)) = \sup\{|s_n^a - s_n^b| : n \in \mathbf{N}\} = d(a, b).$$

Hence, T is an isometry.

3.23. Condition (a) of Theorem 2.22 holds because X is the union of its open balls. For (b), it suffices to note that the intersection of two finite intersections of open balls is a finite intersection of open balls itself.

3.24. Let $x \neq y$ be two points in X and $p_j(x) \neq p_j(y)$ for some $j \in J$ (cf. Section 3.5). For $d = |p_j(x) - p_j(y)| > 0$ we have

$$B_j(x, d/2) \cap B_j(y, d/2) = \varnothing.$$

Indeed, if there is $z \in B_j(x, d/2) \cap B_j(y, d/2)$, then

$$p_j(z - x) < d/2, \quad p_j(z - y) < d/2$$

and

$$\begin{aligned} d = |p_j(x) - p_j(y)| &\leq p_j(x - y) = p_j(x - z + z - y) \\ &\leq p_j(z - x) + p_j(z - y) < d, \end{aligned}$$

a contradiction.

3.25. ***Counterexample!*** Let $X = \mathbf{R}$—the one-dimensional real vector space—with the distance function

$$d(x, y) = |\tan^{-1} x - \tan^{-1} y|, \quad x, y \in \mathbf{R}$$

(cf. Exercise 2.50) and the norm $\|x\| = |x|$, $x \in \mathbf{R}$. Because the inverse tangent function is continuous and strictly increasing on $\mathbf{R}$ the open balls in $(\mathbf{R}, d)$ are open intervals—the same as the open balls in $(\mathbf{R}, \|\cdot\|)$. Therefore the metric space $(\mathbf{R}, d)$ is normable. It can be easily seen that conditions (a) and (b) of the exercise are not satisfied.

Correction. The following is the correct version of the exercise:

Exercise 3.25. If the distance function in a metric vector space (X, d) is induced by a norm, then it satisfies conditions:

(a) $\quad d(x + z, y + z) = d(x, y),$

(b) $\quad d(\lambda x, \lambda y) = |\lambda|\, d(x, y),$

for all $x, y, z \in X$ and $\lambda \in \mathbf{F}$.

Solution. Let $\|\cdot\|$ be a norm on X for which $d(x, y) = \|x - y\|$ for $x, y \in X$. We have

(a) $d(x + z, y + z) = \|x + z - y - z\| = \|x - y\| = d(x, y)$.

(b) $d(\lambda x, \lambda y) = \|\lambda x - \lambda y\| = |\lambda|\, \|x - y\| = |\lambda|\, d(x, y)$.

3.26. Let $B(0, r) = \{x \in \boldsymbol{s} : d(x, 0) < r\}$ be an open ball of radius $r > 0$ in the space $\boldsymbol{s}$. Choose $n \in \mathbf{N}$ such that $1/2^n < r$ and let X be a vector subspace of $\boldsymbol{s}$ consisting of all sequences (x_k) with $x_k = 0$ for $k \leq n$. Clearly, the vector space X is isomorphic to the vector space $\boldsymbol{s}$. For $x = (x_n) \in X$ we have

$$d(x, 0) = \sum_{k=n+1}^{\infty} \frac{1}{2^k} \frac{|x_k|}{1 + |x_k|} < \frac{1}{2^{n+1}} + \frac{1}{2^{n+2}} + \cdots = \frac{1}{2^n} < r.$$

It follows that $X \subseteq B(0, r)$.

3.27. Define a bijection $\varphi : [a, b] \to [0, 1]$ by $\varphi(t) = (t - a)/(b - a)$ and $T : C[0, 1] \to C[a, b]$ by $(Tx)(t) = (x \circ \varphi)(t)$. Clearly, Tx is a continuous function on $[a, b]$. Suppose $Tx = Ty$ for some $x, y \in C[0, 1]$, so $T(x - y) = 0$. Then

$$x(\varphi(t)) - y(\varphi(t)) = 0, \qquad \text{for all } t \in [a, b].$$

Because φ is a bijection, $x = y$, hence, T is injective. For $y \in C[a, b]$ we have $y(\varphi^{-1}(t)) \in C[0, 1]$ and $T(y(\varphi^{-1}(t))) = y(t)$ so T is onto and hence bijective. Finally,

$$\|Tx\| = \sup_{t \in [a,b]} |x(\varphi(t))| = \sup_{s \in [0,1]} |x(s)| = \|x\|.$$

3.28. ***Correction.*** The norm on $C^{(n)}[0,1]$ (pages 65 and 73 in the book) is defined by

$$\|x\| = \sum_{k=0}^{n} \sup\{|x^{(k)}(t)| : t \in [0,1]\},$$

where $x^{(0)}(t) = x(t)$ on $[0,1]$.

Clearly, $\|\cdot\|$ is homogeneous and $\|x\| = 0$ implies that x is a zero function. (This is why it is important to have the sum starting at $k = 0$.) The triangle inequality follows from the triangle inequality for real and complex numbers and Exercise 3.11.

3.29. 1) Space $C(K)$. By Theorem 2.25, the function $\|\cdot\|$ in (3.11) on page 62 in the book is well-defined. As in Section 3.6, it is easy to verify that the space $C(K)$ endowed with the function $\|\cdot\|$ on it is a normed vector space in which convergence is uniform (cf. Theorem 3.17). The space $C(K)$ is Banach because, as in the case of $C[0,1]$, the limit of uniformly convergent sequence of continuous functions on K is a continuous function (Real Analysis theorem, cf. [4, Theorem 7.12]).

2) Space $C^{(n)}[0,1]$. Hint: Use the same steps as in the above proof and a theorem from Real Analysis on uniform convergence of sequences of differentiable functions (cf. [4, Theorem 7.17]).

3.30. ***Correction.*** The seminorms in (3.13) do not separate all distinct points in $C(0,1)$. Indeed, consider two distinct positive functions $x, y \in C(0,1)$ that attain the same maximum at $t = 1/2$. Clearly, $p_k(x) = p_k(y)$ for all $k \in \mathbf{N}$.

However, the statement holds if a weaker concept of a "separating family" is used. Namely, we say that the family $\{p_j\}_{j\in J}$ *separates points* in a vector space if, for $x \neq 0$, there is $j \in J$ such that $p_j(x) \neq 0$. Now, for every nonzero function x in $C(0,1)$ there is $k \in \mathbf{N}$ such that $x|_{I_k}$ is a nonzero function. Hence, $p_k(x) \neq 0$.

3.31. Suppose to the contrary that $f(x_0) > 0$ for some $x_0 \in [a,b]$. Because f is continuous, for $\varepsilon = f(x_0) > 0$ there is a $\delta > 0$ such that $f(x) > \varepsilon$ on $E = [x_0 - \delta, x_0 + \delta] \cap [a,b]$. We have

$$\int_{[a,b]} f \geq \int_E f \geq \int_E \varepsilon > 0,$$

a contradiction.

3.32. Clearly, the total variation of a constant function is zero. On the other hand, if $V_a^b(x) = 0$, then $V_a^b(x,P) = 0$ for every partition P of $[a,b]$. It suffices to consider partitions $P_1 = \{a,b\}$ and $P_2 = \{a,c,b\}$ to complete the solution.

4 Normed Spaces

4.1. Let $x_n = (0, \ldots, 0, 1/n^2, 0, \ldots)$, $n \in \mathbf{N}$, in c_{00}. Because

$$\sum_{k=1}^{\infty} \|x_k\| = \sum_{k=1}^{\infty} \frac{1}{n^2} < \infty,$$

the series $\sum_{k=1}^{\infty} x_k$ converges absolutely. Let $S_n = \sum_{k=1}^{n} x_k$ and suppose that $S_n \to a = (a_n) \in c_{00}$. Because $a \in c_{00}$, there is m such that $a_n = 0$ for all $n > m$. We have

$$\|S_n - a\|_\infty = \sup\{|1 - a_1|, \ldots, |\tfrac{1}{2^m} - a_m|, \tfrac{1}{2^{m+1}}, \ldots\} \geq \tfrac{1}{2^{m+1}},$$

for all $n > m$. This contradicts our assumption that $S_n \to a$. Hence, $\sum_{k=1}^{\infty} x_k$ does not converge in c_{00}.

4.2. (a) Let (x_n) be a Cauchy sequence in $(X, \|\cdot\|)$. Then, for every $\varepsilon_k = 2^{-k}$, $k \in \mathbf{N}$, there is $N_k \in \mathbf{N}$ such that $\|x_n - x_m\| < \varepsilon_k$ for $m, n > N_k$. Hence there are $n_{k+1} > n_k > N_k$ such that $\|x_{n_{k+1}} - x_{n_k}\| < 2^{-k}$. For $y_k = x_{n_k}$, $k \in \mathbf{N}$, the series $\sum_{k=1}^{\infty}(y_{k+1} - y_k)$ converges absolutely, because

$$\sum_{k=1}^{\infty} \|y_{k+1} - y_k\| < \sum_{k=1}^{\infty} \frac{1}{2^k} = 1.$$

By the assumption, the series $\sum_{k=1}^{\infty}(y_{k+1} - y_k)$ converges, so its sequence of nth partial sums $(y_n - y_1)$ converges, implying convergence of (y_n). Because (y_n) is a subsequence of the Cauchy sequence (x_n), the latter also converges. It follows that X is a Banach space.

(b) Let $S = \sum_{k=1}^{\infty} x_k$ be absolutely convergent series in a Banach space $(X, \|\cdot\|)$ and $S_n = \sum_{k=1}^{n} x_k$ for $n = 1, 2, \ldots$. Because S converges absolutely, for every $\varepsilon > 0$ there is $N \in \mathbf{N}$ such that

$$\sum_{k=m+1}^{n} \|x_k\| < \varepsilon, \qquad \text{for all } n > m > N.$$

We have

$$\|S_n - S_m\| = \left\| \sum_{k=m+1}^{n} x_k \right\| \leq \sum_{k=m+1}^{n} \|x_k\| < \varepsilon, \qquad \text{for all } n > m > N.$$

Hence, (S_n) is Cauchy and therefore converges.

4.3. For $x = (x_n) \in \ell_p$ consider the sequence (y_n) defined by

$$y_n = \sum_{k=1}^{n} x_k e_k = (x_1, \cdots, x_n, 0, 0 \ldots), \qquad \text{for } n \in \mathbf{N}.$$

Then

$$\|x - y_n\|_p = \|(0, \ldots, 0, x_{n+1}, x_{n+2}, \ldots)\|_p = \left(\sum_{k=n+1}^{\infty} |x_k|^p \right)^{1/p} \to 0,$$

because the series $\sum_{k=1}^{\infty} |x_k|^p$ converges.

For x in the space c_0 (or c_{00}), we have

$$\|x - y_n\|_\infty = \sup\{|x_k| : k > n\} \to 0,$$

because $x_n \to 0$.

Hint. For the uniqueness property of the above representations, consider coordinates of the zero vectors in these spaces.

4.4. Hint. Apply Theorem 2.2.

4.5. (Necessity.) Follows immediately from the definition.

