PROOF. $(r,\gamma) \sim (r,\gamma)$ is obvious: $r = r$ and $\gamma - \gamma = 0 \cdot 2\pi$. To see $\sim$ is symmetric, let $(r,\gamma) \sim (r',\gamma')$; then $r = r'$ and $\gamma - \gamma' = n \cdot 2\pi$, where $n \in \mathbb{Z}$. Therefore, $r' = r$ and $\gamma' - \gamma = (-n) \cdot 2\pi$; that is, $(r',\gamma') \sim (r,\gamma)$. Finally, to see $\sim$ is transitive, let $(r,\gamma) \sim (r',\gamma')$, and $(r',\gamma') \sim (r'',\gamma'')$. In this case, $r = r' = r''$, so $r = r''$, and

$$\gamma - \gamma' = m \cdot 2\pi, \quad \gamma' - \gamma'' = n \cdot 2\pi,$$

where $m, n \in \mathbb{Z}$. But then $\gamma - \gamma'' = (\gamma - \gamma') + (\gamma' - \gamma'') = (m+n) \cdot 2\pi$. Hence, $(r,\gamma) = (r'',\gamma'')$. The above steps show that $\sim$ is an equivalence relation on P.

Consider an arbitrary element of $P/\sim$, say, $[(r',\gamma')]_\sim$. Since $\gamma' \in \mathbb{R}$, there must exist γ such that $\gamma' - \gamma = n \cdot 2\pi$, where $n \in \mathbb{Z}$. Then, there exists $\gamma \in \mathbb{R}$ such that $\gamma = \gamma' - n \cdot 2\pi$. Hence, we can find a $\widetilde{n} \in \mathbb{Z}$ satisfying $\gamma'/2\pi - 1 \leqslant \widetilde{n} \leqslant \gamma'/2\pi$, and let $(r,\gamma) = (r', \gamma' - \widetilde{n} \cdot 2\pi)$. □

2.5 Orderings

► EXERCISE 44 (2.5.1). a. *Let R be an ordering of A, S be the corresponding strict ordering of A, and R^* be the ordering corresponding to S. Show that $R^* = R$.*

b. *Let S be a strict ordering of A, R be the corresponding ordering, and S^* be the strict ordering corresponding to R. Then $S^* = S$.*

PROOF. **(a)** Let $(a,b) \in R$, where $a, b \in A$. If $a = b$, then $(a,b) \in R^*$ because orderings are reflexive; if $a \neq b$, then $(a,b) \in S$. But then $(a,b) \in R^*$. Hence, $R \subset R^*$. To see the inverse direction, let $(a,b) \in R^*$. Firstly, $a = b$ implies that $(a,b) \in R$ since R is reflexive. So we suppose $a \neq b$. In this case, $(a,b) \in S$. Because S is R's corresponding strict ordering of A, we know $(a,b) \in S$ if and only if $(a,b) \in R$ and $a \neq b$. Hence, $R^* \subset R$. This proves that $R^* = R$.

(b) Let $(a,b) \in S$, then $a \neq b$. Since R is S's corresponding ordering, we have $(a,b) \in R$. Since $(a,b) \in R$ and $a \neq b$, we have $(a,b) \in S^*$. The revers direction can be proven with the same logic. □

► EXERCISE 45 (2.5.2). *State the definitions of incomparable elements, maximal, minimal, greatest, and least elements and suprema and infima in terms of strict orderings.*

SOLUTION. If $(P, <)$ is a partially ordered set, X is a nonempty subset of P, and $a \in P$, then:

- a and b are incomparable in $<$ if $a \neq b$ and neither $a < b$ nor $b < a$ holds;
- a is a *maximal* element of X if $a \in X$ and $(\forall\, x \in X)\; a \nless x$;
- a is a *minimal* element of X if $a \in X$ and $(\forall\, x \in X)\; x \nless a$;

- a is the *greatest* element of X if $a \in X$ and $(\forall\, x \in X)\ x \leqslant a$;
- a is the *least* element of X if $a \in X$ and $(\forall\, x \in X)\ a \leqslant x$;
- a is an *upper bound* of X if $(x \in X)\ x \leqslant a$;
- a is a *lower bound* of X if $(\forall\, x \in X)\ a \leqslant x$;
- a is the *supremum* of X if a is the least upper bound of X;
- a is the *infimum* of X if a is the greatest lower bound of X. □

► EXERCISE 46 (2.5.3). *Let R be an ordering of A. Prove that R^{-1} is also an ordering of A, and for $B \subseteq A$,*

a. *a is the least element of B in R^{-1} if and only if a is the greatest element of B in R;*

b. *Similarly for (minimal and maximal) and (supremum and infimum).*

PROOF. **(a)** (i) $aR^{-1}a$ since aRa. (ii) Suppose $(a,b) \in R^{-1}$ and $(b,a) \in R^{-1}$. Then $(b,a) \in R$ and $(a,b) \in R$, and so $a = b$ since R is antisymmetric. (iii) Let $aR^{-1}b$ and $bR^{-1}c$. Then bRa and cRb. Hence, cRa since R is transitive. But which means that $aR^{-1}c$, i.e., R^{-1} is transitive.

(b) If a is the least element of B in R^{-1}, then $a \in B$ and $aR^{-1}x$ for all $x \in B$. But then xRa for all $x \in B$, i.e., a is the greatest element of B in R; if a be the greatest element of B in R, that is, $a \in B$ and xRa for all $x \in B$, then $aR^{-1}x$ for all $x \in B$, and so a is the least element of B in R^{-1}. With the same logic as (a) we can get (b). □

► EXERCISE 47 (2.5.4). *Let R be an ordering of A and let $B \subseteq A$. Show that $R \cap B^2$ is an ordering of B.*

PROOF. (i) For every $b \in B$ we have $(b,b) \in B^2$ and $(b,b) \in R$; hence, $(b,b) \in R \cap B^2$; that is, $R \cap B^2$ is reflexive. (ii) Let $(a,b) \in R \cap B^2$ and $(b,a) \in R \cap B^2$. Then $(a,b) \in R$ and $(b,a) \in R$ imply that $a = b$. Therefore, $R \cap B^2$ is antisymmetric. (iii) Let $(a,b) \in R \cap B^2$ and $(b,c) \in R \cap B^2$. Then $(a,b) \in R$ and $(b,c) \in R$ implies that $(a,c) \in R$. Furthermore, since both $a \in B$ and $c \in B$, we have $(a,c) \in B^2$. Hence, $(a,c) \in R \cap B^2$; that is, $R \cap B^2$ is transitive. □

► EXERCISE 48 (2.5.5). *Give examples of a finite ordered set $(A, \leqslant)$ and a subset B of A so that*

a. *B has no greatest element.*

b. *B has no least element.*

c. *B has no greatest element, but B has a supremum.*

d. *B has no supremum.*

PROOF. **(a)** Let $A = \{a, b, c, d\}$, $B = \{a, b, c\}$, and

$$\leqslant = \{(a, a), (b, b), (c, c), (d, d), (a, d), (b, d), (c, d)\}.$$

In this example, a is not the greatest element of B because (a, b), (a, c) are incomparable; similarly, b and c are not the greatest elements of B.

(b) As the example in (a), there is no least element.

(c) As the example in (a), there is no greatest element, but d is an upper bound of B, and it is the least upper bound of B, so d is the supremum of B.

(d) Let $A = \{a, b, c, c\}$, $B = \{a, b, c\}$, and $\leqslant = \{(a, a), (b, b), (c, c), (d, d)\}$. Then there is no upper bound of B, and consequently, B has no supremum. □

► EXERCISE 49 (2.5.6). a. *Let $(A, <)$ be a strictly ordered set and $b \notin A$. Define a relation $\prec$ in $B = A \cup \{b\}$ as follows:*

$$x \prec y \text{ if and only if } (x, y \in A \text{ and } x < y) \text{ or } (x \in A \text{ and } y = b).$$

Show that $\prec$ is a strict ordering of B and $\prec \cap A^2 = <$.

b. *Generalize part (a): Let $(A_1, <_1)$ and $(A_2, <_2)$ be strict orderings, $A_1 \cap A_2 = \varnothing$. Define a relation $\prec$ on $B = A_1 \cup A_2$ as follows:*

$$\begin{aligned} x \prec y \text{ if and only if } & x, y \in A_1 \text{ and } x <_1 y \\ & \text{or } x, y \in A_2 \text{ and } x <_2 y \\ & \text{or } x \in A_1 \text{ and } y \in A_2. \end{aligned}$$

Show that $\prec$ is a strict ordering of B and $\prec \cap A_1^2 = <_1$, $\prec \cap A_2^2 = <_2$.

PROOF. **(a)** Let $x \prec y$. Then either $x, y \in A$ and $x < y$ or $x \in A$ and $y = b$. In the first case, $y \nprec x$ because $y \nless x$; in the later case, $y \nprec x$ be definition. Therefore, $\prec$ is asymmetric.

Let $x \prec y$ and $y \prec z$. Then $y \neq b$; otherwise, $y \prec z$ cannot hold. With the same logic, $x \neq b$, too. If $z = b$, then $x \prec z = b$ by definition; if $z \in A$, then $x < y$ and $y < z$ implies $x < z$ and so $x \prec z$.

To prove $(\prec \cap A^2) = <$, let $(x, y) \in (\prec \cap A^2)$. Then $x, y \in A$ and $(x, y) \in \prec$, which means that $(x, y) \in <$. Now let $(x, y) \in <$. Then $x, y \in A \implies (x, y) \in A^2$ and $(x, y) \in \prec$ by definition of $\prec$; hence, $(x, y) \in (\prec \cap A^2)$.

(b) Let $x \prec y$. If $x, y \in A_1$, then $x <_1 y$ and so $y \nprec x$; if $x, y \in A_2$, then $x <_2 y$ and so $y \nprec x$; if $x \in A_1$ and $y \in A_2$, then $y \nprec x$ by definition.

Let $x \prec y$ and $y \prec z$. There are four cases:

- $x, y, z \in A_1$. In this case, $x <_1 y <_1 z \implies x <_1 z \implies x \prec z$.
- $x, y, z \in A_2$. In this case, $x <_2 y <_2 z \implies x <_2 z \implies x \prec z$.
- $x, y \in A_1$ and $z \in A_2$. In this case, $x \prec z$ by definition.

- $x \in A_1$ and $y, z \in A_2$. In this case, $x \prec z$ by definition.

To prove $\left(\prec \cap A_1^2\right) = <_1$, suppose $(x, y) \in \left(\prec \cap A_1^2\right)$ firstly. Then $(x, y) \in \prec$ and $x, y \in A_1$; hence $x \prec y \Longrightarrow x <_1 y$. Now suppose $(x, y) \in <_1$. Then $x \prec y$ and $x, y \in A_1$; that is, $(x, y) \in \left(\prec \cap A_1^2\right)$.

The result that $\left(\prec \cap A_2^2\right) = <_2$ can be proved with the same logic. □

▶ EXERCISE 50 (2.5.7). *Let R be a reflexive and transitive relation in A (R is called a* preordering *of A). Define E in A by*

$$aEb \text{ if and only if } aRb \text{ and } bRa.$$

Show that E is an equivalence relation on A. Define the relation R/E in A/E by

$$[a]_E\,(R/E)[b]_E \text{ if and only if } aRb.$$

Show that the definition does not depend on the choice of representatives for $[a]_E$ and $[b]_E$. Prove that R/E is an ordering of A/E.

PROOF. We first show that E is an equivalence relation on A. (i) E is reflexive since R is. (ii) E is symmetric: if aEb, then aRb and bRa, i.e., bRa and aRb; therefore, bEa by the definition of E. (iii) E is transitive: if aEb and bEc, then aRb and bRa, and bRc and cRb. Hence, aRc and cRa by the transitivity of R. We thus have aEc.

Let $[a]_E(R/E)[b]_E$ if and only if aRb. We show that if $c \in [a]_E$ and $d \in [b]_E$, then $[a]_E(R/E)[b]_E$ if and only if cRd. We firt focus on the "IF" part. Since $c \in [a]_E$, we have cEa, i.e., aRc and cRa; similarly, dRb and bRd. Let cRd. We first have aRd since aRc; we also have dRb; hence aRb, i.e., cRd implies that $[a]_E(R/E)[b]_E$. To prove the "ONLY IF" part, let $[a]_E(R/E)[b]_E$. Then aRb. Since cRa and bRd, we have cRd.

R/E is an ordering of A/E since (i) R/E is reflexive: for any $[a]_E \in A/E$, we have $a \in [a]_E$ and aRa, so $[a]_E(R/E)[a]_E$; (ii) R/E is antisymmetric: if $[a]_E(R/E)[b]_E$ and $[b]_E(R/E)[a]_E$, then aRb and bRa, i.e., aEb. Hence, $[a]_E = [b]_E$; (iii) R/E is transitive: if $[a]_E(R/E)[b]_E$ and $[b]_E(R/E)[c]_E$, then aRb and bRc and so aRc, that is, $[a]_E(R/E)[c]_E$. □

▶ EXERCISE 51 (2.5.8). *Let $A = \mathcal{P}(X)$, $X \neq \varnothing$. Prove:*

a. *Any $S \subseteq A$ has a supremum in the ordering $\subseteq_A$;* $\sup S = \bigcup S$.

b. *Any $S \subseteq A$ has an infimum in $\subseteq_A$;* $\inf S = \bigcap S$ *if $S \neq \varnothing$;* $\inf \varnothing = X$.

PROOF. (a) Let $U = \left\{u \in A \mid s \subseteq_A u,\ \forall\, s \in S\right\}$, i.e., U is the set of all the upper bounds of S according to $\subseteq_A$. Note that $U \neq \varnothing$ since $X \in U$. Now we show that the least element of U exists, and which is $\bigcup S$. Since $s \subseteq_A s \subseteq_A \bigcup S$ for any $s \in S$, we have $\bigcup S \in U$; to see that $\bigcup S$ is the least element of U, take any $u \in U$. Then $s \subseteq_A u$ for all $s \in S$ and so $\bigcup S \subseteq_A u$; therefore, $\sup S = \bigcup S$.

(b) Let $L = \{\ell \in A \mid \ell \subseteq_A s, \ \forall\, s \in S\}$, i.e., L is the set of all the lower bounds of S according to $\subseteq_A$, and $L \neq \varnothing$ since $\varnothing \in L$. We first consider the case that $S \neq \varnothing$, and show that $\sup L = \bigcap S$. Firstly, it is clear that $\bigcap S \in L$; secondly, if $\ell \in L$, then $\ell \subseteq_A s$ for all $s \in S$, so $\ell \subseteq_A \bigcap S$. Therefore, $\inf S = \bigcap S$ if $S \neq \varnothing$.

Finally, let $S = \varnothing$. Then $\inf \varnothing = X$ because for all $B \subseteq X$, $B \subseteq_A C$, $\forall\, C \in \varnothing = S$. Suppose it were not the case. Then there exists $C' \in \varnothing$ such that $B \nsubseteq_A C'$. However, there does not exist such a $C' \in \varnothing$ since there is no element in $\varnothing$. Therefore, all subsets of X, including X itself, is a lower bound of $\varnothing$ according to $\subseteq_A$. Then the greatest element according to $\subseteq_A$ is X. □

► Exercise 52 (2.5.9). *Let* $\mathrm{Fn}(X, Y)$ *be the set of all functions mapping a subset of* X *into* Y *[i.e.,* $\mathrm{Fn}(X, Y) = \bigcup_{Z \subseteq X} Y^Z$*]. Define a relation* $\preccurlyeq$ *in* $\mathrm{Fn}(X, Y)$ *by*

$$f \preccurlyeq g \text{ if and only if } f \subseteq g.$$

a. *Prove that* $\preccurlyeq$ *is an ordering of* $\mathrm{Fn}(X, Y)$.

b. *Let* $F \subseteq \mathrm{Fn}(X, Y)$. *Show that* $\sup F$ *exists if and only if* F *is a compatible system of functions; then* $\sup F = \bigcup F$.

Proof. **(a)** The relation $\preccurlyeq$ is *reflexive* since $f \subseteq f$ for any $f \in \mathrm{Fn}(X, Y)$. If $f \preccurlyeq g$ and $g \preccurlyeq f$, then $f \subseteq g$ and $g \subseteq f$. By the Axiom of Extentionality, we have $f = g$; hence, $\preccurlyeq$ is *antisymmetric*. Finally, let $f \preccurlyeq g$, and $g \preccurlyeq h$, where $f, g, h \in \mathrm{Fn}(X, Y)$. Then $f \subseteq g$ and $g \subseteq h$ implies that $f \preccurlyeq g$; that is, $\preccurlyeq$ is *transitive*. Therefore, $\preccurlyeq$ is an ordering of $\mathrm{Fn}(X, Y)$.

(b) Let $F \subseteq \mathrm{Fn}(X, Y)$. If $\sup F$ exists, there is a function $\sup F \in \mathrm{Fn}(X, Y)$ such that for any $f, g \in \mathrm{Fn}(X, y)$, $f \subseteq \sup F$ and $g \subseteq \sup F$. Suppose $(x, y) \in f$, and $(x, z) \in g$. Then $(x, y) \in \sup F$, and $(x, z) \in \sup F$. Hence, it must be the case that $y = z$; otherwise, $\sup F$ would be not a function. This proves F is a compatible system of functions.

Now suppose F is a compatible system of functions. Then, $\bigcup F$ is a function with $\mathfrak{D}_F = \bigcup\{\mathfrak{D}_f \mid f \in F\} \subseteq X$; therefore, $\bigcup F \in \mathrm{Fn}(X, Y)$. It is easy to see that $\bigcup F$ is an upper bound of F since $f \subseteq \bigcup F \iff f \preccurlyeq \bigcup F$ for any $f \in F$. Finally, let G be any upper bound of F, then $f \subseteq G$ for any $f \in F$; consequently,

$$\left[\bigcup F = \bigcup_{f \in F} f \subseteq G\right] \Longrightarrow \bigcup F \preccurlyeq G,$$

for any upper bound of F. This proves that $\sup F = \bigcup F$. □

► Exercise 53 (2.5.10). *Let* $A \neq \varnothing$*; let* $\mathrm{Pt}(A)$ *be the set of all partitions of* A*. Define a relation* $\preccurlyeq$ *in* $\mathrm{Pt}(A)$ *by*

$$S_1 \preccurlyeq S_2 \text{ if and only if for every } C \in S_1 \text{ there is } D \in S_2 \text{ such that } C \subseteq D.$$

(We say that the partition S_1 *is a* refinement *of the partition* S_2 *if* $S_1 \preccurlyeq S_2$ *holds.)*

a. *Show that* $\preccurlyeq$ *is an ordering.*

b. *Let* $S_1, S_2 \in \mathrm{Pt}(A)$. *Show that* $\{S_1, S_2\}$ *has an infimum. How is the equivalence relation* E_S *related to the equivalence* E_{S_1} *and* E_{S_2}*?*

c. *Let* $T \subseteq \mathrm{Pt}(A)$. *Show that* $\inf T$ *exists.*

d. *Let* $T \subseteq \mathrm{Pt}(A)$. *Show that* $\sup T$ *exists.*

PROOF. (a) It is clear that $\preccurlyeq$ is reflexive. To see $\preccurlyeq$ is antisymmetric, let $S_1 \preccurlyeq S_2$ and $S_2 \preccurlyeq S_1$, where $S_1, S_2 \in \mathrm{Pt}(A)$. Since $S_1 \preccurlyeq S_2$, for every $C_1 \in S_1$ there is $D_2 \in S_2$ such that $C_1 \subseteq D_2$. Suppose that $C_1 \subset D_2$. Since $S_2 \preccurlyeq S_1$, there is $D_1 \in S_1$ such that $D_2 \subseteq D_1$. Then $C_1 \subset D_1$. But then $C_1 \cap D_1 \neq \varnothing$ A contradiction. Hence, $S_1 \subseteq S_2$. Similarly, $S_2 \subseteq S_1$.

To verify that $\preccurlyeq$ is transitive, let $S_1 \preccurlyeq S_2$, and $S_2 \preccurlyeq S_3$, where $S_1, S_2, S_3 \in \mathrm{Pt}(A)$. Then for every $C \in S_1$, there is $D \in S_2$ and $E \in S_3$ such that $C \subseteq D \subseteq E$; that is, $C \subseteq E$. Hence, $S_1 \preccurlyeq S_3$.

(b) Let $S_1, S_2 \in \mathrm{Pt}(A)$. Let $\mathcal{L} = \{S \in \mathrm{Pt}(A) : S \preccurlyeq S_1 \text{ and } S \preccurlyeq S_2\}$. Note that $\mathcal{L} \neq \varnothing$ because $\{\{a\} : a \in A\} \in \mathcal{L}$. We now show

$$M = \{C \cap D : C \in S_1 \text{ and } D \in S_2\}$$

is the greatest element of $\mathcal{L}$. If $m \in M$, then there exist $C \in S_1$ and $D \in S_2$ such that $m = C \cap D$. Then $m \subseteq C$ and $m \subseteq D$; that is, $M \preccurlyeq S_1$ and $M \preccurlyeq S_2$; that is, $M \in \mathcal{L}$.

Pick an arbitrary $N \in \mathcal{L}$. Then for every $n \in N$, there exists $C \in S_1$ such that $n \subseteq C$, and there exists $D \in S_2$ such that $n \subseteq D$; that is, $n \subseteq C \cap D \in M$. Hence, $N \preccurlyeq M$ and so $M = \inf\{S_1, S_2\}$.

(c) The same as (b).

(d) Let $T \subseteq \mathrm{Pt}(A)$. Define $\mathcal{U} = \{S \in \mathrm{Pt}(A) : t \preccurlyeq S \ \forall\, t \in T\}$. Notice that $\mathcal{U} \neq \varnothing$ because $A \in \mathcal{U}$. Now we show that

$$\sup T = \left\{ \bigcup_{C_i \in t_i} C_i : t_i \in T \right\} = P.$$

This can be proved as follows:

- $P \in \mathcal{U}$. For any $C_i \in t_i \in T$, $C_i \subseteq C_i \cup \bigcup_{C_j \in t_j} C_j \in P$, where $j \neq i$; hence $t_i \preccurlyeq P$, $\forall\, t_i \in T$.
- P is the least element of $\mathcal{U}$. Suppose $Q \in \mathcal{U}$. Then $t_i \preccurlyeq Q$, $\forall\, t_i \in T$; then, for any $C_i \in t_i$, there exists $q \in Q$ such that $C_i \subseteq q$, for all $t_i \in T$. But which means that $\bigcup_{C_i \in t_i} C_i \subseteq q, \quad \forall\, t_i \in T$. Hence, $P \preccurlyeq Q$, $\forall\, Q \in \mathcal{U}$. □

► EXERCISE 54 (2.5.11). *Show that if* $(P, <)$ *and* $(Q, \prec)$ *are isomorphic strictly ordered sets and* $<$ *is a linear ordering, then* $\prec$ *is a linear ordering.*

PROOF. Let $h: P \to Q$ be the isomorphism. Pick any $q_1, q_2 \in Q$ with $q_1 \neq q_2$. There exist $p_1, p_2 \in P$ with $p_1 \neq p_2$ such that $q_1 = h(p_1)$ and $q_2 = h(p_2)$. Since $<$ is a linear ordering, p_1 and p_2 are comparable, say, $p_1 < p_2$. Then $h(p_1) = q_1 \prec q_2 = h(p_2)$. □

► EXERCISE 55 (2.5.12). *The identity function on P is an isomorphism between $(P, <)$ and $(P, <)$.*

PROOF. The function $\mathrm{Id}_P : P \to P$ is bijective, and $p_1 < p_2$ iff $\mathrm{Id}_P(p_1) < \mathrm{Id}_P(p_2)$. □

► EXERCISE 56 (2.5.13). *If h is an isomorphism between $(P, <)$ and $(Q, \prec)$, then h^{-1} is an isomorphism between $(Q, \prec)$ and $(P, <)$.*

PROOF. Since $\mathfrak{D}_{h^{-1}} = \mathfrak{R}_h = Q$, and $\mathfrak{R}_{h^{-1}} = \mathfrak{D}_h = P$, the function $h^{-1} : Q \to P$ is bijective. For all $q_1, q_2 \in Q$, there exists unique $p_1, p_2 \in P$ such that $q_1 = h(p_1)$ and $q_2 = h(p_2)$; then

$$q_1 \prec q_2 \iff h(p_1) \prec h(p_2) \iff p_1 < p_2 \iff h^{-1}(q_1) < h^{-1}(q_2). \qquad \square$$

► EXERCISE 57 (2.5.14). *If f is an isomorphism between $(P_1, <_1)$ and $(P_2, <_2)$, and if g is an isomorphism between $(P_2, <_2)$ and $(P_3, <_3)$, then $g \circ f$ is an isomorphism between $(P_1, <_1)$ and $(P_3, <_3)$.*

PROOF. First, $\mathfrak{D}_{g \circ f} = \mathfrak{D}_f \cap f^{-1}[\mathfrak{D}_g] = P_1 \cap f^{-1}[P_2] = P_1$. Next, for every $p_3 \in P_3$, there exists $p_2 \in P_2$ such that $p_3 = g(p_2)$, and for every $p_2 \in P_2$, there exists $p_1 \in P_1$ such that $p_2 = f(p_1)$. Therefore, for every $p_3 \in P_3$, there exists $p_1 \in P_1$ such that $p_3 = g(p_2) = g(f(p_1)) = (g \circ f)(p_1)$. Hence, $g \circ f : P_1 \to P_3$ is surjective.