(Sufficiency.) Let $S = \{z \in X : \|z\| = 1\}$ be the unit sphere in X. The set $T(S) = \{Tz : z \in S\}$ is bounded, so there is a constant C such that $\|Tz\| \leq C$ for all $z \in S$. For any nonzero vector x in X, we have $T(x/\|x\|) \leq C$. Therefore, $\|Tx\| \leq C\|x\|$ for all $x \neq 0$ in X. Hence, T is bounded.

4.6. (Necessity.) Theorem 4.4. (Sufficiency.) Theorem 4.3.

4.7. Clearly, T is a linear operator. Note that the functions $k(s,t)$ and $x(t)$ are continuous on their compact domains. The desired result follows from the following theorem in Analysis.

Theorem. If f is a continuous function on $[0,1] \times [0,1]$ and

$$F(x) = \int_0^1 f(x,y)\,dy, \qquad x \in [0,1],$$

then F is uniformly continuous on $[0,1]$.

Proof. Because f is continuous on the compact set $[0,1] \times [0,1]$, it is uniformly continuous on it. Hence for $\varepsilon > 0$ there is $\delta > 0$ such that

$$|f(x,y) - f(x',y)| < \varepsilon, \quad \text{when } y, x, x' \in [0,1] \text{ and } |x - x'| < \delta.$$

Therefore,

$$|F(x) - F(x')| = \left| \int_0^1 (f(x,y) - f(x',y))\,dy \right| \leq \int_0^1 |f(x,y) - f(x',y)|\,dy < \varepsilon,$$

if $|x - x'| < \delta$. Therefore, F is uniformly continuous on $[0,1]$

4.8. For every $x \in A$, we have $f(x) \leq \sup_{t \in A} f(t)$ and $g(x) \leq \sup_{t \in A} g(t)$. Therefore, $f(x) + g(x) \leq \sup_{t \in A} f(t) + \sup_{t \in A} g(t)$ for all $x \in A$. The result follows.

4.9. For $x = (x_n) \in \ell_p$ we have

$$\|T_l x\|_p = \left(\sum_{k=2}^{\infty} |x_k|^p \right)^{1/p} \leq \left(\sum_{k=1}^{\infty} |x_k|^p \right)^{1/p} = \|x\|_p.$$

Hence, T_l is bounded with $\|T_l\| \leq 1$. We obtain equality in the displayed formula for $x = (0, x_2, \ldots) \in \ell_p$. Therefore, $\|T_l\| = 1$.

A similar argument shows that $\|T_r\| = 1$.

4.10. We may assume that f is not zero. For $y \in \widetilde{X}$ there is a sequence (x_n) in X such that $x_n \to y$. Because f is a bounded linear functional on X, we have

$$|f(x_m) - f(x_n)| = |f(x_m - x_n)| \leq \|f\| \|x_m - x_n\|, \tag{4.1}$$

for $m, n \in \mathbf{N}$. Since convergent sequences are Cauchy, for $\varepsilon > 0$ there is $N \in \mathbf{N}$ such that $\|x_m - x_n\| < \varepsilon / \|f\|$ for all $m, n > N$. By (4.1), $(f(x_n))$ is a Cauchy sequence in $\mathbf{F}$ and therefore converges. We denote its limit by $\widetilde{f}(y)$.

To show that $\widetilde{f}(y)$ is well-defined, that is, its value does not depend on the choice of (x_n), suppose that another sequence, (x'_n), converges to y. We have

$$|f(x'_n) - f(x_n)| = |f(x'_n - x_n)| \leq \|f\| \|x'_n - x_n\| \leq \|f\| (\|x'_n - y\| + \|x_n - y\|) \to 0.$$

It follows that $\lim f(x'_n) = \lim f(x_n) = \widetilde{f}(y)$.

The functional $\widetilde{f}$ is linear. Indeed, for $x, y \in \widetilde{X}$ and $\alpha, \beta \in \mathbf{F}$, there are sequences (x_n) and (y_n) such that $x_n \to x$, $y_n \to y$. Clearly, the sequence $(\alpha x_n + \beta y_n)$ converges to $\alpha x + \beta y$. Therefore,

$$\widetilde{f}(\alpha x + \beta y) = \lim f(\alpha x_n + \beta y_n) = \alpha \lim f(x_n) + \beta \lim f(y_n) = \alpha \widetilde{f}(x) + \beta \widetilde{f}(y).$$

It is clear that $\widetilde{f}|_X = f$. For a sequence (x_n) converging to $y \in \widetilde{X}$, we have

$$|\widetilde{f}(y)| = \lim |f(x_n)| \leq \lim \|f\| \|x_n\| = \|f\| \|y\|.$$

Hence, $\widetilde{f}$ is a bounded linear functional on $\widetilde{X}$ with $\|\widetilde{f}\| \leq \|f\|$. On the other hand,

$$\|\widetilde{f}\| = \sup_{y \neq 0} \frac{|\widetilde{f}(y)|}{\|y\|} \geq \|f\|,$$

because $\widetilde{f}|_X = f$. Therefore, $\|\widetilde{f}\| = \|f\|$.

To show that $\widetilde{f}$ satisfying conditions of the exercise is unique, suppose that $\widehat{f}$ is another extension of f that satisfies the same conditions. For a sequence (x_n) with $x_n \to y \in \widetilde{X}$, we have

$$\widetilde{f}(y) = \lim \widetilde{f}(x_n) = \lim \widehat{f}(x_n) = \widehat{f}(y),$$

by continuity of functionals $\widetilde{f}$ and $\widehat{f}$.

4.11. For $x \in X$,

$$\|(TS)x\| = \|T(Sx)\| \leq \|T\|\|Sx\| \leq \|T\|\|S\|\|x\|.$$

Hence the result.

4.12. We have

$$\sup_{\|z\|<r} \|Tz\| = \sup_{0<\rho<r} \sup_{\|z\|=\rho} \|Tz\| = \sup_{0<\rho<r} \sup_{\|\frac{z}{\rho}\|=1} \left\| T\left(\frac{z}{\rho}\right) \right\| \rho = \|T\|\, r.$$

Furthermore, for vectors $x, z \in X$,

$$z = \frac{z+x}{2} + \frac{z-x}{2}.$$

Hence,

$$\|Tz\| \leq \frac{1}{2}(\|T(x+z)\| + \|T(x-z)\|) \leq \max\{\|T(x+z)\|, \|T(x-z)\|\}.$$

By taking the supremum over $z \in B(0, r)$, we obtain the desired result. (Note that vectors $x + z$ and $x - z$ belong to the ball $B(x, r)$.)

4.13. Hint. Apply Theorem 1.6 (Minkowski's Inequality).

4.14. We have

$$||x\|' = \sqrt{x_1^2/a^2 + x_2^2/b^2} = \frac{1}{a}\sqrt{x_1^2 + x_2^2 \cdot \frac{a^2}{b^2}} \geq \frac{1}{a}\sqrt{x_1^2 + x_2^2} = \frac{1}{a}\|x\|.$$

Similarly, $\|x\|' \leq \dfrac{1}{b}\|x\|$.

4.15. For $x = (x_1, \ldots, x_n) \in \ell_2^n$, the real function $f(x) = \sum_{k=1}^n |x_k|$ is continuous. By Exercise 2.45, the sets

$$A = \{x \in \ell_2^n : f(x) - 1 > 0\} \quad \text{and} \quad B = \{x \in \ell_2^n : 1 - f(x) > 0\}$$

are open. The set S is the complement of the open set $A \cup B$ and therefore is closed.

Note, that for $x \in S$, $|x_k| \leq 1$, $1 \leq k \leq n$. Hence,

$$\|x\|_2 = \Big(\sum_{k=1}^n |x_k|^2\Big)^{1/2} \leq \Big(\sum_{k=1}^n |x_k|\Big)^{1/2} = 1,$$

for every $x \in S$. Therefore, S is bounded.

4.16. Hint. Apply Theorem 4.5.

4.17. ***Correction.*** This is Theorem 3.4 in Chapter 3.

4.18. Let $\|\cdot\|$ and $\|\cdot\|'$ be equivalent norms on a vector space X, that is, for some $0 < a \leq b$,

$$a\|x\| \leq \|x\|' \leq b\|x\|, \qquad \text{for all } x \in X.$$

For a bounded linear operator (with respect to both norms) $T : X \to X$, we have ($x \neq 0$)

$$\|T\|' \geq \frac{\|Tx\|'}{\|x\|'} \geq \frac{a\|Tx\|}{b\|x\|}.$$

Hence, $\|T\|' \geq (a/b)\|T\|$. Similarly, $\|T\| \geq (a/b)\|T\|'$. We obtained the desired result:

$$\frac{a}{b}\|T\| \leq \|T\|' \leq \frac{b}{a}\|T\|.$$

4.19. ***Correction.*** X_0 is the null space, $\mathcal{N}(f)$, of a bounded nonzero linear functional f on X. Otherwise, the statement of the exercise does not hold. Counterexample: Consider in the real space ℓ_2^3 the subspace $X_0 = \{\lambda e_1 : \lambda \in \mathbf{R}\}$ and vectors $x = e_3$, $z = e_2$.

Solution. Note that $f(x) \neq 0$ and $f(z) \neq 0$. Let $y = \alpha x - z$, for $\alpha = f(z)/f(x)$. We have $y \in X_0$, because $f(y) = \alpha f(x) - f(z) = 0$.

4.20. For $x \neq 0$ in X, we have

$$\frac{|f(x)|}{\|x\|} = \frac{\left|\int_0^1 x(t)\,dt\right|}{\sup_{t\in[0,1]}|x(t)|} \leq \int_0^1 \frac{|x(t)|}{\sup_{t\in[0,1]}|x(t)|}\,dt \leq 1.$$

Therefore, $\|f\| \leq 1$.

Let (x_n) be a sequence of vectors in X defined by

$$x_n(t) = \begin{cases} nt, & \text{for } 0 \leq t \leq 1/n, \\ 1, & \text{for } 1/n \leq t \leq 1, \end{cases} \qquad \text{for } n \in \mathbf{N}.$$

Clearly, $\sup_{t\in[0,1]} x_n(t) = 1$ and $\int_0^1 x_n(t)\,dt = 1 - \frac{1}{2}\,n$ for $n \in \mathbf{N}$. Hence, $\|f\| = 1$.

4.21. Because the nonnegative function x on $[0,1]$ is continuous, there is $\delta > 0$ such that $x(t) < 1/2$ for $0 \leq t < \delta$. We may assume that $\delta < 1$. Since x is bounded by 1, we have

$$\int_0^1 x(t)\,dt = \int_0^\delta x(t)\,dt + \int_\delta^1 x(t)\,dt \leq \frac{1}{2}\,\delta + 1 - \delta < 1.$$

5 Linear Functionals

5.1. Let $\{e_1, \ldots, e_n\}$ be a basis of an n-dimensional normed space X and $x = c_1e_1 + \cdots + c_ne_n$ a vector in X. For a linear functional f on X, we have

$$|f(x)| = \Big|\sum_{k=1}^{n} c_k f(e_k)\Big| \le \max_k |f(e_k)| \sum_{k=1}^{n} |c_k| \le \frac{\max_k |f(e_k)|}{c} \, \|x\|,$$

where c is the constant from Lemma 4.1. Hence, f is bounded.

5.2. Clearly, each function $f_k : x \mapsto x_k$ is linear. For the boundedness, see Theorem 6.4.

5.3. The key idea: For a vector $x = (x_n) \in c$ and $x_0 = \lim x_n$, consider the vector $(x_n - x_0) \in c_0$, then apply steps similar to the proof of Theorem 5.2.

Let $x = (x_n)$ be a vector in the normed space c and f a bounded linear functional on c. We denote $x_0 = \lim x_n$, $b' = f(e)$, where $e = (1, 1, \ldots, 1, \ldots)$, and $b_n = f(e_n)$, for $n \in \mathbf{N}$. Here, (e_n) is the standard Schauder basis in c_0.

Step 1. We establish a formula for a bounded linear functional on c. The vector $x - x_0e$ belongs to the space c_0. Therefore,

$$x = x_0 e + \sum_{k=1}^{\infty} (x_k - x_0) e_k,$$

so

$$f(x) = x_0 b' + \sum_{k=1}^{\infty} (x_k - x_0) b_k. \tag{5.1}$$

Assuming that the series $\sum_{k=1}^{\infty} b_k$ converges absolutely we can rewrite (5.1) as

$$f(x) = x_0 b' + \sum_{k=1}^{\infty} x_k b_k - x_0 \sum_{k=1}^{\infty} b_k,$$

or, for $b_0 = b' - \sum_{k=1}^{\infty} b_k$, as

$$f(x) = x_0 b_0 + \sum_{k=1}^{\infty} x_k b_k, \tag{5.2}$$

where $b = (b_0, b_1, \ldots)$ is a vector in ℓ^1.

To show that the series $\sum_{k=1}^{\infty} b_k$, indeed, converges absolutely, we define a sequence $(\lambda^{(n)})$ in c by $\lambda^{(n)} = (\lambda_1, \ldots, \lambda_n, 0, 0 \ldots)$, where

$$\lambda_k = \begin{cases} 1, & \text{if } b_k = 0, \\ |b_k|/b_k, & \text{otherwise.} \end{cases}$$

(Note that this notation should be used in the proof of Theorem 5.2, instead of $x^{(n)}$.) We have

$$\sum_{k=1}^{n} |b_k| = f(\lambda^{(n)}) \leq \|f\| \|\lambda^{(n)}\| = \|f\|.$$

Hence, $\sum_{k=1}^{\infty} b_k$ converges absolutely.

Step 2. Let us show that every vector $d = (d_0, d_1, \ldots) \in \ell^1$ defines a bounded linear functional on c by (5.2):

$$g(x) = x_0 d_0 + \sum_{k=1}^{\infty} x_k d_k.$$

Clearly, g is a linear functional on c. We have

$$\sum_{k=1}^{n} |x_k||d_k| \leq \|x\| \sum_{k=1}^{n} |d_k| \leq \|x\| \sum_{k=1}^{\infty} |d_k|, \qquad \text{for } n \in \mathbf{N}.$$

Hence, the series $\sum_{k=1}^{\infty} x_k d_k$ converges absolutely. Furthermore,

$$|g(x)| \leq |x_0||d_0| + \sum_{k=1}^{n} |x_k||d_k| \leq \|x\| \Big(|d_0| + \sum_{k=1}^{\infty} |d_k| \Big)$$

It follows that g is a bounded linear functional on c.

By applying the above inequalities to the functional f in (5.2), we obtain

$$\|f\| \leq |b_0| + |b_1| + \cdots + |b_n| + \cdots = \|b\|_1, \tag{5.3}$$

the norm of $b = (b_0, b_1, \ldots)$ in the space ℓ_1. To prove that c^* is isomorphic to ℓ_1 we need to prove that, in fact, $\|f\| = \|b\|_1$.

Step 3. Let f be the functional in (5.2). We define a sequence $\mu^{(n)}$ of vectors in c by

$$\mu_k^{(n)} = \begin{cases} 1, & \text{if } b_k = 0, \\ |b_k|/b_k, & \text{otherwise}, \end{cases} \qquad \text{for } k \leq n,$$

and

$$\mu_k^{(n)} = \begin{cases} 1, & \text{if } b_0 = 0, \\ |b_0|/b_0, & \text{otherwise}, \end{cases} \qquad \text{for } k > n.$$

Because

$$\lim_{k \to \infty} \mu_k^{(n)} = \begin{cases} 0, & \text{if } b_0 = 0, \\ |b_0|/b_0, & \text{otherwise}, \end{cases}$$

we have

$$f(\mu^{(n)}) = |b_0| + \sum_{k=1}^{n} |b_k| + \sum_{k=n+1}^{\infty} b_k \mu_k^{(n)}.$$

Because $|f(\mu^{(n)})| \leq \|f\|\|\mu^{(n)}\| = \|f\|$, we have

$$f(\mu^{(n)}) = |b_0| + \sum_{k=1}^{n} |b_k| + \sum_{k=n+1}^{\infty} b_k \mu_k^{(n)} \leq \|f\|,$$

for every $n \in \mathbf{N}$. Therefore, $\|b\|_1 = |b_0| + |b_1| + \cdots \leq \|f\|$. By (5.3), $\|f\| = \|b\|_1$.

5.4. Let f be a linear functional on $\boldsymbol{s}$ defined by (5.8), that is, $f(x) = \sum_{k=1}^{\infty} x_k b_k$ for $x = (x_n) \in \boldsymbol{s}$ and $(b_n) \in c_{00}$. Clearly, f is a finite (!) sum of linear functionals $f_m(x) = b_m x_m$, $m \in \mathbf{N}$. Inasmuch as convergence in $\boldsymbol{s}$ is coordinatewise (cf. page 61 in the text), the functionals f_m's are continuous. Hence, f is continuous.

5.5. From the forth displayed inequality on page 108, we obtain

$$\sup_{y' \in Y} [-p(x_0 - y') + f(y')] \leq \inf_{y'' \in Y} [p(x_0 - y'') + f(y'')].$$

Therefore, c could be either of these two numbers (or any number in between).

5.6. Let $Y = \operatorname{span}\{x, y\}$. This is a two- or one-dimensional subspace of X. It is not difficult to construct a linear functional f' on Y such that $f'(x) \neq f(y)$. If $Y = X$, then we are done. Otherwise, define $f = \widetilde{f'}$, the extension of f'.

5.7. Let $X = c_0$, $Y = c_{00}$, and $x_0 = (1, 1, \ldots, 1, \ldots) \in X \setminus Y$. Suppose to the contrary that there is a bounded linear functional f on X such that $f(x) = 0$ for $x \in Y$ and $f(x_0) \neq 0$. By Theorem 5.2, $f(x) = \sum_{k=1}^{\infty} x_k b_k$ for $(x_n) \in X$, $(b_n) \in \ell_1$. Because f is zero function on Y, all partial sums of $\sum_{k=1}^{\infty} b_k$ are zero. This contradicts our assumption that $f(x_0) = \sum_{k=1}^{\infty} b_k \neq 0$.

5.8. We have an equivalent inequality:

$$y^2(\beta^2 - 1) + 2\beta xy \leq 0, \qquad \text{for all } (x, y) \in \mathbf{R}^2.$$

Clearly, $\beta \neq 1$. If $\beta \neq 0$ and $x \neq 0$, then the discriminant $4\beta^2 x^2$ of the quadratic function in y on the left hand side is positive, so the above inequality cannot hold for all real x and y. Hence, $\beta = 0$.

5.9. By Corollary 5.2,

$$\sup \left\{ \frac{|f(x)|}{\|f\|} : f \in X^*,\ f \neq 0 \right\} \geq \|x\|.$$

Because $|f(x)| \leq \|f\|\|x\|$,

$$\sup \left\{ \frac{|f(x)|}{\|f\|} : f \in X^*,\ f \neq 0 \right\} \leq \|x\|.$$

Hence the result.

5.10. Hints. For the triangle inequality, use Lemma 1.1. To show that (X, d) is a metric space, follow the steps in the proof of Theorem 3.2.

5.11. For $u, v \in Y + x$, we have $u = y + x$, $v = y' + x$, where $y, y' \in Y$. For $\lambda \in [0, 1]$,

$$\lambda u + (1 - \lambda)v = \lambda(y + x) + (1 - \lambda)(y' + x) = \lambda y + (1 - \lambda)y' + x \in Y + x,$$

because $\lambda y + (1 - \lambda)y' \in Y$.

5.12. Let $b = (b_n) \in \ell_\infty$, $x = (x_n) \in \ell_p$ for $0 < p < 1$, and

$$f(x) = \sum_{k=1}^{\infty} b_k x_k.$$

We may assume that $b \neq 0$. By Exercise 3.15, $\sum_{k=1}^{\infty} |x_k| < \infty$, because $\sum_{k=1}^{\infty} |x_k|^p < \infty$. Therefore,

$$|f(x)| = \left| \sum_{k=1}^{\infty} b_k x_k \right| \leq \|b\|_\infty \|x\|_1.$$

For $\varepsilon > 0$, let $\delta = \min\left\{ \frac{\varepsilon}{\|b\|_\infty}, 1 \right\}$. If $d(x, y) < \delta$, then

$$\|x - y\|_1 = \sum_{k=1}^{\infty} |x_k - y_k| \leq \sum_{k=1}^{\infty} |x_k - y_k|^p = d(x, y) < \delta,$$

and we have

$$|f(x) - f(y)| = |f(x - y)| \leq \|b\|_\infty \|x - y\|_1 < \varepsilon.$$

Hence the functional f is uniformly continuous on ℓ_p.

5.13. Clearly, p is positive homogeneous. Let (x_n) and (y_n) be elements of ℓ_∞. For each $i \geq n$, we have

$$x_i + y_i \leq \sup_{k \geq n} x_k + \sup_{k \geq n} y_k.$$

Hence,

$$\sup_{k \geq n} (x_k + y_k) \leq \sup_{k \geq n} x_k + \sup_{k \geq n} y_k.$$

We prove subadditivity of p by taking limits as $n \to \infty$ in the above inequality.

5.14. By positive homogeneity, $p(0) = p(0 \cdot x) = 0p(x) = 0$. By subadditivity,

$$0 = p(0) = p(x + (-x)) \leq p(x) + p(-x).$$

Therefore, $p(-x) \geq -p(x)$.

5.15. Let $x_n \to x$ in X. We have

$$p(x) = p(x - x_n + x_n) \leq p(x - x_n) + p(x_n),$$

so

$$p(x) - p(x_n) \leq p(x - x_n).$$

By exchanging x and x_n, we obtain

$$p(x_n) - p(x) \leq p(x_n - x),$$

or, equivalently,

$$-p(x_n - x) \leq p(x) - p(x_n).$$

In summary,

$$-p(x_n - x) \leq p(x) - p(x_n) \leq p(x - x_n).$$

Because p is continuous at zero, $p(0) = 0$, and $x_n \to x$, we have $p(x_n) \to p(x)$. Hence, p is continuous on X.