To see that $g \circ f$ is injective, let $p_1 \neq p_1'$. Then $f(p_1) \neq f(p_1')$, and so $g(f(p_1)) \neq g(f(p_1'))$.

Finally, to see $g \circ f$ is order-preserving, notice that

$$p_1 <_1 p_1' \iff f(p_1) <_2 f(p_1') \iff (g \circ f)(p_1) <_3 (g \circ f)(p_1'). \qquad \square$$

3

NATURAL NUMBERS

3.1 Introduction to Natural Numbers

► Exercise 58 (3.1.1). *$x \subseteq S(x)$ and there is no z such that $x \subset z \subset S(x)$.*

Proof. It is clear that $x \subseteq x \cup \{x\} = S(x)$. Given x, suppose there exists a set z such that $x \subset z$. Then there must exist some set $a \neq \varnothing$ such that $z = x \cup a$. If $a = \{x\}$, then $z = S(x)$; if $a \neq \{x\}$, then there must exist $d \in a$ such that $d \neq x$. Therefore, we have $a \nsubseteq \{x\}$. Consequently, $z = x \cup a \nsubseteq x \cup \{x\} = S(x)(d)$. □

3.2 Properties of Natural Numbers

► Exercise 59 (3.2.1). *Let $n \in \mathbb{N}$. Prove that there is no $k \in \mathbb{N}$ such that $n < k < n + 1$.*

Proof. (Method 1) Let $n \in \mathbb{N}$. Suppose there exists $k \in \mathbb{N}$ such that $n < k$. Then $n \in k$; that is, $n \subset k$ [See Exercise 65]. If $k < n + 1$, then $k \subset n + 1 = S(n)$. That is impossible by Exercise 58.

(Method 2) Suppose there exists k such that $k < n+1$. By Lemma 2.1, $k < n+1$ if and only if $k < n$ or $k = n$. Therefore, it cannot be the case that $n < k$. □

► Exercise 60 (3.2.2). *Use Exercise 59 to prove for all $m, n \in \mathbb{N}$: if $m < n$, then $m + 1 \leqslant n$. Conclude that $m < n$ implies $m + 1 < n + 1$ and that therefore the successor $S(n) = n + 1$ defines a one-to-one function in $\mathbb{N}$.*

Proof. $m < m + 1$ for all $m \in \mathbb{N}$. It follows from Exercise 59 that there is no $n \in \mathbb{N}$ satisfying $m < n < m + 1$. Since $<$ is linear on $\mathbb{N}$, it must be the case that $m + 1 \leqslant n$. Then $m + 1 \leqslant n < n + 1$ implies that $m + 1 < n + 1$. To see $S(n)$ is one-to-one, let $m < n$. Then $S(m) = m+1$, $S(n) = n+1$, and so $m+1 < n+1$. □

► Exercise 61 (3.2.3). *Prove that there is a one-to-one mapping of $\mathbb{N}$ onto a proper subset of $\mathbb{N}$.*

PROOF. Just consider $S: n \mapsto n+1$. By Exercise 60, S is injective. By definition, S is defined on $\mathbb{N}$, i.e., $\mathfrak{D}_S = \mathbb{N}$, and by the following Exercise 62, $\mathfrak{R}_S = \mathbb{N} \smallsetminus \{0\}$. Therefore, $S: \mathbb{N} \to \mathbb{N} \smallsetminus \{0\}$, as desired. □

► EXERCISE 62 (3.2.4). *For every $n \in \mathbb{N}$, $n \neq 0$, there is a unique $k \in \mathbb{N}$ such that $n = k + 1$.*

PROOF. We use the induction principle in Exercise 69 to prove this claim. Let $\mathbf{P}(x)$ be "there is a unique $k \in \mathbb{N}$ such that $x = k+1$". It is clear that $\mathbf{P}(1)$ holds since $1 = 0+1$. The uniqueness of $0 = \varnothing$ is from Lemma 3.1 in Chapter 1. Now suppose that $\mathbf{P}(n)$ holds and consider $\mathbf{P}(n+1)$. We have $n+1 = (k+1)+1$ by the induction assumption $\mathbf{P}(n)$. Note that $k+1 = S(k) \in \mathbb{N}$. Let $k+1 = k'$. The uniqueness of k implies that k' is unique. We thus complete the proof. □

► EXERCISE 63 (3.2.5). *For every $n \in \mathbb{N}$, $n \neq 0, 1$, there is a unique $k \in \mathbb{N}$ such that $n = (k+1)+1$.*

PROOF. We know from Exercise 62 that for every nonzero $n \in \mathbb{N}$ there is a unique $k' \in \mathbb{N}$, such that $n = k' + 1$. Now consider $k' \in \mathbb{N}$. If $n \neq 1$, then $k' \neq 0$. Therefore, we can impose the result of Exercise 62 on k'; that is, there is a unique $k \in \mathbb{N}$ such that $k' = k+1$. Combining these above two steps, we know for all $n \in \mathbb{N}$, $n \neq 0, 1$, there is a unique $k \in \mathbb{N}$ such that $n = (k+1)+1$. □

► EXERCISE 64 (3.2.6). *Prove that each natural number is the set of all smaller natural numbers; i.e., $n = \{m \in \mathbb{N} : m < n\}$.*

PROOF. Let $\mathbf{P}(x)$ denote "$x = \{m \in \mathbb{N} : m < x\}$". It is evident that $\mathbf{P}(0)$ holds trivially. Assume that $\mathbf{P}(n)$ holds and let us consider $\mathbf{P}(n+1)$. We have

$$n+1 = n \cup \{n\} = \{m \in \mathbb{N} : m < n\} \cup \{n\} = \{m \in \mathbb{N} : m < n+1\}.$$ □

► EXERCISE 65 (3.2.7). *For all $m, n \in \mathbb{N}$, $m < n$ if and only if $m \subset n$.*

PROOF. Let $\mathbf{P}(x)$ be the property "$m < x$ if and only if $m \subset x$". It is clear that $\mathbf{P}(0)$ holds trivially. Assume that $\mathbf{P}(n)$ holds. Let us consider $\mathbf{P}(n+1)$. First let $m < n+1$; then $m < n$ or $m = n$. If $m < n$, then $m \subset n \subset (n+1)$ by the induction assumption $\mathbf{P}(n)$; if $m = n$, then $m = n \subset (n+1)$, too. Now assume that $m \subset (n+1)$. Then either $m = n$ or $m \subset n$. We get $m < n+1$ in either case. □

► EXERCISE 66 (3.2.8). *Prove that there is no function $f: \mathbb{N} \to \mathbb{N}$ such that for all $n \in \mathbb{N}$, $f(n) > f(n+1)$. (There is no infinite decreasing sequence of natural numbers.)*

PROOF. Suppose there were such a function f. Then $\varnothing \neq \{f(n) \in \mathbb{N} : n \in \mathbb{N}\} \subseteq \mathbb{N}$. Because $(\mathbb{N}, <)$ is well-ordered, the set $\{f(n) \in \mathbb{N} : n \in \mathbb{N}\}$ has a least element α; that is, there is $m \in \mathbb{N}$ such that $f(m) = \alpha$. But $f(m+1) < f(m) = \alpha$, which contradicts the assumption that α is the least element. □

► EXERCISE 67 (3.2.9). *If* $X \subseteq \mathbb{N}$*, then* $\left\langle X, < \cap X^2 \right\rangle$ *is well-ordered.*

PROOF. Let $Y \subseteq X$ be nonempty. Y has a least element y when Y is embedded in $\mathbb{N}$. But clearly y is still a least element of Y when Y is embedded in $X \subseteq \mathbb{N}$. □

► EXERCISE 68 (3.2.10). *In Exercise 49, let* $A = \mathbb{N}$*,* $b = \mathbb{N}$*. Prove that* $\prec$ *as defined there is a well-ordering of* $B = \mathbb{N} \cup \{\mathbb{N}\}$*. Notice that* $x \prec y$ *if and only if* $x \in y$ *holds for all* $x, y \in B$*.*

PROOF. The relation $\prec$ in $B = \mathbb{N} \cup \{\mathbb{N}\}$ is defined as

$$x \prec y \iff (x, y \in \mathbb{N} \text{ and } x < y) \text{ or } (x \in \mathbb{N} \text{ and } y = \mathbb{N}) \iff x \in y.$$

Let $X \subseteq B = \mathbb{N} \cup \{\mathbb{N}\}$ be nonempty. There are two cases:

- If $\mathbb{N} \notin X$, then $X \subseteq \mathbb{N}$, and so X has a least element since $(\mathbb{N}, <)$ is well-ordered.
- If $\mathbb{N} \in X$, then $X = Y \cup \{\mathbb{N}\}$, where $Y \subseteq \mathbb{N}$. Hence, Y has a least element α. But $\alpha \prec \mathbb{N}$ since $\alpha \in \mathbb{N}$; that is, α is the least element of X. □

► EXERCISE 69 (3.2.11). *Let* $\mathbf{P}(x)$ *be a property. Assume that* $k \in \mathbb{N}$ *and*

a. $\mathbf{P}(k)$ *holds.*

b. *For all* $n \geqslant k$*, if* $\mathbf{P}(n)$ *then* $\mathbf{P}(n+1)$*.*

Then $\mathbf{P}(n)$ *holds for all* $n \geqslant k$*.*

PROOF. If $k = 0$, then this is the original Induction Principle. So assume that $k > 0$ and $\mathbf{P}(k)$ holds. Then, by Exercise 62, there is a unique $k' \in \mathbb{N}$ such that $k' + 1 = k$. Define

$$B = \{n \in \mathbb{N} : n \leqslant k'\}, \quad \text{and} \quad C = \left\{n \in \mathbb{N} : n \geqslant k \text{ and } \mathbf{P}(n)\right\}.$$

Notice that $B \cap C = \varnothing$.

We now show that $A = B \cup C$ is inductive. Obviously, $0 \in A$. If $n \in B$, then either $n < k'$ and so $n + 1 \in B$, or $n = k'$ and so $n + 1 = k \in C$. If $n \in C$, then $n + 1 \in C$ by assumption. Hence, $\mathbb{N} = A$ (since $A \subseteq \mathbb{N}$), and so $\{n \in \mathbb{N} : n \geqslant k\} = \mathbb{N} \setminus B = C$. □

► EXERCISE 70 (3.2.12, Finite Induction Principle). *Let* $\mathbf{P}(x)$ *be a property. Assume that* $k \in \mathbb{N}$ *and*

a. $\mathbf{P}(0)$.

b. *For all* $n < k$*,* $\mathbf{P}(n)$ *implies* $\mathbf{P}(n+1)$*.*

Then $\mathbf{P}(n)$ *holds for all* $n \leqslant k$*.*

PROOF. Suppose there were $n < k$ such that $\neg\mathbf{P}(n)$. Then it must be the case that $\neg\mathbf{P}(m)$, where $m + 1 = n$. Thus, $X = \{a \in \mathbb{N} : a < k \text{ and } \neg\mathbf{P}(a)\} \neq \varnothing$, and so X has a least element, α. Also $\alpha \neq 0$ since $\mathbf{P}(0)$ holds by assumption.

However, if $\neg\mathbf{P}(\alpha)$, then $\neg\mathbf{P}(\beta)$, where $\beta + 1 = \alpha$, is also true. But $\beta < \alpha$, which contradicts the assumption that α is the least element of X. Therefore, $\mathbf{P}(n)$ holds for all $n < k$.

To see $\mathbf{P}(k)$ holds, too, notice that there exists $m \in \mathbb{N}$ and $m < k$ such that $m + 1 = k$ (by Exercise 62). Because we have shown that $\mathbf{P}(m)$ holds, $\mathbf{P}(m + 1) = \mathbf{P}(k)$ also holds. □

► EXERCISE 71 (3.2.13, Double Induction). *Let $\mathbf{P}(x, y)$ be a property. Assume*

$$\text{If } \mathbf{P}(k, \ell) \text{ holds for all } k, \ell \in \mathbb{N} \text{ such that } k < m \text{ or } (k = m \text{ and } \ell < n), \text{ then } \mathbf{P}(m, n) \text{ holds.} \tag{$**$}$$

Conclude that $\mathbf{P}(m, n)$ holds for all $m, n \in \mathbb{N}$.

PROOF. We proceed by induction on m. Fix $n \in \mathbb{N}$. Then $\mathbf{P}(m, n)$ is true for all $m \in \mathbb{N}$ by the second version of Induction Principle. Now for every $m \in \mathbb{N}$, (m, n) is true for all n by the second version of Induction Principle. Hence, $\mathbf{P}(m, n)$ holds for all $m, n \in \mathbb{N}$. □

3.3 THE RECURSION THEOREM

► EXERCISE 72 (3.3.1). *Let f be an infinite sequence of elements of A, where A is ordered by $\prec$. Assume that $f_n \prec f_{n+1}$ for all $n \in \mathbb{N}$. Prove that $n < m$ implies $f_n \prec f_m$ for all $n, m \in \mathbb{N}$.*

PROOF. We proceed by induction on m in the form of Exercise 69. For an arbitrary $n \in \mathbb{N}$, let $\mathbf{P}(x)$ denote "$f_n \prec f_x$ if $n < x$". Let $k = n + 1$. then $\mathbf{P}(k)$ holds since $f_n \prec f_{n+1} = f_k$ by assumption.

Suppose that $\mathbf{P}(m)$ holds, where $m \geqslant k$, and consider $\mathbf{P}(m + 1)$. Since $f_m \prec f_{m+1}$ by the assumption of the exercise, and $f_n \prec f_m$ by induction hypothesis of $\mathbf{P}(m)$, we have $f_n \prec f_{m+1}$.

Using the Induction Principle in the form of Exercise 69, we conclude that $\mathbf{P}(m)$ holds for all $m \geqslant k = n + 1 > n$. □

► EXERCISE 73 (3.3.2). *Let $(A, \prec)$ be a linearly ordered set and $p, q \in A$. We say that q is a* successor *or p if $p \prec q$ and there is no $r \in A$ such that $p \prec r \prec q$. Note that each $p \in A$ can have at most one successor. Assume that $(A, \prec)$ is nonempty and has the following properties:*

a. *Every $p \in A$ has a successor.*

b. *Every nonempty subset of A has a $\prec$-least element.*

c. *If $p \in A$ is not the $\prec$-least element of A, then p is a successor of some $q \in A$.*

Prove that $(A, \prec)$ is isomorphic to $(\mathbb{N}, <)$. Show that the conclusion need not hold if one of the conditions (a)–(c) is omitted.

PROOF. We first show that each $p \in A$ can have at most one successor. If q_1 and q_2 are both the successors of p, and $q_1 \neq q_2$, say, $q_1 \prec q_2$, then $p \prec q_1 \prec q_2$, in contradiction to the assumption that q_2 is a successor of p.

Let a be the least element of A (by (b)) and let $g(x, n)$ be the successor of x (for all n). Then $a \in A$ and $g\colon A \times \mathbb{N} \to A$ is well defined by (a). The Recursion Theorem guarantees the existence of a function $f\colon \mathbb{N} \to A$ such that

- $f_0 = a =$ the least element of A;
- $f_{n+1} = g(f_n, n) =$ the successor of f_n.

By definition, $f_n \prec f_{n+1}$ for all $n \in \mathbb{N}$; by Exercise 72 $f_n \prec f_m$ whenever $n < m$. Consequently, f is injective. It remains to show that f is surjective.

If not, $A \smallsetminus \mathfrak{R}_f \neq \varnothing$; let p be the least element of $A \smallsetminus \mathfrak{R}_f$. Then $p \neq a$, the least element of A. It follows from (c) that there exists $q \in A$ such that p is the successor of q. There exists $m \in \mathbb{N}$ such that $f_m = q$; for otherwise $q \in A \smallsetminus \mathfrak{R}_f$ and $q \prec p$. Hence, $f_{m+1} = p$ by the recursive condition. Consequently, $p \in \mathfrak{R}_f$, a contradiction. □

► EXERCISE 74 (3.3.3). *Give a direct proof of Theorem 3.5 in a way analogous to the proof of the Recursion Theorem.*

PROOF. We first show that there exists a unique infinite sequence of finite sequences $\langle F^n \in \mathrm{Seq}(S) : n \in \mathbb{N} \rangle = F$ satisfying

$$F^0 = \langle\,\rangle, \tag{A}$$

$$F^{n+1} = G\left(F^n, n\right), \tag{B}$$

where

$$G\left(F^n, n\right) = \begin{cases} F^n \cup \left\{\left\langle n, g(F^n_0, \dots, F^n_{n-1})\right\rangle\right\} & \text{if } F^n \text{ is a sequence of length } n \\ \langle\,\rangle & \text{otherwise.} \end{cases}$$

It is easy to see that $G\colon \mathrm{Seq}(S) \times \mathbb{N} \to \mathrm{Seq}(S)$.

Let $T\colon (m+1) \to \mathrm{Seq}(S)$ be an m-step computation based on $F_0 = \langle\,\rangle$ and G. Then

$$T^0 = \langle\,\rangle, \quad \text{and} \quad T^{k+1} = G(T^k, k) \text{ for } 0 \leqslant k < m.$$

Notice that $T \in \mathcal{P}(\mathbb{N} \times \mathrm{Seq}(S))$. Let

$$\mathcal{F} = \left\{T \in \mathcal{P}(\mathbb{N} \times \mathrm{Seq}(S)) : T \text{ is an } m\text{-step computation for some } m \in \mathbb{N}\right\}.$$

Let $F = \bigcup \mathcal{F}$. Then

- *F is a function.* We need only to prove the system of functions $\mathcal{F}$ is compatible. Let $T, U \in \mathcal{F}$, $\mathfrak{D}_T = m \in \mathbb{N}$, $\mathfrak{D}_U = n \in \mathbb{N}$. Assume, e.g., $m \leqslant n$; then $m \subseteq n \Longrightarrow m \cap n = m$, and it suffices to show that

$$\left\langle T_0^k, \ldots, T_{k-1}^k \right\rangle = T^k = U^k = \left\langle U_0^k, \ldots, U_{k-1}^k \right\rangle$$

  for all $k < m$. This can be done by induction [Exercise 70]. Surely, $T^0 = \langle\,\rangle = U^0$. Next let k be such that $k + 1 < m$, and assume $T^k = U^k$. Then

$$T^{k+1} = T^k \cup \left\{ \left\langle k, g(T^k) \right\rangle \right\} = U^k \cup \left\{ \left\langle k, g(U^k) \right\rangle \right\} = U^{k+1}.$$

  Thus, $T^k = U^k$ for all $k < m$.

- *$\mathfrak{D}_F = \mathbb{N}$; $\mathfrak{R}_F \subseteq \mathrm{Seq}(S)$.* We know that $\mathfrak{D}_F = \bigcup \left\{ \mathfrak{D}_T \mid T \in \mathcal{F} \right\} \subseteq \mathbb{N}$, and $\mathfrak{R}_F \subseteq \mathbb{N}$. To show that $\mathfrak{D}_F = \mathbb{N}$, it suffices to prove that for each $n \in \mathbb{N}$ there is an n-step computation T. We use the Induction Principle. Clearly, $T = \{\langle 0, \langle\,\rangle \rangle\}$ is a 0-step computation.

  Assume that T is an n-step computation. Then the following function T_+ on $(n + 1) + 1$ is an $(n + 1)$-step computation:

$$\begin{cases} T_+^k = T^k, & \text{if } k \leqslant n \\ T_+^{n+1} = T^n \cup \left\{ \langle n, g(T^n) \rangle \right\}. \end{cases}$$

  We conclude that each $n \in \mathbb{N}$ is in the domain of some computation $T \in \mathcal{F}$, so $\mathbb{N} \subseteq_{T \in \mathcal{F}} \mathfrak{D}_T = \mathfrak{D}_F$.

- *F satisfies condition (A) and (B).* Clearly, $F_0 = \langle\,\rangle$ since $T^0 = \langle\,\rangle$ for all $T \in \mathcal{F}$. To show that $F_{n+1} = G(F_n, n)$ for any $n \in \mathbb{N}$, let T be an $(n + 1)$-step computation; then $T^k = F_k$ for all $k \in \mathfrak{D}_T$, so $F_{n+1} = T^{n+1} = G(T^n, n) = G(F_n, n)$.

  Let $H : \mathbb{N} \to \mathrm{Seq}(S)$ be such that

$$H_0 = \langle\,\rangle, \tag{A$'$}$$

and

$$H_{n+1} = G(H_n, n) \quad \forall\, n \in \mathbb{N}. \tag{B$'$}$$

We show that $F_n = H_n$, $\forall\, n \in \mathbb{N}$, again using induction. Certainly $F_0 = H_0$. If $F_n = H_n$, then $F_{n+1} = G(F_n, n) = G(H_n, n) = H_{n+1}$; therefore, $F = H$, as claimed.

Now we can define a function f by

$$f = \bigcup_{n \in \mathbb{N}} F^n. \qquad \square$$

► EXERCISE 75 (3.3.4). *Derive the "parametric" version of the Recursion Theorem: Let* $a\colon P \to A$ *and* $g\colon P \times A \times \mathbb{N} \to A$ *be functions. There exists a unique function* $f\colon P \times \mathbb{N} \to A$ *such that*

a. $f(p,0) = a(p)$ *for all* $p \in P$*;*

b. $f(p,n+1) = g\big(p, f(p,n), n\big)$ *for all* $n \in \mathbb{N}$ *and* $p \in P$*.*

PROOF. Define $G\colon A^P \times \mathbb{N} \to A^P$ by

$$G(x,n)(p) = g(p, x(p), n)$$

for $x \in A^P$ and $n \in \mathbb{N}$. Define $F\colon \mathbb{N} \to A^P$ by recursion:

$$F_0 = a \in A^P, \quad F_{n+1} = G(F_n, n). \tag{3.1}$$

Then, by the Recursion Theorem, there exists a unique $F\colon \mathbb{N} \to A^P$ satisfying (3.1). Now let $f(p,n) = F_n(p)$. Then

- $f(p,0) = F_0(p) = a(p)$, and
- $f(p,n+1) = F_{n+1}(p) = G(F_n,n)(p) = g(p, F_n(p), n) = g(p, f(p,n), n)$. □

► EXERCISE 76 (3.3.5). *Prove the following version of the Recursion Theorem:*

Let g *be a function on a subset of* $A \times \mathbb{N}$ *into* A*,* $a \in A$*. Then there is a unique sequence* f *of elements of* A *such that*

a. $f_0 = a$*;*

b. $f_{n+1} = g(f_n, n)$ *for all* $n \in \mathbb{N}$ *such that* $(n+1) \in \mathfrak{D}_f$*;*

c. f *is either an infinite sequence or* f *is a finite sequence of length* $k+1$ *and* $g(f_k, k)$ *is undefined.*

PROOF. Let $\overline{A} = A \cup \{\overline{a}\}$ where $\overline{a} \notin A$. Define $\overline{g}\colon \overline{A} \times \mathbb{N} \to \overline{A}$ as follows:

$$\overline{g}(x,n) = \begin{cases} g(x,n) & \text{if defined} \\ \overline{a} & \text{otherwise.} \end{cases} \tag{3.2}$$

Then, by the Recursion Theorem, there exists a unique infinite sequence $\overline{f}\colon \mathbb{N} \to \overline{A}$ such that

$$\overline{f}_0 = a, \quad \overline{f}_{n+1} = \overline{g}(\overline{f}_n, n).$$

If $\overline{f}_\ell = \overline{a}$ for some $\ell \in \mathbb{N}$, consider $\overline{f} \restriction \ell$ for the least such ℓ. □

► EXERCISE 77 (3.3.6). *Prove: If* $X \subseteq \mathbb{N}$*, then there is a one-to-one (finite or infinite) sequence* f *such that* $\mathfrak{R}_f = X$*.*

PROOF. Define $g\colon X \times \mathbb{N} \to X$ by

$$g(x,n) = \min\{y \in X : y > x\}.$$

Let $a = \min X$. Then, by Exercise 76, there exists a unique function f satisfying $f_0 = a$ and $f_{n+1} = g(f_n, n)$.

For every $m \in \mathbb{N}$, we have $f_{m+1} \geqslant f_m + 1 > f_m$; hence, f is injective. It follows from the previous exercise that f is surjective. □

3.4 Arithmetic of Natural Numbers

► Exercise 78 (3.4.1). *Prove the associative low of addition:* $(k + m) + n = k + (m + n)$ *for all* $k, m, n \in \mathbb{N}$.