5.16. ***Correction.*** Cf. [2, Section 4.2, Problem 7]. "Linear combination" should be "linear combination with nonnegative coefficients".

Straightforward verification of the defining properties.

5.17. Hint. By Exercise 5.9, $x = 0$ if $f(x) = 0$ for all $f \in X^*$.

5.18. ***Correction.*** $\sum_{i=1}^k \lambda_i = 1$ not 0.

(Necessity.) Let $\{x_1, \dots, x_k\}$ be a finite subset of $S = Y + a$. Then each $x_i = y_i + a$ where $y_i \in Y$. We have

$$\sum_{i=1}^k \lambda_i x_i = \sum_{i=1}^k \lambda_i (y_i + a) = \sum_{i=1}^k \lambda_i y_i + \Big(\sum_{i=1}^k \lambda_i\Big) a = \sum_{i=1}^k \lambda_i y_i + a.$$

Hence, $\sum_{i=1}^k \lambda_i x_i \in S$, because $\sum_{i=1}^k \lambda_i y_i \in Y$.

(Sufficiency.) Let $Y = S - a$ for some $a \in S$. Clearly, $0 = a - a \in Y$. For $x, y \in Y$ there are $x', y' \in S$ such that $x = x' - a$, $y = y' - a$. Because $x', y', a \in S$, for $\lambda, \mu \in \mathbf{F}$, we have

$$\lambda x' + \mu y' + (1 - \lambda - \mu) a \in S.$$

Hence, $\lambda x + \mu y \in Y$. It follows that Y is a subspace.

5.19. (Necessity.) Let $S = Z + a$, where Z is a subspace of X, and

$$\pi : X \to X/Z$$

be the quotient map (cf. [1, Section 3.E]). We have $S = \pi^{-1}(\{S\})$, because $S \in X/Z$.

(Sufficiency.) Let $T : X \to Z$ ($Z \neq \{0\}$) be a linear map. For a nonzero vector $z_0 \in Z$, let $S = T^{-1}(\{z_0\}) = \{x \in X : Tx = z_0\}$. For $a \in S$, we have $S = \mathcal{N}(T) + a$, so S is an affine set. ($\mathcal{N}(T) = T^{-1}(\{0\})$ is the null space of T.)

5.20. (Necessity.) Let S be a hyperplane in X. There is a subspace Y of X and a vector $a \in X \setminus Y$, such that $S = Y + a$ and $\operatorname{span}\{Y, a\} = X$. We have $X = \{x \in X : x = y + \lambda a,\ y \in Y,\ \lambda \in \mathbf{R}\}$. Let $f(x) = f(y + \lambda a) = \lambda$ on X. Clearly, f is a linear functional and $S = \{x \in X : f(x) = 1\}$.

(Sufficiency.) For a nonzero functional f on X and $c \in \mathbf{R}$, let

$$S = \{x \in X : f(x) = c\} \qquad \text{and} \qquad Y = \{x \in X : f(x) = 0\}.$$

Clearly, $Y + a \subseteq S$, for $a \in S$. On the other hand, for $x \in S$, we have

$$c = f(x) = f((x - a) + a) = f(x - a) + c.$$

Therefore, $x - a \in Y$, which implies that $S \subseteq Y + a$. Thus, $S = Y + a$, that is, S is a hyperplane.

5.21. A one-dimensional subspace of ℓ_2^2.

5.22. Hint. Suppose that $0 \notin U$ and let $a \in U$. The translation $x \mapsto x - a$ maps convex sets to convex sets and hyperplanes to hyperplanes.

5.23. Let $T : \mathbf{F} \to X$ be defined by $\lambda \mapsto \lambda x_0$. Because multiplication by scalar is a continuous operation, a neighborhood U of point λx_0 contains the product of neighborhoods U_λ and U_{x_0} of $\lambda \in \mathbf{F}$ and $x_0 \in X$:

$$U_\lambda U_{x_0} \subseteq U.$$

Hence, $T^{-1}(U)$ contains U_λ. Therefore, T is continuous.

5.24. (a) Let X and Y be vector spaces, $T : X \to Y$ a linear mapping , and S a convex subset of X. For $x, y \in S$ and $\lambda \in [0, 1]$, we have

$$\lambda Tx + (1 - \lambda)Ty = T(\lambda x + (1 - \lambda)y) \in T(S).$$

Hence. $T(x)$ is a convex subset of Y.

(b) Let S be a convex subset of $\mathbf{R}$. We assume that S is bounded and denote $a = \inf S$, $b = \sup S$. We also assume that $a < b$. Let $x \in S \setminus \{a, b\}$. By the definitions of infimum and supremum, there are x_1 and x_2 in S such that $a < x_1 \leq x \leq x_2 < b$. Hence, every $x \in S$ belongs to (a, b). It follows that S is one of the intervals (a, b), $[a, b)$, $(a, b]$, or $[a, b]$. The case of an unbounded set S is treated similarly.

5.25. Let $S_x = \{t > 0 : x \in tU\}$, so $p(x) = \inf S_x$.

(a) If $p(x) = 0$, then, clearly, $x \in tU$ for every $t > 0$. Conversely, if $x \in tU$ for every $t > 0$, then $p(x) = \inf S_x = 0$.

(b) For $x \in U$, $1 \in S_x$, so $p(x) = \inf S_x \leq 1$. Similarly, if $x \notin U$, then $1 \notin S_x$, and $p(x) \geq 1$.

5.26. Hint. Apply Exercises 5.24 (a) and (b).

5.27. ***Correction.*** A "vector space" is a "topological vector space".

That the set $V - W$ is convex follows from the following identity:

$$\lambda(v - w) + (1 - \lambda)(v' - w') = [\lambda v + (1 - \lambda)v'] - [\lambda w + (1 - \lambda)w'],$$

where $\lambda \in [0, 1]$, $v, v' \in V$, and $w, w' \in W$. For every $w \in W$, the set $V - w$ is open because operations of addition and scalar multiplication are continuous. It remains to note that $V - W = \bigcup_{w \in W}(V - w)$.

5.28. ***Corrections.*** Instead of $x_0 = v_0 - w_0$, we must have $x_0 = w_0 - v_0$ in the exercise and in the proof of Theorem 5.13. In the last sentence, replace $x_0 \in U$ with $0 \in U$.

The statement follows from Exercise 5.27. Also, $0 = w_0 - v_0 + x_0 \in U$.

5.29. ***Clarification.*** The "leading nonzero term" is the last(!) nonzero term.

Let $(x_n), (y_n) \in M$. Clearly, $(\lambda x_n + (1 - \lambda)y_n) \in M$ for $\lambda \in [0, 1]$. Thus, M and $-M$ are convex sets. Obviously, they are disjoint. Let (e_n) be the Schauder basis in c_{00}. Suppose that there is a nonzero linear functional

$$f(x) = \sum_{k=1}^{\infty} c_k x_k, \quad \text{where } c_k = f(e_k),\ k \in \mathbf{N},$$

such that $M = \{x \in c_{00} : f(x) \geq c\}$ and $-M = \{x \in c_{00} : f(x) \leq c\}$ for some $c \in \mathbf{R}$. We may assume that $f(e_1) \neq 0$. We have

$$f(\varepsilon e_1) = \varepsilon f(e_1) \geq c \quad \text{and} \quad f(-\varepsilon e_1) = -\varepsilon f(e_1) \leq c, \quad \text{for every } \varepsilon > 0,$$

because $\varepsilon e_1 \in M$ and $-\varepsilon e_1 \in -M$. It follows that $c = 0$.

Let $x = (x_1, \ldots, x_n, 0, \ldots) \in M$, so $x_n > 0$, and $y = e_{n+1} \in M$. We have

$$f(x + \varepsilon y) = f(x) + \varepsilon f(y) \geq 0 \quad \text{and} \quad f(-x + \varepsilon y) = -f(x) + \varepsilon f(y) \geq 0,$$

for every $\varepsilon > 0$, because vectors $x + \varepsilon y$ and $-x + \varepsilon y$ belong to M. It follows that $f(x) = 0$ for all $x \in M$. Every vector in $-M$ is a negative of a vector in M. Hence, $f(x) = 0$ for all $x \in -M$. Thus, f is the zero functional, a contradiction.

5.30. a) Clearly, $|\delta(x)| = |x(0)| \leq \sup_{t \in [0,1]} |x(t)| = \|x\|$. Hence, $\|\delta\| \leq 1$. For $x_0(t) \equiv 1$, $\|x_0\| = 1$ and $\delta(x_0) = 1$. Therefore, $\|\delta\| = 1$.

b) We define, for $n \in \mathbf{N}$,

$$x_n(t) = \begin{cases} (-n^2/2)t + n, & \text{for } t \in [0, 2/n], \\ 0, & \text{for } t \in (2/n, 1]. \end{cases}$$

It is easy to verify that $\|x_n\| = \int_0^1 x_n(t)\,dt = 1$. Because $\delta(x_n) = n$, the functional δ is unbounded.

5.31. Let X, X^*, X^{**}, and $X^{***} = (X^{**})^* = (X^*)^{**}$ be a normed space X, its dual, its second dual, and its third dual spaces, respectively. We denote by $J : X \to X^{**}$ and $J' : X^* \to X^{***}$ the respective canonical mappings. Thus, by definition,

$$J(x)(f) = f(x), \qquad \text{for } x \in X,\ f \in X^*, \tag{5.4}$$

and

$$J'(f)(g) = g(f), \qquad \text{for } f \in X^*,\ g \in X^{**}. \tag{5.5}$$

We need to show that J' is onto, that is, for every $F \in X^{***}$ there is $f \in X^*$ such that $J'(f) = F$. Because X is reflexive, for every $g \in X^{**}$ there is a unique $x_g \in X$, such that $g = J(x_g)$. Note that $F \circ J$ is a linear functional on X, $F \circ J \in X^*$. We have, by (5.5) and (5.4),

$$J'(F \circ J)(g) = g(F \circ J) = J(x_g)(F \circ J) = (F \circ J)(x_g) = F(g),$$

for every $g \in X^{**}$. Hence, $F = J'(F \circ J)$, so J' is onto.

5.32. Below, Y is a closed subspace of a reflexive Banach space X.

Let $J : X \to X^{**}$ and $J' : Y \to Y^{**}$ be canonical mappings. We need to show that for every $y^{**} \in Y^{**}$ there is $y \in Y$ such that $J'(y) = y^{**}$, that is, (cf. (5.4)) $y^{**}(y^*) = y^*(y)$ for all $y^* \in Y^*$. We define $x^{**} \in X^{**}$ by $x^{**}(x^*) = y^{**}(x^*|_Y)$ for $x^* \in X^*$. Because X is a reflexive space, there is $x \in X$ such that $J(x) = x^{**}$. Hence,

$$y^{**}(x^*|_Y) = x^{**}(x^*) = x^*(x), \qquad \text{for all } x^* \in X^*.$$

It remains to show that $x \in Y$. Assume to the contrary that $x \notin Y$. By Corollary 5.4, there is an $x^* \in X^*$ which is zero on Y and $x^*(x) \neq 0$. For this x^* we have from the above formula

$$0 = y^{**}(x^*|_Y) = x^{**}(x^*) = x^*(x) \neq 0,$$

a contradiction.