Proof. We use induction on n. So fix $k, m \in \mathbb{N}$. If $n = 0$, then

$$(k + m) + 0 = k + m,$$

and

$$k + (m + 0) = k + m.$$

Assume that $(k + m) + n = k + (m + n)$ and consider $n + 1$:

$$\begin{aligned}(k + m) + (n + 1) &= [(k + m) + n] + 1\\ &= [k + (m + n)] + 1\\ &= k + [(m + n) + 1]\\ &= k + [m + (n + 1)].\end{aligned}$$

□

► Exercise 79 (3.4.2). *If* $m, n, k \in \mathbb{N}$, *then* $m < n$ *if and only if* $m + k < n + k$.

Proof. We first need to prove the following proposition: for any $m, n \in \mathbb{N}$,

$$m < n \iff m + 1 < n + 1. \tag{3.3}$$

The "$\Longrightarrow$" half has been proved in Exercise 60, so we need only to show the "$\Longleftarrow$" part. Assume that $m + 1 < n + 1$. Then $m < m + 1 \leqslant n$. Hence, $m < n$.

For the "$\Longrightarrow$" half we use induction on k. Consider fixed $m, n \in \mathbb{N}$ with $m < n$. Clearly, $m < n \iff m + 0 < n + 0$. Assume that $m < n \Longrightarrow m + k < n + k$. Then by (3.3), $(m + k) + 1 < (n + k) + 1$, i.e., $m + (k + 1) < n + (k + 1)$.

For the "$\Longleftarrow$" half we use the trichotomy law and the "$\Longrightarrow$" half. If $m + k < n + k$, then we cannot have $m = n$ (lest $n + k < n + k$) nor $n < m$ (lest $n + k < m + k < n + k$). The only alternative is $m < n$. □

► Exercise 80 (3.4.3). *If* $m, n \in \mathbb{N}$ *then* $m \leqslant n$ *if and only if there exists* $k \in \mathbb{N}$ *such that* $n = m + k$. *This* k *is unique, so we can denote it* $n - m$, *the* difference *of* n *and* m.

Proof. For the "$\Longrightarrow$" half we use induction on n. If $n = 0$, the proposition trivially hods since there is no natural number $m < 0$. Assume that $m < n$ implies that there exists a unique $k_{m,n} \in \mathbb{N}$ such that $m + k_{m,n} = n$. Now consider $n+1$. If $m < n+1$, then $m = n$ or $m < n$. If $m = n$, let $k_{m,n+1} = 1$ and so $m + k_{m,n+1} = n+1$; if $m < n$, then by the induction hypothesis, there exists a unique $k_{m,n} \in \mathbb{N}$ such that $m + k_{m,n} = n$. Let $k_{m,n+1} = k_{m,n} + 1$. Then

$$m + k_{m,n+1} = m + (k_{m,n} + 1) = (m + k_{m,n}) + 1 = n + 1.$$

For the "$\Longleftarrow$" half we use induction on k. If $k = 0$, it is obvious that $m = n$. Now assume that $m + k = n$ implies that $m \leqslant n$. Let us suppose that for all $m, n \in \mathbb{N}$ there exists a unique $k+1$ such that $m + (k+1) = n$. Then by Exercise 79 we have

$$\begin{aligned} 0 < k+1 &\Longrightarrow m + 0 < m + (k+1) \\ &\Longrightarrow m < n. \end{aligned}$$ □

► EXERCISE 81 (3.4.4). *There is a unique function* $\star$ *(*multiplication*) from* $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$ *such that*

$$m \star 0 = 0 \quad \textit{for all } m \in \mathbb{N};$$
$$m \star (n+1) = m \star n + m \quad \textit{for all } m, n \in \mathbb{N}.$$

PROOF. We use the parametric version of the Recursion Theorem. Let $a\colon \mathbb{N} \to \mathbb{N}$ be defined as $a(p) = 0$, and $g\colon \mathbb{N} \times \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ be defined as $g(p, x, n) = x + p$. Then, there exists a unique function $\star\colon \mathbb{N} \to \mathbb{N}$ such that

$$m \star 0 = a(m) = 0,$$

and

$$m \star (n+1) = g(m, m \star n, n) = m \star n + m.$$ □

► EXERCISE 82 (3.4.5). *Prove that multiplication is commutative, associative, and distributive over addition.*

PROOF. ($\cdot$ is commutative) We first show that 0 commutes by showing $0 \cdot m = 0$ (since $m \cdot 0 = 0$) for all $m \in \mathbb{N}$. Clearly, $0 \cdot 0 = 0$, and if $0 \cdot m = 0$, then

$$0 \cdot (m+1) = 0 \cdot m + 0 = 0.$$

Let us now assume that n commutes, and let us show that $n+1$ commutes. We prove, by induction on m, that

$$m \cdot (n+1) = (n+1) \cdot m \quad \text{for all } m \in \mathbb{N}. \tag{3.4}$$

If $m = 0$, then (3.4) holds, as we have already shown. Thus let us assume that (3.4) holds for m, and let us prove that

$$(m+1) \cdot (n+1) = (n+1) \cdot (m+1). \tag{3.5}$$

We derive (3.5) as follows:

$$\begin{aligned}(m+1)\cdot(n+1) = [(m+1)\cdot n] + (m+1) &= [n\cdot(m+1)] + (m+1)\\ &= (n\cdot m + n) + (m+1)\\ &= (n\cdot m + m) + (n+1)\\ &= (m\cdot n + m) + (n+1)\\ &= m\cdot(n+1) + (n+1)\\ &= (n+1)\cdot m + (n+1)\\ &= (n+1)\cdot(m+1).\end{aligned}$$

($\cdot$ is distributive over addition) We show that for all $m, n, p \in \mathbb{N}$,

$$m\cdot(n+p) = m\cdot n + m\cdot p. \tag{3.6}$$

Fix $m, n \in \mathbb{N}$. We use induction on p. It is clear that $m\cdot(n+0) = m\cdot n = m\cdot n + 0 = m\cdot n + m\cdot 0$. Now assume that (3.6) holds for p, and let us consider $p+1$:

$$\begin{aligned}m\cdot[n+(p+1)] &= m\cdot[(n+p)+1]\\ &= m\cdot(n+p) + m\\ &= m\cdot n + m\cdot p + m\\ &= m\cdot n + (m\cdot p + m)\\ &= m\cdot n + m\cdot(p+1).\end{aligned}$$

($\cdot$ is associative) Fix $m, n \in \mathbb{N}$. We use induction on p. Clearly, $m\cdot(n\cdot 0) = m\cdot 0 = 0$, and $(m\cdot n)\cdot 0 = 0$ as well. Now suppose that

$$m\cdot(n\cdot p) = (m\cdot n)\cdot p.$$

Then

$$\begin{aligned}m\cdot[n\cdot(p+1)] &= m\cdot(n\cdot p + n)\\ &= m\cdot(n\cdot p) + m\cdot n\\ &= (m\cdot n)\cdot p + m\cdot n\\ &= (m\cdot n)\cdot(p+1).\end{aligned}$$ □

► EXERCISE 83 (3.4.6). *If $m, n \in \mathbb{N}$ and $k > 0$, then $m < n$ if and only if $m\cdot k < n\cdot k$.*

PROOF. For the "$\Longrightarrow$" half we fix $m, n \in \mathbb{N}$ and use induction on k. Clearly, $m\cdot 1 < n\cdot 1$ since

$$m\cdot 1 = m\cdot(0+1) = m\cdot 0 + m = m,$$

and similarly for $n\cdot 1$. Let us assume that $m < n$ implies $m\cdot k < n\cdot k$ with $k > 0$, and let us consider $k+1$:

$$\begin{aligned} m \cdot (k+1) &= m \cdot k + m \\ &< n \cdot k + m \\ &< n \cdot k + n \\ &= n \cdot (k+1), \end{aligned}$$

where the inequalities follow from Exercise 79.

The other half then follows exactly as in Exercise 79. □

► EXERCISE 84 (3.4.7). *Define exponentiation of nature numbers as follows:*

$$m^0 = 1 \quad \textit{for all } m \in \mathbb{N} \textit{ (in particular, } 0^0 = 1\textit{);}$$
$$m^{n+1} = m^n \cdot m \quad \textit{for all } m, n \in \mathbb{N} \textit{ (in particular, } 0^n = 0 \textit{ for } n > 0\textit{).}$$

Prove the usual laws of exponents.

PROOF. We show that $m^{n+p} = m^n \cdot m^p$ for all $m, n, p \in \mathbb{N}$ using induction on p. It is evident that

$$m^{n+0} = m^n = m^n \cdot 1 = m^n \cdot m^0,$$

so let us assume $m^{n+p} = m^n \cdot m^p$ and consider $p + 1$:

$$\begin{aligned} m^{n+(p+1)} &= m^{(n+p)+1} \\ &= m^{n+p} \cdot m \\ &= (m^n \cdot m^p) \cdot m \\ &= m^n \cdot (m^p \cdot m) \\ &= m^n \cdot m^{p+1}. \end{aligned}$$

□

3.5 OPERATIONS AND STRUCTURES

► EXERCISE 85 (3.5.1). *Which of the following sets are closed under operations of addition, subtraction, multiplication, and division of real number?*

a. *The set of all positive integers.*

b. *The set of all integers.*

c. *The set of all rational numbers.*

d. *The set of all negative rational numbers.*

e. *The empty set.*

SOLUTION. See the following table:

	$+$	$-$	$\times$	division of real numbers
(a)	Yes	No	Yes	No
(b)	Yes	Yes	Yes	No
(c)	Yes	Yes	Yes	No
(d)	Yes	No	Yes	No
(e)	Yes	Yes	Yes	Yes

□

► EXERCISE 86 (3.5.4). *Let $A \neq \varnothing$, $B = \mathcal{P}(A)$. Show that $(B, \cup_B, \cap_B)$ and $(B, \cap_B, \cup_B)$ are isomorphic structures.*

PROOF. Define a function $h\colon B \to B$ as $h(x) = B \smallsetminus x$. It is evident that h is injective. To see h is surjective, notice that if $y \in B$, then $y \subseteq A$ and so $A \smallsetminus y \in B$; hence $h(A \smallsetminus y) = y$.

Since $B = \mathcal{P}(A)$, both $\cup_B$ and $\cap_B$ are well defined. For all $x, y \in B$,

$$h(x \cup_B y) = B \smallsetminus (x \cup_B y) = (B \smallsetminus x) \cap_B (B \smallsetminus y) = h(x) \cap_B h(y),$$

and similarly, $h(x \cap_B y) = h(x) \cup_B h(y)$. □

► EXERCISE 87 (3.5.5). *Refer to Example 5.7 for notation.*

a. *There is a real number $a \in A$ such that $a + a = a$ (namely, $a = 0$). Prove from this that there is $a' \in A'$ such that $a' \times a' = a'$. Find this a'.*

b. *For every $a \in A$ there is $b \in A$ such that $a + b = 0$. Show that for every $a' \in A'$ there is $b' \in A'$ such that $a' \times b' = 1$. Find this b'.*

PROOF. It is from Example 5.7 that $(A, \leqslant_A, +) \cong (A', \leqslant_{A'}, \times)$, and the isomorphism $h\colon A \to A'$ is $h(x) = e^x$.

(a) If $a + a = a$, then

$$h(a + a) = e^{a+a} = e^a \times e^a = e^a.$$

Hence, there exists $a' = e^a = e^0 = 1$ such that $a' \times a' = a'$.

(b) For every $a' \in A'$, there exists a unique $a \in A$ such that $h(a) = a'$. Let $b \in A$ such that $a + b = 0$. Then

$$h(a + b) = h(a) \times h(b) = a' \times h(b) = e^0 = 1.$$

Hence, for every $a' \in A'$, there exists $b' = h(b)$ such that $a' \times b' = 1$. □

► EXERCISE 88 (3.5.6). *Let $\mathbb{Z}^+$ and $\mathbb{Z}^-$ be, respectively, the sets of all positive and negative integers. Show that $(\mathbb{Z}^+, <, +)$ is isomorphic to $(\mathbb{Z}^-, >, +)$ (where $<$ is the usual ordering of integers).*

PROOF. Define $h\colon \mathbb{Z}^+ \to \mathbb{Z}^-$ by letting $h(z) = -z$. Then h is bijective. Let $z_1, z_2 \in \mathbb{Z}^+$. Then $z_1 < z_2$ iff $-z_1 > -z_2$ iff $h(z_1) > h(z_2)$. It is evident that both operations on $\mathbb{Z}^+$ and $\mathbb{Z}^-$ are well defined, and $h(z_1 + z_2) = -(z_1 + z_2) = (-z_1) + (-z_2) = h(z_1) + h(z_2)$. Thus, $(\mathbb{Z}^+, <, +) \cong (\mathbb{Z}^-, >, +)$. □

► EXERCISE 89 (3.5.14). *Construct the sets C_0, C_1, C_2, and C_3 in Theorem 5.10 for*

a. $\mathfrak{A} = (\mathbb{R}, S)$ *and* $C = \{0\}$.

b. $\mathfrak{A} = (\mathbb{R}, +, -)$ *and* $C = \{0, 1\}$.