5.33. By Exercise 5.31, it suffices to show that reflexivity of X^* implies reflexivity of X. We use notations from Exercise 5.31.

Suppose to the contrary that $J(X)$ is a proper subspace of X^{**}. Inasmuch as X is a Banach space and J is an isometry, $J(X)$ is a closed subspace of X^{**}. By Corollary 5.4, there is a nonzero $F \in X^{***}$ which is zero on $J(X)$. Because X^* is reflexive, there is a nonzero $f \in X^*$ such that $J'(f) = F$. By (5.4) and (5.5), for every $x \in X$,

$$0 = F(J(x)) = J'(f)(J(x)) = J(x)(f) = f(x).$$

Hence, f is the zero functional, a contradiction.

5.34. Let $\{x_1, \dots, x_n\}$ be a set of n linearly independent vectors in a normed space X and $Y = \operatorname{span}\{x_1, \dots, x_n\}$. By Exercise 5.1, linear functionals f_k, $1 \le k \le n$, defined by $f_k(x) = \lambda_k$ for $x = \sum_{k=1}^n \lambda_k x_k$ are bounded on Y. Suppose that $\sum_{k=1}^n \mu_k f_k = 0$. For $x = x_k$, we have $f_k(x) = 1$ and $f_i(x) = 0$ for $i \neq k$. It follows that $\mu_k = 0$ for $1 \le k \le n$, so the set $\{f_k\}_{1 \le k \le n}$ is linearly independent in Y^*. It is not difficult to see that the set $\{\widetilde{f}_k\}_{1 \le k \le n}$ is linearly independent in X^*.

6 Fundamental Theorems

6.1. Zero terms of the sequence (a_n) do not affect the convergence properties of the respective series.

6.2. For $m > n$, $\|x^{(m)} - x^{(n)}\| = 1/(n+1)$. Hence, $(x^{(n)})$ is Cauchy in c_{00}. For $x = (x_1, \ldots, x_n, 0, 0, \ldots) \in c_{00}$, we have

$$\|x^{(m)} - x\| = \sup\{|x_1 - 1|, |x_2 - 1/2|, \ldots, |x_n - 1/n|, 1/(n+1), \ldots\} \geq \frac{1}{n+1},$$

for all $m > n$. Hence, $(x^{(n)})$ does not converge to x in c_{00}.

6.3. Suppose that X is a proper dense Banach subspace of a Banach space Y and let $x \in Y \setminus X$. Because X is dense in Y, there is a sequence (x_n) in X that converges to x. Because X is Banach, $x \in X$, a contradiction.

6.4. Clearly, T is a linear operator. The sequence $(T_n x)$ converges for every $x \in X$. Therefore, the sequence $(\|T_n x\|)$ is bounded (cf. Theorem 2.7). By Theorem 6.1, there is M such that $\|T_n\| < M$ for all $n \in \mathbf{N}$. Because $T_n x \to Tx$ for all $x \in X$, the sequence $\|T_n x - Tx\|$ is bounded for every $x \in X$. By Theorem 6.1, there is M' such that $\|T_n - T\| < M'$. We have,

$$\begin{aligned}\|Tx\| &= \|T_n x + Tx - T_n x\| \leq (\|T_n\| + \|T - T_n\|)\|x\| \\ &\leq (M + M')\|x\|,\end{aligned}$$

for all $x \in X$. It follows that T is bounded.

6.5. Follows immediately from Theorem 6.1.

6.6. By Theorem 2.7, the sequence $(T_n x)$ is bounded for every $x \in X$. By Theorem 6.1, the sequence $(\|T_n\|)$ is bounded.

6.7. (a)$\Rightarrow$(b). Follows from $\|T_n x\| = \|T_n\| ||x||$.

(b)$\Rightarrow$(a). Theorem 6.1.

(b)$\Rightarrow$(c). Follows from $|f(T_n x)| \leq \|f\| \|T_n x\|$.

(c)$\Rightarrow$(b). Suppose to the contrary that, for some $x_0 \in X$, the sequence $(\|T_n x_0\|)$ is unbounded in Y. By Corollary 5.2, for each $n \in \mathbf{N}$ there is a bounded linear functional f_n on Y such that $f_n(T_n x_0) = \|T_n x_0\|$. This contradicts (c).

6.8. Follows immediately from Theorem 6.2.

6.9. (a) For $x, y \in B(x_0, r)$ and $\lambda \in [0, 1]$, we have

$$\begin{aligned}\|\lambda x + (1-\lambda)y - x_0\| &= \|\lambda x + (1-\lambda)y - (\lambda + (1-\lambda))x_0\| \\ &= \|\lambda(x - x_0) + (1-\lambda)(y - x_0)\| \\ &\leq \lambda\|x - x_0\| + (1-\lambda)\|y - y_0\| < r.\end{aligned}$$

Hence, $B(x_0, r)$ is a convex set.

(b) We have the chain of equivalences:

$$\begin{aligned} x \in r^{-1}(B(x_0,r) - x_0) &\Leftrightarrow rx + x_0 \in B(r, x_0) \\ &\Leftrightarrow \|rx + x_0 - x_0\| < r \Leftrightarrow \|x\| < 1 \Leftrightarrow x \in B(0,1). \end{aligned}$$

Hence the result.

6.10. It suffices to show that, for each $x_0 \in X$ and $\lambda_0 \in \mathbf{F}$, $\lambda_0 \neq 0$, the mapping $T : x \mapsto \lambda_0 x + x_0$ is a homeomorphism of X onto itself.

Clearly, T is onto and, by continuity of vector addition and scalar multiplication, continuous with continuous inverse $T^{-1} : x \mapsto \lambda_0^{-1}(x - x_0)$.

6.11. (Necessity.) Suppose that T^{-1} is bounded and let (Tx_n) be a convergent sequence in $T(X)$. Because T^{-1} is continuous the sequence (x_n) converges to some $x \in X$. Since T is continuous, $Tx_n \to Tx \in T(X)$. Hence, $T(X)$ is closed in Y.

(Sufficiency.) Let $T(X)$ be a closed subset of Y. By Theorem 2.8, $T(X)$ is a Banach subspace of Y. The result follows from Theorem 6.3.

6.12. (a) We need to verify the triangle inequality only. We have

$$\begin{aligned} \|(x,y) + (u,v)\| &= \|(x+u, y+v)\| = \max\{\|x+u\|_1, \|y+v\|_2\} \\ &\leq \max\{\|x\|_1 + \|u\|_1, \|y\|_2 + \|v\|_2\} \\ &\leq \max\{\|x\|_1, \|y\|_2\} + \max\{\|u\|_1, \|v\|_2\} = \|(x,y)\| + \|(u,v)\|, \end{aligned}$$

because, clearly,

$$\|x\|_1 + \|u\|_1 \leq \max\{\|x\|_1, \|y\|_2\} + \max\{\|u\|_1, \|v\|_2\}$$

and

$$\|y\|_2 + \|v\|_2 \leq \max\{\|x\|_1, \|y\|_2\} + \max\{\|u\|_1, \|v\|_2\}.$$

Hint. To prove that the topology defined by the max-norm (6.2) coincides with the product topology, show that

$$B((x_0, y_0), r) = B(x_0, r) \times B(y_0, r).$$

(b) Let $((x_n, y_n))$ be a Cauchy sequence in $X \times Y$, that is, for $\varepsilon > 0$ there is $N \in \mathbf{N}$ such that

$$\|(x_n, y_n) - (x_m, y_m)\| < \varepsilon, \qquad \text{for all } m, n > N.$$

From this, we immediately obtain

$$\|x_n - x_m\|_1 < \varepsilon \quad \text{and} \quad \|y_n - y_m\|_2 < \varepsilon \qquad \text{for all } m, n > N.$$

Thus, (x_n) and (y_n) are Cauchy in spaces X and Y, respectively, and therefore, $x_n \to x$ and $y_n \to y$ for some $x \in X$, $y \in Y$. It is readily verified that $(x_n, y_n) \to (x, y)$ in $X \times Y$. Hence, $X \times Y$ is Banach.

(c) The graph $G(T)$ is a closed subspace of the Banach space $X \times Y$ (cf. (b)) and therefore is a Banach space.

6.13. Suppose that $x_n \to x$ in X and $Tx_n \to y$ in Y. Then the sequence (x_n, Tx_n) converges to (x, y). Indeed, we have

$$\|(x_n, Tx_n) - (x, y)\| = \|(x_n - x, Tx_n - y)\| = \max\{\|x_n - x\|_1, \|Tx_n - y\|_2\} \to 0.$$

Inasmuch as $y = Tx$, $(x, Tx) \in G(T)$. Hence, $G(T)$ is closed in $X \times Y$. By Theorem 6.5, T is bounded.

6.14. (a) Let C be a compact subset of X and $K = T(C) \subseteq Y$. We need to show that the limit y of a convergent sequence (y_n) in K is in K. Let (x_n) be a sequence in C such that $y_n = Tx_n$, $n \in \mathbf{N}$. Because C is compact, there is a subsequence (x_{n_k}) of (x_n) that converges to some $x \in C$. Clearly, $y_{n_k} \to y$. Since $G(T)$ is closed, $(x_{n_k}, y_{n_k}) \to (x, y) \in G(T)$. Hence, $y \in K$.

(b) Similar to (a).

6.15. Let (x_n) be a sequence of functions on $[0, 1]$ defined by

$$x_n(t) = \sqrt{t + \frac{1}{n}} - \sqrt{\frac{1}{n}}$$

Clearly, $x_n \in Y$ for all n. We will show that (x_n) uniformly converges to the function $x(t) = \sqrt{t} \notin C^1[0, 1]$.

For $n \in \mathbf{N}$, let

$$f_n(t) = x_n(t) - x(t) = \sqrt{t + \frac{1}{n}} - \sqrt{\frac{1}{n}} - \sqrt{t},$$

and $\varepsilon > 0$. Clearly, $f_n'(t) < 0$ for $t \in (0, 1]$, so the functions f_n's are decreasing. Because $f_n(0) = 0$, we have $f_n(t) < \varepsilon$ for all $t \in [0, 1]$ and $n \in \mathbf{N}$. Now, to prove that (x_n) converges uniformly to x it suffices to show that there is $N \in \mathbf{N}$ such that $-\varepsilon < f_n(t)$ for all $t \in [0, 1]$ and $n > N$. We have

$$f_n(t) + \varepsilon = \sqrt{t + \frac{1}{n}} - \sqrt{t} + \varepsilon - \sqrt{\frac{1}{n}} > 0,$$

for $n > \dfrac{1}{\varepsilon^2}$. Hence, we may choose $N = \lceil 1/\varepsilon^2 \rceil$.