PROOF. (a) $C_0 = C = \{0\}$, $C_1 = C_0 \cup S[C_0] = \{0\} \cup \{1\} = \{0, 1\}$, $C_2 = C_1 \cup S[C_1] = \{0, 1\} \cup \{1, 2\} = \{0, 1, 2\}$, and $C_3 = C_2 \cup S[C_2] = \{0, 1, 2\} \cup \{1, 2, 3\} = \{0, 1, 2, 3\}$.

(b) $C_0 = C = \{0, 1\}$, $C_1 = C_0 \cup +[C_0^2] \cup -[C_0^2] = \{0, 1\} \cup \{0, 1, 2\} \cup \{-1, 0, 1\} = \{-1, 0, 1, 2\}$, $C_2 = C_1 \cup +[C_1^2] \cup -[C_1^2] = \{-1, 0, 1, 2\} \cup \{-2, -1, 0, 1, 2, 3, 4\} \cup \{-3, -2, -1, 0, 1, 2, 3\} = \{-3, -2, -1, 0, 1, 2, 3, 4\}$, and $C_3 = C_2 \cup +[C_2^2] \cup -[C_2^2] = \{-7, -6, \cdots, 7, 8\}$. □

4

FINITE, COUNTABLE, AND UNCOUNTABLE SETS

4.1 Cardinality of Sets

▶ Exercise 90 (4.1.1). *Prove Lemma 1.5.*

a. *If* $|A| \leqslant |B|$ *and* $|A| = |C|$*, then* $|C| \leqslant |B|$.

b. *If* $|A| \leqslant |B|$ *and* $|B| = |C|$*, then* $|A| \leqslant |C|$.

c. $|A| \leqslant |A|$.

d. *If* $|A| \leqslant |B|$ *and* $|B| \leqslant |C|$*, then* $|A| \leqslant |C|$.

Proof. **(a)** If $|A| = |C|$, then $|C| = |A|$, and so there is a bijection $f: C \to A$. Since $|A| \leqslant |B|$, there is an injection $g: A \to B$. Then $g \circ f: C \to B$ is an injection and so $|C| \leqslant |B|$.

(b) Since $|A| \leqslant |B|$, there is a bijection $g: A \to \mathcal{R}_g$, where $\mathcal{R}_g \subseteq B$ is the image of A under g. Since $|B| = |C|$, there is a bijection $f: B \to C$. Let $h := f \restriction \mathcal{R}_g$ be the restriction of f on $\mathcal{R}_g$. Let $D' := \mathcal{R}_h \subseteq C$. Then $h: \mathcal{R}_g \to D'$ is a bijection. To prove $|A| \leqslant |C|$, consider $h \circ g: A \to D'$. This is a one-to-one correspondence from A to $D' \subseteq C$.

(c) This claim follows two facts: (i) Id_A is a one-to-one mapping of A onto A, and (ii) $A \subseteq A$.

(d) Since $|A| \leqslant |B|$, there is a bijection $f: A \to \mathcal{R}_f$, where $\mathcal{R}_f \subseteq B$. Since $|B| \leqslant |C|$, there is a bijection $g: B \to \mathcal{R}_g$, where $\mathcal{R}_g \subseteq C$. Let $h := g \restriction \mathcal{R}_f$. Then $h \circ f: A \to C$ is a injection and so $|A| \leqslant |C|$. □

▶ Exercise 91 (4.1.2). *Prove*

a. *If* $|A| < |B|$ *and* $|B| \leqslant |C|$*, then* $|A| < |C|$.

b. *If* $|A| \leqslant |B|$ *and* $|B| < |C|$*, then* $|A| < |C|$.

Proof. **(a)** $|A| < |B|$ means $|A| \leqslant |B|$ and $|A| \neq |B|$. We thus have $|A| \leqslant |C|$ by Exercise 90 (d). If $|A| = |C|$, then $|B| \leqslant |A|$ by Exercise 90 (b). But then $|A| = |B|$ by the Cantor-Bernstein Theorem. A contradiction.

(b) $|B| < |C|$ means $|B| \leqslant |C|$ and $|B| \neq |C|$. We thus have $|A| \leqslant |C|$ by Exercise 90 (d). If $|A| = |C|$, then $|C| \leqslant |B|$ by Exercise 90 (a). But then $|B| = |C|$ by the Cantor-Bernstein Theorem. A contradiction. □

► EXERCISE 92 (4.1.3). *If $A \subseteq B$, then $|A| \leqslant |B|$.*

PROOF. Just consider Id_A. This is an embedding on B, and so $|A| \leqslant |B|$. □

► EXERCISE 93 (4.1.4). *Prove:*

a. $|A \times B| = |B \times A|$.

b. $|(A \times B) \times C| = |A \times (B \times C)|$.

c. $|A| \leqslant |A \times B|$ *if* $B \neq \varnothing$.

PROOF. **(a)** Let $f: (a,b) \mapsto (b,a)$ for all $(a,b) \in A \times B$. It is easy to see f is a function. To see f is injective, let $(a_1,b_1) \neq (a_2,b_2)$. Then $f(a_1,b_1) = (b_1,a_1) \neq (b_2,a_2) = f(a_2,b_2)$. To see f is surjective, let $(b,a) \in B \times A$. There must exist $(a,b) \in A \times B$ such that $f(a,b) = (b,a)$. We thus proved that $f: A \times B \to B \times A$ is bijective; consequently, $|A \times B| = |B \times A|$.

(b) Remember that $(A \times B) \times C \neq A \times (B \times C)$ [see Exercise 26 (b)], but as we are ready to prove, these two sets are equipotent. Let

$$f: \big((a,b),c\big) \mapsto \big(a,(b,c)\big), \quad \forall\ \big((a,b),c\big) \in (A \times B) \times C.$$

With the same logic as in (a), we see that f is bijective and so $|(A \times B) \times C| = |A \times (B \times C)|$.

(c) If $B \neq \varnothing$, we can choose some $b \in B$. Let $f: a \mapsto (a,b)$ for all $a \in A$. Then $f: A \to A \times b \subseteq A \times B$ is bijective, and so $|A| \leqslant |A \times B|$ if $B \neq \varnothing$. □

► EXERCISE 94 (4.1.5). *Show that $|S| \leqslant |\mathcal{P}(S)|$.*

PROOF. If $a \in S$, then $\{a\} \subseteq S$; hence, $\{a\} \in \mathcal{P}(S)$ for each $a \in S$. Define

$$\mathcal{A} = \big\{\{a\}: a \in S\big\}.$$

It is clear that $\mathcal{A} \subseteq \mathcal{P}(S)$. Consider the embedding $f: a \mapsto \{a\}$ for all $a \in S$. Then $f: S \to \mathcal{A}$ is bijective, and so $|S| \leqslant |\mathcal{P}(S)|$.

In fact, $|S| < |\mathcal{P}(S)|$. To prove this, we need the following claim.

CLAIM. There is a one-to-one mapping from $A \neq \varnothing$ to B iff there is a mapping from B onto A.

Proof. If $f: A \to B$ is one-to-one, and $\mathcal{R}_f = B^* \subseteq B$, then let

$$g(x) = \begin{cases} f^{-1}(x) & \text{if } x \in B^* \\ a_0 & \text{if } x \in B \smallsetminus B^*, \text{ where } a_0 \in A. \end{cases}$$

Then this g is a mapping from B onto A.

Conversely, let $g\colon B \to A$ be a mapping of B onto A. The relation "$x \sim y$ if $g(x) = g(y)$" is an equivalence relation on B [See Exercise 42, p. 22]. Let h be a choice function on the set of equivalence classes, i.e., if $[x]_\sim$ is an equivalence class, then $h\left([x]_\sim\right)$ is an element of $[x]_\sim$. It is clear that the map $f(x) = \left(h \circ g^{-1}\right)(x)$ is a one-to-one mapping of A into B.

To verify $|S| < |\mathcal{P}(S)|$, we want to show that there is no mapping from S onto $\mathcal{P}(S)$ [note that $\mathcal{P}(S) \neq \varnothing$ since $\varnothing \in \mathcal{P}(S)$ at least; hence here $\mathcal{P}(S)$ takes the role of A in the above claim]. Let $f\colon S \to \mathcal{P}(S)$ be any mapping. We have to show that f is not onto $\mathcal{P}(S)$. Let

$$A := \left\{a \in S : a \notin f(a)\right\} \in \mathcal{P}(S).$$

[Notice that by the Axiom Schema of Comprehension, A is a subset of S, and so is an element of $\mathcal{P}(S)$ by the Axiom of Power Set.] We claim that A does not have a preimage under f. In fact, suppose that is not the case, and $f(a_0) = A$ with some $a_0 \in S$. Then, because $A \subseteq S$, there are two possibilities:

- $a_0 \in A$, i.e., $a_0 \in f(a_0)$ which is not possible for then a_0 cannot be in A by the definition of A.
- $a_0 \notin A$, which is gain not possible, for then $a_0 \notin f(a_0)$, so a_0 should belong to A.

Thus, in either case we have arrived at a contradiction, which means that a_0 with the property $f(a_0) = A$ does not exist. □

► EXERCISE 95 (4.1.6). *Show that* $|A| \leqslant \left|A^S\right|$ *for any* A *and any* $S \neq \varnothing$.

PROOF. For every $a \in A$, we construct a constant function $f_a\colon S \to A$ by letting $f_a(s) = a$ for all $s \in S$. Now $F := \{f_a\colon a \in A\} \subseteq A^S$. Let $g\colon a \mapsto f_a$. It is easy to see that g is surjective. To see g is injective, let $a, a' \in A$ and $a \neq a'$; then $g(a) = f_a \neq f_{a'} = g\left(a'\right)$. This proves that $|A| = |F|$; that is, $|A| \leqslant \left|A^S\right|$, where $S \neq \varnothing$. □

► EXERCISE 96 (4.1.7). *If* $S \subseteq T$, *then* $\left|A^S\right| \leqslant \left|A^T\right|$; *in particular,* $|A^n| \leqslant |A^m|$ *if* $n \leqslant m$.

PROOF. For any $f \in A^S$, we define a corresponding function $g_f \in A^T$ as follows

$$g_f(x) = \begin{cases} f(x) & \text{if } x \in S \\ a_0 & \text{if } x \in T \smallsetminus S, \text{where } a_0 \in A. \end{cases}$$

Then $B := \left\{g_f \in A^T : f \in A^S\right\} \subseteq A^T$. Hence, we have a bijection $A^S \to B$. If $n \leqslant m$, then either $n = m$ or $n \in m$. Therefore, $|A^n| \leqslant |A^m|$ if $n < m$. □

► EXERCISE 97 (4.1.8). $|T| \leqslant \left|S^T\right|$ *if* $|S| \geqslant 2$.

PROOF. Since $|S| \geqslant 2$, we can pick $u, v \in S$ with $u \neq v$. For any $t \in T$, define a function $f_t \in S^T$ as follows

$$f_t(x) = \begin{cases} u & \text{if } x = t \\ v & \text{if } x \neq t. \end{cases}$$

Notice that $A := \left\{ f_t \in S^T : t \in T \right\} \subseteq S^T$. Then we can define a function $g \colon T \to A$ as $g(t) = f_t$. It is clear g is a one-to-one mapping from T onto B; therefore, $|T| \leqslant \left| S^T \right|$. □

► EXERCISE 98 (4.1.9). *If $|A| \leqslant |B|$ and if A is nonempty then there exists a mapping f of B onto A.*

PROOF. $|A| \leqslant |B|$ implies that there is a one-to-one correspondence f from $A \neq \varnothing$ onto $f[A] \subseteq B$. Define $g \colon B \to A$ as follows:

$$g(x) = \begin{cases} f^{-1}(x) & \text{if } x \in f[A] \\ a_0 & \text{if } x \in B \smallsetminus f[A], \end{cases}$$

where $a_0 \in A$. See also the claim in Exercise 94. □

(For Exercise 99–Exercise 101) Let F be a function on $\mathcal{P}(A)$ into $\mathcal{P}(A)$. A set $X \subseteq A$ is called *a fixed point* of F if $F(X) = X$. The function F is called *monotone* if $X \subseteq Y \subseteq A$ implies $F(X) \subseteq F(Y)$.