7 Hilbert Spaces

7.1. Straightforward verification.

7.2. Inasmuch as $x, y \in \ell_2$, we have by the Cauchy–Schwarz inequality (Theorem 1.8)

$$\sum_{k=1}^{\infty} |x_k y_k| \leq \sqrt{\sum_{k=1}^{\infty} |x_k|^2} \cdot \sqrt{\sum_{k=1}^{\infty} |y_k|^2}.$$

Hence the series $\sum_{k=1}^{\infty} x_k \overline{y}_k$ converges absolutely (note that $|x_k \overline{y}_k| = |x_k y_k|$), so $\langle x, y\rangle$ is well-defined.

Properties 1–4 of Definition 7.1 are verified directly.

7.3. Clearly, the integral of the zero function is zero. Let $x(t)$ be a continuous nonnegative function on $[a, b]$ such that $x(t_0) > 0$ for some $t_0 \in [a, b]$. Because x is continuous at t_0, for $\varepsilon = x(t_0)/2$ there is a $\delta > 0$ such that $x(t) > \varepsilon$ for $t \in A = [a, b] \cap (t_0 - \delta, t_0 + \delta)$. We have

$$\int_a^b x(t)\, dt \geq \int_A \varepsilon = \varepsilon \cdot \text{length}(A) \geq \varepsilon \cdot \delta > 0.$$

This contradiction shows that x is the zero function.

7.4. For $\varepsilon > 0$, let $\delta = \varepsilon/\sqrt{b-a}$. Suppose that

$$\|x - x_0\|_\infty = \sup_{t\in[a,b]} |x(t) - x_0(t)| < \delta.$$

Then

$$\|x - x_0\| = \sqrt{\int_a^b |x(t) - x_0(t)|^2\, dt} < \sqrt{\delta^2(b-a)} = \varepsilon.$$

Therefore, the identity mapping $X_1 \to X_2$ is continuous at $x = x_0$.

7.5. We have $\langle x, 0\rangle = \langle x, x - x\rangle = \langle x, x\rangle - \langle x, x\rangle = 0$, and $\langle 0, y\rangle = \overline{\langle y, 0\rangle} = 0$.

7.6. We have $\|\lambda x\| = \sqrt{\langle \lambda x, \lambda x\rangle} = \sqrt{\lambda\overline{\lambda}\langle x, x\rangle} = |\lambda|\|x\|$.

7.7. Straightforward verification of properties 1–4 of Definition 7.1.

7.8. Hint. Use Theorems 2.8 and 4.13.

7.9. By the Cauchy–Schwarz inequality (Theorem 7.1),

$$\frac{1}{4}(\langle x, y\rangle + \overline{\langle x, y\rangle})^2 = (\text{Re}\langle x, y\rangle)^2 \leq |\langle x, y\rangle|^2 \leq \|x\|^2\|y\|^2 = \langle x, x\rangle\langle y, y\rangle.$$

7.10. Let $u = z - x$, $v = z - y$, so $x - y = v - u$, $v + u = 2z - (x + y)$. In this notations, we need to show that

$$\|u\|^2 + \|v\|^2 = \frac{1}{2}\|v - u\|^2 + \frac{1}{2}\|v + u\|^2,$$

which is the parallelogram equality (cf. (7.7)).

7.11. We have

$$\|x - x_n\|^2 = \langle x - x_n, x - x_n\rangle = \|x\|^2 + \|x_n\|^2 - 2\mathrm{Re}\langle x, x_n\rangle \to 2\|x\|^2 - 2\|x\|^2 = 0.$$

Hence, $x_n \to x$.

7.12. The sequence $(1/n)$ belongs to the space ℓ_2. Therefore, the function

$$f(x) = \sum_{k=1}^{\infty} \frac{x_k}{k}, \qquad x = (x_1, x_2, \ldots) \in \ell_2,$$

is a continuous linear functional on ℓ_2 (Corollary 7.1). By Theorem 4.6, its null space $X = \{x \in \ell_2 : f(x) = 0\}$ is a closed subspace of ℓ_2.

7.13. We have

$$\begin{aligned}\|tx + (1-t)y\|^2 &= \langle tx + (1-t)y, tx + (1-t)y\rangle \\ &= t^2 + (1-t)^2 + 2\mathrm{Re}\langle tx, (1-t)y\rangle \\ &\le t^2 + (1-t)^2 + 2|\langle tx, (1-t)y\rangle| \\ &< t^2 + (1-t)^2 + 2t(1-t) = 1,\end{aligned}$$

where we have a strict inequality in the last line by Theorem 7.1 (the Cauchy–Schwarz inequality).

7.14. Follows from $\langle x, \sum_{k=1}^{n} \lambda_k x_k\rangle = \sum_{k=1}^{n} \overline{\lambda}_k \langle x, x_k\rangle$.

7.15. (Necessity.) Suppose that $\langle x, y\rangle = 0$. We have

$$\langle x \pm \lambda y, x \pm \lambda y\rangle = \|x\|^2 \pm \overline{\lambda}\langle x, y\rangle \pm \lambda\langle y, x\rangle + |\lambda|^2\|y\|^2 = \|x\|^2 + |\lambda|^2\|y\|^2.$$

Hence, $\|x + \lambda y\| = \|x - \lambda y\|$.

(Sufficiency.) From the first displayed identity we obtain

$$\overline{\lambda}\langle x, y\rangle + \lambda\langle y, x\rangle = 0.$$

Take $\lambda = 1$ if the space is real to obtain $\langle x, y\rangle = 0$. For a complex space use $\lambda = 1$ and $\lambda = \boldsymbol{i}$.

7.16. Suppose that $x \neq 0$, $y \neq 0$, and $x \perp y$. Because the space is 1-dimensional, there is $\lambda \neq 0$ such that $y = \lambda x$. We have $0 = \langle x, \lambda x\rangle = \overline{\lambda}\|x\|^2$, a contradiction. Clearly, the zero vector is orthogonal to every vector.

7.17. By definition, $y \in E^\perp$ if and only if $y \perp x$ for all $x \in E$, that is, y in $\bigcap_{x \in E} \{x\}^\perp$. Hence the result.

7.18. (a) ***Correction.*** Replace $L_2^\perp$ with $X_2^\perp$.

If $x \in (X_1 + X_2)^\perp$, then $x \perp X_1$ and $x \perp X_2$, that is, $x \in X_1^\perp \cap X_2^\perp$. On the other hand, if $x \in X_1^\perp \cap X_2^\perp$, then $x \perp x_1$ and $x \perp x_2$ for all $x_1 \in X_1$, $x_2 \in X_2$. Therefore, $x \perp (x_1 + x_2)$, that is, $x \in (X_1 + X_2)^\perp$.

(b) By part (a), $(X_1^\perp + X_2^\perp)^\perp = X_1 \cap X_2$ (cf. Corollary 7.1). By Corollary 7.2, $\overline{X_1^\perp + X_2^\perp} = (X_1^\perp + X_2^\perp)^{\perp\perp} = (X_1 \cap X_2)^\perp$.

7.19. (a) Clearly, $x \perp \operatorname{span} E$. Let x_0 be a vector in $\overline{\operatorname{span} E}$ and (x_n) a sequence in $\operatorname{span} E$ converging to x_0. By Theorem 7.4, $0 = \langle x, x_n \rangle \to \langle x, x_0 \rangle$. Hence, $x \perp x_0$.

(b) Hint. Apply the argument from (a).

(c) Clearly, $E^\perp$ is a subspace of X. Let x_0 be a limit point of $E^\perp$ and (x_n) a sequence in $E^\perp$ converging to x_0. By Theorem 7.4, for every $x \in E$, we have $0 = \langle x, x_n \rangle \to \langle x, x_0 \rangle$. Hence, $x_0 \in E^\perp$. It follows that $E^\perp$ is a closed subspace of X.

7.20. Let (x_n) be a sequence of functions in E that converges to some function $x \in X$. Suppose that $x(t)$ is not a zero function on $[-1, 0]$. We have

$$\begin{aligned}\|x - x_n\|^2 &= \int_{-1}^{1} (x(t) - x_n(t))^2 \, dt = \int_{-1}^{0} x^2(t) \, dt + \int_{0}^{1} (x(t) - x_n(t))^2 \, dt \\ &\geq \int_{-1}^{0} x^2(t) \, dt > 0\end{aligned}$$

(cf. Exercise 7.3), a contradiction. It follows that $x \in E$, that is, E is a closed subspace of X.

Every function $x(t)$ in $E \oplus E^\perp$ is the sum of a function in E and a function in $E^\perp$ (cf. Example 7.8). Hence, $x(0) = 0$ and $E \oplus E^\perp = \{x \in X : x(0) = 0\}$.

7.21. (a) We have

$$\|x + y\|^2 = \langle x + y, x + y \rangle = \|x\|^2 + \langle x, y \rangle + \langle y, x \rangle + \|y\|^2 = \|x\|^2 + \|y\|^2.$$

(b) Straightforward mathematical induction.

(c) We have, by (b) and Theorem 2.12,

$$\sum_{k=1}^{n} \|x_k\|^2 = \left\| \sum_{k=1}^{n} x_k \right\|^2 \to \left\| \sum_{k=1}^{\infty} x_k \right\|^2.$$

Hence the result.

7.22. (a) ***Correction.*** Holds only for a real inner product space.

(Necessity.) Follows from Exercise 7.21(a).

(Sufficiency.) Suppose to the contrary that $\|x_1 + x_2\|^2 = \|x_1\|^2 + \|x_2\|^2$ for all $x_1 \in E_1$, $x_2 \in E_2$, and E_1 is not orthogonal to E_2. Then there are $x \in E_1$, $y \in E_2$ such that $\langle x, y\rangle \neq 0$. Therefore,

$$\|x + y\|^2 = \langle x + y, x + y\rangle = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle \neq \|x\|^2 + \|y\|^2,$$

a contradiction.

(b) Let x be a nonzero vector in a complex inner product space. We have

$$\|x + (\boldsymbol{i}x)\|^2 = \|x\|^2 + 2\mathrm{Re}\langle x, \boldsymbol{i}x\rangle + \|\boldsymbol{i}x\|^2 = \|x\|^2 + \|\boldsymbol{i}x\|^2,$$

because $\langle x, \boldsymbol{i}x\rangle = -\boldsymbol{i}\|x\|^2$. However, vectors x and $\boldsymbol{i}x$ are not orthogonal.

7.23. Let $\delta = \inf_{w\in E} \|x - w\|$. By the definition of an infimum, there is a sequence (u_n) in E such that $\delta_n = \|x - u_n\| \to \delta$. By Apollonius's Identity (cf. Exercise 7.10),

$$\|x - u_n\|^2 + \|x - u_m\|^2 = \frac{1}{2}\|u_n - u_m\|^2 + 2\left\|x - \frac{u_n + u_m}{2}\right\|^2$$

Therefore,

$$\begin{aligned} 0 \leq \|u_n - u_m\|^2 &= 2\|x - u_n\|^2 + 2\|x - u_m\|^2 - 4\left\|x - \frac{u_n + u_m}{2}\right\|^2 \\ &< 2\delta_n^2 + 2\delta_m^2 - 4\delta^2 = 2(\delta_n^2 - \delta^2) + 2(\delta_m^2 - \delta^2) \to 0, \end{aligned}$$

because $(u_n + u_m)/2 \in E$, by convexity of E. It follows that (u_n) is Cauchy in E and therefore converges, $u_n \to u \in E$. We have

$$\delta \leq \|x - u\| \leq \|x - u_n\| + \|u_n - u\| \to \delta.$$

Hence, $\|x - u\| = \delta$, that is, u is a desired vector in E.

To show the uniqueness of u, suppose that there is $v \neq u$ in E such that $\|x - v\| = \delta$. By Apollonius's Identity again,

$$\|x - u\|^2 + \|x - v\|^2 = \frac{1}{2}\|u - v\|^2 + 2\left\|x - \frac{u + v}{2}\right\|^2.$$

Therefore,

$$\delta^2 + \delta^2 - \frac{1}{2}\|u - v\|^2 = 2\left\|x - \frac{u + v}{2}\right\|^2 \geq 2\delta^2,$$

a contradiction. It follows that $v = u$, which proves the uniqueness of u.

Counterexamples in $\mathbf{R}$: (a) $E = (0, 1)$, $x = 1$, (b) $E = \{-1\} \cup \{1\}$, $x = 0$.

7.24. Hint. Use the Table of Integrals in a Calculus textbook.

7.25. Let x be the sum of the summable family $\{b_i\}_{i\in J}$. By Definition 7.6, there is a finite subset A of J such that (for $\varepsilon = 1$)

$$\sum_{i\in B} b_i < x + 1, \qquad \text{for all finite sets } B \subseteq J \text{ containing the set } A.$$

Let C be a finite subset of J. Then, $A \subseteq C \cup A$. Hence,

$$\sum_{i \in C} b_i \leq \sum_{i \in C \cap A} b_i < x + 1.$$

By Lemma 7.2, $x = \sup\{\sum_{i \in C} b_i : C \text{ is finite subset of } J\}$.

Clearly, $\sum_{i \in C} a_i \leq \sum_{i \in C} b_i$ for every finite subset C of J. By Lemma 7.2, the family $\{a_i\}_{i \in J}$ is summable with $\sum_{i \in J} a_i \leq \sum_{i \in J} b_i$.

7.26. By Theorem 7.9 (Bessel's inequality) and Theorem 1.8 (Cauchy–Schwarz inequality),

$$\sum_{k=1}^{\infty} |\langle x, e_k \rangle \langle y, e_k \rangle| \leq \sqrt{\sum_{k=1}^{\infty} |\langle x, e_k \rangle|^2} \cdot \sqrt{\sum_{k=1}^{\infty} |\langle y, e_k \rangle|^2} \leq \|x\| \|y\|.$$

7.27. Because $\{x_i\}_{i \in J}$ is summable with the sum x, for every $\varepsilon > 0$ and every finite subset C of J, $\sum_{i \in C} x_i < x + \varepsilon$ (cf. solution of Exercise 7.25). By Lemma 7.2, $x = \sup_C \{\sum_{i \in C} x_i\}$, where supremum is taken over all finite subsets C of J. By the same lemma, the family $\{x_i\}_{i \in J_0}$ is summable with the sum

$$\sum_{i \in J_0} x_i = \sup_{C \subseteq J_0} \Big\{ \sum_{i \in C} x_i \Big\} \leq \sup_{C \subseteq J} \Big\{ \sum_{i \in C} x_i \Big\} = x.$$

7.28. We prove the following statement: Let $\{x_i\}_{i \in J}$ be a summable family of elements of $\mathbf{F}$ and let $\varphi : K \to J$ be a bijection of an index set K onto J. We denote $y_k = x_{\varphi(k)}$, $k \in K$ and show that the family $\{y_k\}_{k \in K}$ is summable with $\sum_{k \in K} y_k = \sum_{i \in J} x_i$.

Let $x = \sum_{i \in J} x_i$. If $\big| \sum_{i \in B} x_i - x \big| < \varepsilon$ (cf. Definition 7.6) for all finite sets B containing a finite set A, then $\big| \sum_{k \in L} y_k - x \big| < \varepsilon$ for every finite subset L of K containing $\varphi^{-1}(A)$. Hence the result.

7.29. It is easy to verify that $\langle e^{\boldsymbol{i}\lambda t}, e^{\boldsymbol{i}\mu t} \rangle = 0$, if $\lambda \neq \mu$, and $\langle e^{\boldsymbol{i}\lambda t}, e^{\boldsymbol{i}\lambda t} \rangle = 1$ for every $\lambda, \mu \in \mathbf{R}$. Suppose that $f(t) = \sum_{k=1}^{n} c_k e^{\boldsymbol{i}\lambda_k t}$ is orthogonal to every vector in X. Then f must be orthogonal to every vector $e^{\boldsymbol{i}\lambda t}$. In particular, for every $1 \leq m \leq n$,

$$\Big\langle \sum_{k=1}^{n} c_k e^{\boldsymbol{i}\lambda_k t}, e^{\boldsymbol{i}\lambda_m t} \Big\rangle = \sum_{k=1}^{n} c_k \langle e^{\boldsymbol{i}\lambda_k t}, e^{\boldsymbol{i}\lambda_m t} \rangle = c_m = 0.$$

Hence, f is the zero function on $\mathbf{R}$.

7.30. Properties 1–4 of Definition 7.1 are verified directly (cf. Example 7.2).

We apply Gram–Schmidt process to $x_1(t) = 1$, $x_2(t) = t$, $x_3(t) = t^2$ on $[-1, 1]$ with the inner product $\langle x, y \rangle = \int_{-1}^{1} x(t) y(t)\, dt$ (cf. Example 7.2).

Step 1. Clearly, $e_1 = x_1/\|x_1\| = 1/\sqrt{2}$.

Step 2. The vector $v_2 = x_2 - \langle x_2, e_1\rangle e_1$ is orthogonal to e_1. We have

$$v_2(t) = t - \int_{-1}^{1} t(1/\sqrt{2})\,dt \cdot \frac{1}{\sqrt{2}} = t.$$

Hence, $e_2 = v_2/\|v_2\| = \sqrt{3/2}\,t$ ($\|t\|^2 = \frac{2}{3}$).

Step 3. The vector $v_3 = x_3 - \langle x_3, e_1\rangle e_1 + \langle x_3, e_2\rangle e_2$ is orthogonal to e_1 and e_2. We have

$$v_3(t) = t^2 - \int_{-1}^{1} t^2 \frac{1}{\sqrt{2}}\,dt \cdot \frac{1}{\sqrt{2}} - \int_{-1}^{1} t^2 \sqrt{\frac{3}{2}}\,t\,dt \cdot \sqrt{\frac{3}{2}} = t^2 - \frac{1}{3}.$$

It is easy to verify that $\|v_3\| = \frac{2\sqrt{2}}{3\sqrt{5}}$. Hence, $e_3 = \frac{3\sqrt{5}}{2\sqrt{2}}\left(t^2 - \frac{1}{3}\right)$.

7.31. Choose L as a Hamel basis (cf. Theorem 1.18) of the vector space $\operatorname{span} S$.

7.32. We have $\langle x_1, y\rangle - \langle x_2, y\rangle = \langle x_1 - x_2, y\rangle = 0$. Therefore, for $y = x_1 - x_2$, $\|x_1 - x_2\|^2 = \langle x_1 - x_2, x_1 - x_2\rangle = 0$. Hence, $x_1 = x_2$.

7.33. We verify properties 1–4 from Definition 7.1.

Property 1.

$$\begin{aligned}\langle Tx + Ty, Tz\rangle_1 &= \langle T(x+y), Tz\rangle_1 = \langle z, x+y\rangle = \langle z, x\rangle + \langle z, y\rangle \\ &= \langle Tx, Tz\rangle_1 + \langle Ty, Tz\rangle_1.\end{aligned}$$

Property 2. Recall that T is conjugate–linear (Theorem 7.17.)

$$\langle \lambda Tx, Ty\rangle_1 = \langle T\overline{\lambda}x, Ty\rangle_1 = \langle y, \overline{\lambda}x\rangle = \lambda\langle y, x\rangle = \lambda\langle Tx, Ty\rangle_1.$$

Property 3. $\langle Tx, Ty\rangle_1 = \langle y, x\rangle = \overline{\langle x, y\rangle} = \overline{\langle Ty, Tx\rangle_1}$.

Property 4. $\langle Tx, Tx\rangle_1 = \langle x, x\rangle$. Hence $\langle Tx, Tx\rangle_1 \geq 0$, and $\langle Tx, Tx\rangle_1 = 0$ if and only if $Tx = 0$.

7.34. Because

$$\frac{S(x,y)}{\|x\|\|y\|} = S\left(\frac{x}{\|x\|}, \frac{y}{\|y\|}\right) \qquad x \neq 0\ y \neq 0,$$

the formula

$$\|S\| = \sup\left\{\frac{|S(x,y)|}{\|x\|\|y\|} : x \in X \setminus \{0\},\ y \in Y \setminus \{0\}\right\}$$

follows from the definition on p. 173. The above formula immediately implies $|S(x,y)| \leq \|S\|\|x\|\|y\|$.

7.35. (Necessity.) Trivial.

(Sufficiency.) Suppose there is $x \in X$ such that $Tx \neq 0$. Then,

$$0 = \langle Tx, Tx \rangle = \|Tx\|^2,$$

a contradiction.

7.36. For $\lambda \in \mathbf{C}$ and $x, y \in X$, we have

$$\begin{aligned} 0 &= \langle T(\lambda x + y), \lambda x + y \rangle = \langle \lambda Tx + Ty, \lambda x + y \rangle \\ &= |\lambda|^2 \langle Tx, x \rangle + \langle Ty, y \rangle + \lambda \langle Tx, y \rangle + \overline{\lambda} \langle Ty, x \rangle \\ &= \lambda \langle Tx, y \rangle + \overline{\lambda} \langle Ty, x \rangle. \end{aligned}$$

For $\lambda = 1$, $\langle Tx, y \rangle + \langle Ty, x \rangle = 0$, and, for $\lambda = \boldsymbol{i}$, $\langle Tx, y \rangle - \langle Ty, x \rangle = 0$. Hence, $\langle Tx, y \rangle = 0$ for all $x, y \in X$ and the result follows from Exercise 7.35.

Hint. Let T be a 90° rotation of $\mathbf{R}^2$.

7.37. Because the series $\sum_{k=1}^{\infty} 1/k^2$ converges, $x = (1/n) \in \ell_2$. Clearly, the vector x is not in $\mathcal{R}(T)$. Let $x^{(n)} = (1, 1/2, 1/3, \ldots, 1/n, 0, 0, \ldots) \in \mathcal{R}(T)$. We have

$$\|x - x^{(n)}\|^2 = \left\| (0, \ldots, 0, \frac{1}{n+1}, \frac{1}{n+2}, \ldots) \right\|^2 = \sum_{k=n+1}^{\infty} \frac{1}{k^2} \to 0,$$

since $\sum_{k=1}^{\infty} 1/k^2 < \infty$. It follows that x is a limit point of $\mathcal{R}(T)$ that does not belong to this set.