► EXERCISE 99 (4.1.10). *Let $F \colon \mathcal{P}(A) \to \mathcal{P}(A)$ be monotone. Then F has a fixed point.*

PROOF. Let $\mathcal{T} = \{X \subseteq A \colon F(X) \subseteq X\}$. Note that $\mathcal{T} \neq \varnothing$ since, e.g., $A \in \mathcal{T}$. Now let $\bar{X} = \bigcap \mathcal{T}$ and so $\bar{X} \subseteq X$ for any $X \in \mathcal{T}$. Since F is monotone, we have $F(\bar{X}) \subseteq F(X) \subseteq X$ for every $X \in \mathcal{T}$. Then

$$F\left(\bar{X}\right) \subseteq \bar{X}. \tag{4.1}$$

Hence, $\bar{X} \in \mathcal{T}$.

On the other hand, (4.1) and the monotonicity of F implies that

$$F\left(F\left(\bar{X}\right)\right) \subseteq F\left(\bar{X}\right). \tag{4.2}$$

But (4.2) implies that $F\left(\bar{X}\right) \in \mathcal{T}$, too. Then, by the definition of $\bar{X}$, we have

$$\bar{X} \subseteq F\left(\bar{X}\right). \tag{4.3}$$

Therefore, (4.1) and (4.3) imply that $F\left(\bar{X}\right) = \bar{X}$, i.e., $\bar{X}$ is a fixed point of F. □

► EXERCISE 100 (4.1.11). *Use Exercise 99 to give an alternative proof of the Cantor-Bernstein Theorem.*

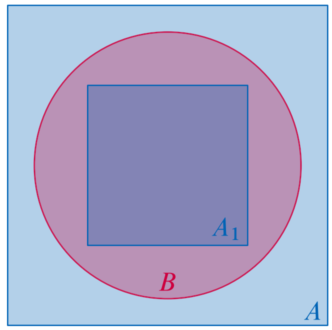


FIGURE 4.1. Cantor-Bernstein Theorem

PROOF. We use Exercise 99 to prove Lemma 4.1.7: *If* $A_1 \subseteq B \subseteq A$ *and* $|A_1| = |A|$, *then* $|B| = |A|$. Let $F: \mathcal{P}(A) \to \mathcal{P}(A)$ be defined by

$$F(X) = (A \smallsetminus B) \cup f[X],$$

where $f: A \to A_1$ is a bijection from A onto A_1. Then F is monotone since f is, and so there exists a fixed point $C \subseteq A$ of F such that

$$C = (A \smallsetminus B) \cup f[C].$$

Let $D = A \smallsetminus C$. Define a function $g: A \to B$ as

$$g(x) = \begin{cases} f(x), & \text{if } x \in C \\ x, & \text{if } x \in D. \end{cases}$$

We now show that g is bijective.

g is surjective We have

$$\begin{aligned} \mathcal{R}_g = f[C] \cup D = f[C] \cup (A \smallsetminus C) &= f[C] \cup \left\{A \smallsetminus \left[(A \smallsetminus B) \cup f[C]\right]\right\} \\ &= f[C] \cup \left[(A \cap B) \cap f^c[C]\right] \\ &= f[C] \cup \left(B \cap f^c[C]\right) \\ &= f[C] \cup B \\ &= B, \end{aligned}$$

where the last equality holds since $f[C] \subseteq A_1 \subseteq B$ [remember that $f: A \leftrightarrow A_1$]. Thus, g is surjective indeed.

g is injective Both $g \upharpoonright C$ and $g \upharpoonright D$ are injective functions, so we need only to show $f[C] \cap D = \varnothing$. This holds because

$$f[C] \cap D = f[C] \cap \left(B \cap f^c[C]\right) = \varnothing.$$

Therefore, $g\colon A \to B$ is bijective, and so $|B| = |A|$. □

▶ EXERCISE 101 (4.1.12). *Prove that $\bar{X}$ in Exercise 99 is the least fixed point of F, i.e., if $F(X) = X$ for some $X \subseteq A$, then $\bar{X} \subseteq X$.*

PROOF. Notice that if $F(X) = X$, then $F(X) \subseteq X$, and so $X \in \mathcal{T}$. Then we obtain the conclusion just because $\bar{X} = \bigcap \mathcal{T}$. □

(For Exercise 102 and Exercise 103) A function $F : \mathcal{P}(A) \to \mathcal{P}(A)$ is *continuous* if

$$F\left(\bigcup_{i \in \mathbb{N}} X_i\right) = \bigcup_{i \in \mathbb{N}} F(X_i)$$

holds for any nondecreasing sequence of subsets of A. [$\langle X_i : i \in \mathbb{N}\rangle$ is *nondecreasing* if $X_i \subseteq X_j$ holds whenever $i \leqslant j$.]

▶ EXERCISE 102 (4.1.13). *Prove that F used in Exercise 100 is continuous.*

PROOF. Let $\langle X_i : i \in \mathbb{N}\rangle \subseteq \mathcal{P}(A)$ be a nondecreasing sequence of A. Then

$$\begin{aligned} F\left(\bigcup_{i \in \mathbb{N}} X_i\right) = (A \smallsetminus B) \cup f\left[\bigcup_{i \in \mathbb{N}} X_i\right] &= (A \smallsetminus B) \cup \left[\bigcup_{i \in \mathbb{N}} f[X_i]\right] \\ &= \bigcup_{i \in \mathbb{N}} \left[(A \smallsetminus B) \cup f[X_i]\right] \\ &= \bigcup_{i \in \mathbb{N}} F(X_i). \end{aligned}$$ □

▶ EXERCISE 103 (4.1.14). *Prove that if $\bar{X}$ is the least fixed point of a monotone continuous function $F\colon \mathcal{P}(A) \to \mathcal{P}(A)$, then $\bar{X} = \bigcup_{i \in \mathbb{N}} X_i$, where we define recursively $X_0 = \varnothing$, $X_{i+1} = F(X_i)$.*

PROOF. We prove this statement with several steps.

(1) We first show that the infinite sequence $\langle X_i : i \in \mathbb{N}\rangle$ defined by $X_0 = \varnothing$, $X_{i+1} = F(X_i)$ is nondecreasing [$\langle X_i : i \in \mathbb{N}\rangle$ exists by the Recursion Theorem]. We use the Induction Principle to prove this property. Let $\mathbf{P}(x)$ denote "$X_x \subseteq X_{x+1}$". Then

- $\mathbf{P}(0)$ holds because $X_0 = \varnothing$.
- Assume that $\mathbf{P}(n)$ holds, i.e., $X_n \subseteq X_{n+1}$. We need to show $\mathbf{P}(n+1)$. Notice that

$$X_{(n+1)+1} = F(X_{n+1}) \overset{\langle 1 \rangle}{\supseteq} F(X_n) = X_{n+1},$$

where ⟨1⟩ holds because $X_n \subseteq X_{n+1}$ by $\mathbf{P}(n)$ and since F is monotone. We thus prove $\mathbf{P}(n+1)$

Therefore, by the Induction Principle, $X_n \subseteq X_{n+1}$, for any $n \in \mathbb{N}$. Then by Exercise 72, $X_i \subseteq X_j$ holds whenever $i \leqslant j$, i.e., $\langle X_i : i \in \mathbb{N}\rangle$ is a nondecreasing infinite sequence.

(2) We now show $\bigcup_{i\in\mathbb{N}} X_i$ is a fixed point of F. Since F is continuous and $\langle X_i : i \in \mathbb{N}\rangle$ is nondecreasing, we have

$$\begin{aligned} F\left(\bigcup_{i\in\mathbb{N}} X_i\right) = \bigcup_{i\in\mathbb{N}} F(X_i) = F(X_0) \cup F(X_1) \cup \cdots &= \varnothing \cup F(X_0) \cup F(X_1) \cdots \\ &= X_0 \cup X_1 \cup X_2 \cup \cdots \\ &= \bigcup_{i\in\mathbb{N}} X_i; \end{aligned}$$

therefore, $\bar{X} := \bigcup_{i\in\mathbb{N}} X_i$ is a fixed point of F.

(3) To see $\bar{X}$ is the least fixed point of F, let X be any fixed point of F, that is, $F(X) = X$. Then, since $\varnothing \subseteq X$, we have $F(\varnothing) \subseteq F(X) = X$ as F is monotone and X is a fixed point of F. Furthermore, $X_1 := F(\varnothing) \subseteq X$ means that $X_2 = F(X_1) \subseteq F(X) = X$. With this process, we have $X_{i+1} = F(X_i) \subseteq X$. Therefore, $\bar{X} = \bigcup_{i\in\mathbb{N}} X_i \subseteq X$ for any fixed point X of F; that is, $\bar{X}$ is the least fixed point of F.

(4) Till now, we have just proved that $\bar{X} = \bigcup_{i\in\mathbb{N}} X_i$ is a least fixed point of F, but the exercise asks us to prove the inverse direction. However, that direction must hold because there is only one least element in the set of all fixed points of F. □

4.2 Finite Sets

► Exercise 104 (4.2.1). *If $S = \{X_0, \ldots, X_{n-1}\}$ and the elements of S are mutually disjoint, then* $\left|\bigcup S\right| = \sum_{i=0}^{n-1} |X_i|$.

Proof. We use the Induction Principle to prove this claim. The statement is true if $|S| = 0$. Assume that it is true for all S with $|S| = n$, and let $S = \{X_0, \ldots, X_{n-1}, X_n\}$ be a set with $n+1$ elements, where each $X_i \in S$ is finite, and the elements of S are mutually disjoint. By the induction hypothesis, $|\bigcup_{i=1}^{n-1} X_i| = \sum_{i=0}^{n-1} |X_i|$, and we have

$$|S| = \left|\left(\bigcup_{i=1}^{n-1} X_i\right) \cup X_n\right| \overset{\langle 1\rangle}{=} \left|\bigcup_{i=1}^{n-1} X_i\right| + |X_i| \overset{\langle 2\rangle}{=} \sum_{i=1}^{n-1} |X_i| + |X_n| = \sum_{i=1}^{n} |X_i|,$$

where ⟨1⟩ is from Theorem 4.2.7, and ⟨2⟩ is from the induction hypothesis. □

► EXERCISE 105 (4.2.2). *If X and Y are finite, then $X \times Y$ is finite, and $|X \times Y| = |X| \times |Y|$.*

PROOF. Let $X = \{x_0, \dots, x_{m-1}\}$, and let $Y = \{y_0, \dots, y_{n-1}\}$, where $\langle x_0, \dots, x_{m-1}\rangle$ and $\langle y_0, \dots, y_{n-1}\rangle$ are injective finite sequences. Then

$$X \times Y = \big\{(x, y) : x \in X \text{ and } y \in Y\big\} = \bigcup_{x' \in X} \big\{(x', y) : y \in Y\big\}.$$

Note that $\{(x', y) : y \in Y\}$ is finite for a fixed $x' \in X$ since Y is finite. Precisely, since $|Y| = m$, there is a bijective function $f : m \to Y$, so we can construct a bijective function $g : m \to \{(x', y) : y \in Y\}$ as $g_i = (x', f_i)$ for all $i \leqslant m - 1$. Therefore, $\big|\{(x', y) : y \in Y\}\big| = m$ for all $x' \in X$. Thus, by Theorem 4.2.7, a finite union of finite sets is finite, we conclude that $X \times Y$ is finite, and

$$|X \times Y| = \left|\bigcup_{x' \in X} \big\{(x', y) : y \in Y\big\}\right| = \sum_{x' \in X} \left|\big\{(x', y) : y \in Y\big\}\right| = \sum_{x' \in X} |Y| = |X| \times |Y|,$$

where the second equality comes from Exercise 104 because $\{(x', y) : y \in Y\} \cap \{(x'', y) : y \in Y\} = \varnothing$ whenever $x', x'' \in X$ and $x' \neq x''$. □

► EXERCISE 106 (4.2.3). *If X is finite, then $|\mathcal{P}(X)| = 2^{|X|}$.*

PROOF. We proceed by induction on the number of elements of X. The statement is true if $|X| = 0$: in this case, $\mathcal{P}(\varnothing) = \{\varnothing\}$, and so $|\mathcal{P}(\varnothing)| = 1 = 2^0$. Assume that it is true for all X with $|X| = n$. Let Y be a set with $n + 1$ elements, i.e., $Y = \{y_0, \dots, y_{n-1}, y_n\}$. Let $X = \{y_0, \dots, y_{n-1}\}$ and $\mathcal{U} = \{U : U \subseteq Y \text{ and } y_n \in U\}$. Then $\mathcal{P}(Y) = \mathcal{P}(X) \cup \mathcal{U}$. Since $\mathcal{P}(X) \cap \mathcal{U} = \varnothing$, and $|\mathcal{P}(X)| = |\mathcal{U}|$, we have by Exercise 104

$$|\mathcal{P}(Y)| = |\mathcal{P}(X)| + |\mathcal{U}| = |\mathcal{P}(X)| + |\mathcal{P}(X)| = 2^n + 2^n = 2^{n+1} = 2^{|Y|}. \qquad \square$$