7.38. By Theorem 7.23 and Corollary 7.2,

$$\mathcal{N}(T)^{\perp} = \mathcal{R}(T^*)^{\perp\perp} = \overline{\mathcal{R}(T^*)},$$

because $\mathcal{R}(T^*)$ is the span of itself. Similarly,

$$\mathcal{N}(T^*)^{\perp} = \mathcal{R}(T)^{\perp\perp} = \overline{\mathcal{R}(T)}.$$

7.39. Suppose that $Sx = Sy$, that is, $T^*Tz = -z$, where $z = y - x$. We have

$$-\|z\|^2 = -\langle z, z \rangle = \langle T^*Tz, z \rangle = \langle Tz, Tz \rangle = \|Tz\|^2,$$

which implies $z = 0$. Hence, S is one-to-one.

7.40. By Theorem 6.3 (the Open Mapping Theorem), T^{-1} is a bounded operator. By Theorem 7.22 (c),

$$I = I^* = (TT^{-1})^* = (T^{-1})^* T^* = (T^{-1})^* T,$$

so $(T^{-1})^* = T^{-1}$.

7.41. Hint. Use Example 7.15 and elementary operations on 2×2 matrices.

7.42. Hint. Apply Theorem 7.22 (a) and (b).

7.43. Straightforward verification.

7.44. For a nonzero vector $x = (x_1, x_2, \ldots) \in \ell_2$, the equation $Sx = \lambda x$ yields

$$(0, x_1, x_2, \ldots) = (\lambda x_1, \lambda x_2, \lambda x_3, \ldots),$$

with no solutions.

From the equation $Tx = \lambda x$, we easily obtain $x_n = x_1 \lambda^{n-1}$, $n > 1$. The geometric series $\sum_{k=1}^{\infty} x_1^2 \lambda^{2(k-1)}$ converges if and only if $|\lambda| < 1$.

7.45. Clearly, the required result follows from $(T^n)^* = (T^*)^n$, $n \in \mathbf{N}$. This is trivially true for $n = 1$. Suppose that, for some $k \in \mathbf{N}$, $(T^k)^* = (T^*)^k$. By Theorem 7.22 (c),

$$(T^{k+1})^* = (TT^k)^* = (T^k)^* T^* = (T^*)^k T^* = (T^*)^{k+1}.$$

The result follows by mathematical induction.

7.46. ***Correction.*** The second operator in part (a) must be T_2.

(a) Follows immediately from Theorem 7.22 (a) and (b).

(b) (Necessity.) If T is normal, then

$$T_1 T_2 = \frac{1}{4\boldsymbol{i}}(T^2 + T^*T - TT^* - (T^*)^2) = \frac{1}{4\boldsymbol{i}}(T^2 - (T^*)^2)$$

and

$$T_2 T_1 = \frac{1}{4\boldsymbol{i}}(T^2 + TT^* - T^*T - (T^*)^2) = \frac{1}{4\boldsymbol{i}}(T^2 - (T^*)^2)$$

Hence, $T_1 T_2 = T_2 T_1$.

(Sufficiency.) From the above displayed identities, we have

$$0 = T_1 T_2 - T_2 T_1 = \frac{1}{4\boldsymbol{i}}(2T^*T - 2TT^*).$$

Therefore, $TT^* = T^*T$.

7.47. By Theorem 7.22 (e), $\|T^2\| = \|T^*T\| = \|T\|^2$. Therefore, the square of any nonzero self-adjoint operator is a nonzero self-adjoint operator (cf. Exercise 7.45).

Suppose to the contrary that there is a power of T that is the zero operator, and let n be the least integer such that $T^n = 0$. Then $T^k \neq 0$ for $k < n$, and $T^k = 0$ for all $k \geq n$. The square of the nonzero operator T^{n-1} is the operator T^{2n-2} which is the zero operator because $2n - 2 \geq n$, a contradiction. Hence, $T^n \neq 0$ for all $n \in \mathbf{N}$.

7.48. Let $X = \mathcal{R}(T)$ and $y \in \overline{X}$. There is a sequence (y_n) in X that converges to y. Let $y_n = Tx_n$ for $n \in \mathbf{N}$. Inasmuch as T is an isometry and (y_n) is Cauchy, the sequence (x_n) is also Cauchy. Because H is complete, (x_n) converges to some $x \in H$. Then $Tx_n \to Tx$, so $y = Tx$. It follows that X is a closed subspace of H. It is proper because otherwise T would be unitary.

7.49. Let $\{e_1, \ldots, e_n\}$ be a basis in X. Suppose that vectors $\{Te_1, \ldots, Te_n\}$ are linearly dependent, that is, there are elements $c_1, \ldots, c_n$ in $\mathbf{F}$ not all zero such that $\sum_{k=1}^n c_k Te_k = 0$. Then $Tx = 0$ for $x = \sum_{k=1}^n c_k e_k \neq 0$. This contradicts isometry property of T, because $0 = \|Tx - T0\| \neq \|x - 0\|$. Therefore, T maps X onto X. By Exercise 7.48, T is unitary (see also Theorem 7.28).

7.50. Note that $T_n \to T$ implies $T_n^* \to T^*$, because $\|T_n^* - T^*\| = \|(T_n - T)^*\| = \|T_n - T\|$.

We have

$$\begin{aligned} 0 \leq \|TT^* - T^*T\| &\leq \|TT^* - T_nT_n^*\| + \|T_nT_n^* - T_n^*T_n\| + \|T_n^*T_n - T^*T\| \\ &= \|TT^* - T_nT_n^*\| + \|T_n^*T_n - T^*T\| \to 0, \end{aligned}$$

because $T_nT_n^* \to TT^*$ and $T_n^*T_n \to T^*T$.

7.51. From $ST^* = T^*S$ we obtain $TS^* = S^*T$, because $(ST^*)^* = (T^*S)^*$. Therefore,

$$\begin{aligned} (S+T)(S^* + T^*) &= SS^* + TS^* + ST^* + TT^* = S^*S + S^*T + T^*S + T^*T \\ &= (S^* + T^*)(S+T), \end{aligned}$$

and

$$(ST)(ST)^* = STT^*S^* = ST^*TS^* = T^*SS^*T = T^*S^*ST = (ST)^*(ST).$$

7.52. Hint. Apply Theorem 7.20 and the argument from the paragraph preceding this theorem.

7.53. Hint. $\|P\| = 1$.

7.54. (Necessity.) Suppose that $E \perp E'$. Then for every vector x, $P_{E'}x \in E'$. and therefore $P_E(P_{E'}x) = 0$, so $P_EP_{E'}$ is the zero operator.

(Sufficiency.) If $P_EP_{E'} = 0$, then, for $x \in E$, $x' \in E'$, we have

$$\langle x, x' \rangle = \langle P_E x, P_{E'}x' \rangle = \langle x, P_EP_{E'}x' \rangle = \langle x, 0 \rangle = 0,$$

so $E \perp E'$.

7.55. We have

$$(P + Q - PQ)^* = P^* + Q^* - (PQ)^* = P + Q - QP = P + Q - PQ.$$

Hence, $P+Q-PQ$ is a self-adjoint operator. Furthermore, because $I-Q$ is a projection operator,

$$\begin{aligned}
(P+Q-PQ)^2 &= [P(I-Q)+Q]^2 \\
&= P^2(I-Q)^2 + P(I-Q)Q + QP(I-Q) + Q^2 \\
&= P(I-Q) + PQ - PQ + QP - QPQ + Q \\
&= P(I-Q) + PQ - PQ + QP - PQ + Q \\
&= P + Q - PQ.
\end{aligned}$$

By Theorem 7.29, $P+Q-PQ$ is a projection operator.

For every $x \in H$, we have $Px \in E$, $Qx \in E'$, and, by Theorem 7.30, $(PQ)x \in E \cap E'$. Hence, $(P+Q-PQ)x \in E+E'$, that is,

$$(P+Q-PQ)(H) \subseteq E+E'.$$

Every vector $z \in E+E'$ is the sum of some vectors $x \in E$ and $x' \in E'$, $z = x + x'$.

For $x \in E$,

$$(P+Q-PQ)x = Px + Q(I-P)x = Px = x,$$

because $(I-P)x = 0$. Similarly, $(P+Q-PQ)x' = x'$ for $x' \in E'$. Therefore,

$$(P+Q-PQ)z = (P+Q-PQ)x + (P+Q-PQ)x' = x + x' = z.$$

It follows that $P+Q-PQ$ is a projection of H onto $E+E'$.

7.56. First, we note that conditions $P_E P_{E'} = P_E$ and $P_{E'} P_E = P_E$ are equivalent because projection operators are self-adjoint.

(Necessity.) If $x \in E$, then

$$x = P_E x = (P_{E'} P_E)x = P_{E'} x,$$

so $x \in E'$. Hence, $E \subseteq E'$.

(Sufficiency.) For every $z \in H$, we have $(P_{E'} P_E)z = P_E z$, because $E \subseteq E'$.

7.57. Let U be a self-adjoint unitary operator. Then $U^2 = I$ (cf. Definition 7.12).

(Necessity.) Let U be a self-adjoint unitary operator and $P = \frac{1}{2}(U+I)$. We show that P is a projection operator. Clearly, P is bounded and self-adjoint. Furthermore,

$$P^2 = \tfrac{1}{4}(U^2 + 2U + I) = \tfrac{1}{2}(U+I) = P.$$

By Theorem 7.29, P is a projection operator. Obviously, $U = 2P - I$.

(Sufficiency.) Let P be a projection operator and $U = 2P - I$. Clearly, U is self-adjoint and

$$U^2 = (2P-I)^2 = 4P^2 - 4P + I = I,$$

because $P^2 = P$. Hence, U is a unitary operator.

7.58. By Exercise 7.57, U is a self-adjoint unitary operator if and only if $U = 2P - I$ for a projection operator P.

(Necessity.) Let $E_1 = P(H)$, $E_2 = E_1^{\perp}$, $x = x_1 + x_2 \in H$, where $x_1 \in E_1$, $x_2 \in E_2$. Then

$$Ux = 2P(x_1 + x_2) - I(x_1 + x_2) = 2x_1 - x_1 - x_2 = x_1 - x_2.$$

(Sufficiency.) Let P be the projection operator onto the subspace E_1. Then

$$Ux = x_1 - x_2 = 2x_1 - x_1 - x_2 = (2P - I)x, \qquad \text{for all } x \in H.$$

References

[1] S. Axler, *Linear Algebra Done Right*, Third Edition, Springer, 2015.

[2] E. Kreyzig, *Introductory Functional Analysis with Applications*, Wiley, 1989.

[3] H.L. Royden and P.M. Fitzpatrick, *Real Analysis*, Fourth Edition, Prentice Hall, 2010.

[4] W. Rudin, *Principles of Mathematical Analysis*, Third Edition, McGraw-Hill, 1976.