► EXERCISE 107 (4.2.4). *If X and Y are finite, then X^Y has $|X|^{|Y|}$ elements.*

PROOF. Let $X = \{x_0, \dots, x_{m-1}\}$ and $Y = \{y_0, \dots, y_n\}$, where $\langle x_0, \dots, x_{m-1}\rangle$ and $\langle y_0, \dots, y_n\rangle$ are injective finite sequences. We use the Induction Principle on Y to prove this claim. If $|Y| = 0$, then $X^Y = X^{\varnothing} = \{\langle\,\rangle\} = \{\varnothing\}$, and so $\left|X^Y\right| = 1 = |X|^0 = |X|^{|Y|}$. Assume that for any finite X, $\left|X^Y\right| = |X|^{|Y|}$ if $|Y| = n \in \mathbb{N}$. Now consider a finite set Y with $|Y| = n + 1$. Let $Y = \{y_0, \dots, y_n\}$. Let $Y' = \{y_0, \dots, y_{n-1}\}$; that is, $\left|Y'\right| = n$. By the induction hypothesis, $\left|X^{Y'}\right| = |X|^{\left|Y'\right|} = m^n$, i.e., there are m^n functions in $X^{Y'}$. For any $f \in X^{Y'}$, we can construct a set $F(f)$ as follows:

$$F(f) := \left\{ g_i \in X^Y : g_i(y) = \begin{cases} f(y) & \text{if } y \in Y' \\ x_i & \text{if } y = y_n \end{cases}, \text{ and } i \leqslant m - 1 \right\}.$$

It is easy to see that $X^Y = \bigcup_{f \in X^{Y'}} F(f)$, and $|F(f)| = |X| = m$. Since $\left|X^{Y'}\right| = m^n$ by induction hypothesis, and for each f there is a corresponding set $F(f)$ with m elements; furthermore, $F(f) \cap F(f') = \varnothing$ whenever $f \neq f'$. It then follows from Exercise 104 that

$$\left|X^Y\right| = \sum_{f \in X^{Y'}} |F(f)| = m^n \cdot m = m^{n+1} = |X|^{|Y|}. \qquad \square$$

► EXERCISE 108 (4.2.5). *If $|X| = n \geqslant k = |Y|$, then the number of one-to-one functions $f : Y \to X$ is $n \cdot (n-1) \cdot \cdots \cdot (n-k+1)$.*

PROOF. Let $X = \{x_0, \ldots, x_{n-1}\}$ and $Y = \{y_0, \ldots, y_{k-1}\}$, where $\langle x_0, \ldots, x_{n-1}\rangle$ and $\langle y_0, \ldots, y_{k-1}\rangle$ are injective finite sequences. To construct a injective function $f : Y \to X$, we just pick k different elements from X. Because there are $n \cdot (n-1) \cdot \cdots \cdot (n-k+1)$ different ways to pick n elements from $k \geqslant n$ elements, there are $n \cdot (n-1) \cdot \cdots \cdot (n-k+1)$ injective functions $f : Y \to X$. $\square$

► EXERCISE 109 (4.2.6). *X is finite iff every nonempty system of subsets of X has a $\subseteq$-maximal elements.*

PROOF. To see the $\Longrightarrow$ half, let $X = \{x_0, \ldots, x_{n-1}\}$. If $\varnothing \neq \mathcal{U} \subseteq \mathcal{P}(X)$, let $m := \max\{|Y| : Y \in \mathcal{U}\}$. Such a set m exists since X is finite, so $Y \subseteq X$ is finite [see Theorem 4.2.4], and $\mathcal{P}(X)$ is finite, too [see Theorem 4.2.8]. Let $\widetilde{Y} \in \mathcal{U}$ satisfying $\left|\widetilde{Y}\right| = m$. Now we show $\widetilde{Y}$ is a $\subseteq$-maximal element in $\mathcal{U}$. Suppose not; then there exists $Y' \in \mathcal{P}(X)$ such that $\widetilde{Y} \subset Y'$, but then $\left|\widetilde{Y}\right| < |Y'|$. A contradiction.

For the $\Longleftarrow$ half, assume that X is infinite, and every nonempty system of X has a $\subseteq$-maximal element. Let

$$\mathcal{V} := \{Y \subseteq X : Y \text{ is finite}\}.$$

However, there are no maximal elements in $\mathcal{V}$. To see this, suppose $Y \in \mathcal{V}$ is a $\subseteq$-maximal element, then consider $Y' = Y \cup \{y\}$, where $y \notin Y$ [such a y exists since X is infinite]; then $Y \subset Y'$ and Y' is finite. A contradiction. $\square$

► EXERCISE 110 (4.2.7). *Use Lemma 2.6 and Exercise 105 and Exercise 107 to give easy proofs of commutativity and associativity for addition and multiplication of natural numbers, distributivity of multiplication over addition, and the usual arithmetic properties of exponentiation.*

PROOF. As an example, we only prove the commutativity of addition of natural numbers. Let $|X| = m$ and $|Y| = n$, where $X \cap Y = \varnothing$ and $m, n \in \mathbb{N}$. It follows from Lemma 2.6 that

$$|X \cup Y| = |X| + |Y| = m + n.$$

Similarly, we have $|Y \cup X| = |Y| + |X| = n + m$. Since $|X \cup Y| = |Y \cup X|$, we know that $m + n = n + m$. $\square$

► EXERCISE 111 (4.2.8). *If A, B are finite and $X \subseteq A \times B$, then* $|X| = \sum_{a \in A} k_a$, *where* $k_a = \big|X \cap (\{a\} \times B)\big|$.

PROOF. Let $K_a = X \cap (\{a\} \times B)$ for all $a \in A$. We first show $\bigcup_{a \in A} K_a = X$. Since $K_a \subseteq X$ for all $a \in A$, we have $\bigcup_{a \in A} K_a \subseteq X$. Let $(a, b) \in X$. Then $a \in A$ and $b \in B$, so there exists K_a such that $(a, b) \in \{a\} \times B$; therefore, $(a, b) \in X \cap K_a$. Consequently, $X \subseteq \bigcup_{a \in A} K_a$. We then show that $K_a \cap K_{a'} = \varnothing$ if $a \neq a'$, but this is straightforward because $(a, b) \neq (a', b')$ for any $b, b' \in B$ when $a \neq a'$. Now, follows Exercise 104, we have

$$|X| = \left|\bigcup_{a \in A} K_a\right| = \sum_{a \in A} |K_a| = \sum_{a \in A} k_a. \qquad \square$$

4.3 COUNTABLE SETS

REMARK. We verify that the mapping $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ defined by

$$f(x, y) = \frac{(x + y)(x + y + 1)}{2} + x$$

is bijective (see Figure 4.2).

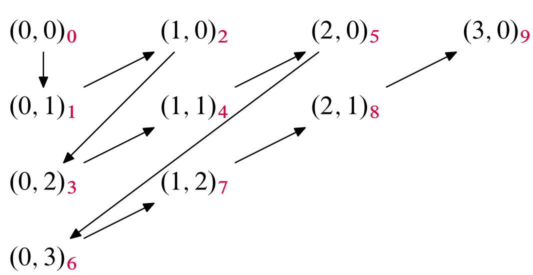


FIGURE 4.2. $(x, y) \mapsto (x + y)(x + y + 1)/2 + x$.

Look at the diagonal where $x + y = 3$ (positions 6, 7, 8, 9 in the diagram). $(x+y)(x+y+1)/2 = 6$ is the sum of the first $x+y = 3$ integers, which accounts for all previous diagonals ($x + y = 0, 1, 2$). Then x locates the position within the diagonal; e.g., $x = 0$ yields position 6, $x = 1$ position 7, $x = 2$ position 8, $x = 3$ position 9.

To go backwards, say we are given the integer 11. Since $1 + 2 + 3 + 4 = 10 < 11 < 1+2+3+4+5$, we are on the diagonal with $x + y = 4$; $x = 0$ gives position 10, $x = 1$ gives 11. Therefore $x = 1$, $y = 4 - 1 = 3$.

► EXERCISE 112 (4.3.1). *Let* $|A_1| = |B_1|, |A_2| = |B_2|$. *Prove*

a. *If* $A_1 \cap A_2 = \varnothing$, $B_1 \cap B_2 = \varnothing$, *then* $|A_1 \cup A_2| = |B_1 \cup B_2|$.

b. $|A_1 \times A_2| = |B_1 \times B_2|$.

c. $|\mathrm{Seq}(A_1)| = |\mathrm{Seq}(B_1)|$.

REMARK. See the original exercise. I am afraid that there are some mistakes in the original one.

PROOF. **(a)** Let $f\colon A_1 \to A_2$, and $g\colon B_1 \to B_2$ be bijections. Define a function $h : (A_1 \cup A_2) \to (B_1 \cup B_2)$ as follows:

$$h(a) = \begin{cases} f(a) & \text{if } a \in A_1 \\ g(a) & \text{if } a \in A_2. \end{cases}$$

It can be see that

$$h = f \cup g\colon (A_1 \cup A_2) \to (B_1 \cup B_2)$$

is bijective since $A_1 \cap A_2 = B_1 \cap B_2 = \varnothing$.

(b) Let f and g be defined as in part (a). We define a function $h : A_1 \times A_2 \to B_1 \times B_2$ as follows:

$$h(a_1, a_2) = \big(f(a_1), g(a_2)\big).$$

Then h is bijective.

(c) We know that $\left|A_1^n\right| = \left|B_1^n\right|$, $\forall\, n \in \mathbb{N}$ [see Lemma 5.1.6]. Notice that $\mathrm{Seq}(A_1) = \bigcup_{n\in\mathbb{N}} A^n$, and $\mathrm{Seq}(B_1) = \bigcup_{n\in\mathbb{N}} B_1^n$, and

$$A_1^m \cap A_1^n = \varnothing, \quad B_1^m \cap B_1^n,$$

for any $m \neq n$, $m, n \in \mathbb{N}$ [because, say, A_1^m and A_1^n have different domains]. Therefore,

$$|\mathrm{Seq}(A_1)| = \left|\bigcup_{n\in\mathbb{N}} A_1^n\right| = \sum_{n\in\mathbb{N}} \left|A_1^n\right| = \sum_{n\in\mathbb{N}} \left|B_1^n\right| = \left|\bigcup_{n\in\mathbb{N}} B_1^n\right| = |\mathrm{Seq}(B_1)|. \qquad \square$$

► EXERCISE 113 (4.3.2). *The union of a finite set and a countable set is countable.*

PROOF. Let $|A| = m$, $|B| = \aleph_0$, and $A' = B \smallsetminus A$. Then $C = A \cup B = A' \cup B$. Since A is finite, $\left|A'\right| = n \leqslant m$. Then there exists two bijections $f : n \to A'$ and $g\colon \mathbb{N} \to B$. Define a function $h : \mathbb{N} \to A' \cup B$ as follows

$$h(i) = \begin{cases} f(i) & \text{if } i < n \\ g(i-n) & \text{if } i \geqslant n. \end{cases}$$

It is easy to see that h is a bijection; thus $|A \cup B| = \left|A' \cup B\right| = \aleph_0$. $\square$

► EXERCISE 114 (4.3.3). *If $A \neq \varnothing$ is finite and B is countable, then $A \times B$ is countable.*

PROOF. Write A as $A = \{a_0, \ldots, a_{n-1}\}$, where $\langle a_0, \ldots, a_{n-1}\rangle$ is a one-to-one finite sequence. Since B is countable, there is a bijection $f : \mathbb{N} \to B$. Pick $a_i \in A$ and consider the set

$$A_i = \left\{(a_i, f(n)) \colon f(n) \in B \text{ and } n \in \mathbb{N}\right\}.$$

Then A_i is countable because there is a bijection $g \colon n \mapsto (a_i, f(n))$.

Since $A \times B = \bigcup_{i \in n} A_i$, that is, $A \times B$ is the union of a finite system of countable sets, and so it is countable by Corollary 4.3.6. □

► EXERCISE 115 (4.3.4). *If $A \neq \varnothing$ is finite, then* Seq(A) *is countable.*

PROOF. It suffices to prove for $A = n \in \mathbb{N}$. We first show that $|\mathrm{Seq}(n)| \geqslant \aleph_0$. Because $n \neq 0$, we can pick an $i \in n$. Consider the following set of finite sequences on n:

$$S = \left\{s_0 = \langle\ \rangle, s_1 = \langle i\rangle, s_2 = \langle i, i\rangle, s_3 = \langle i, i, i\rangle, \ldots\right\}.$$

Define $f : \mathbb{N} \to S$ by letting $f(n) = s_n$; then f is bijective. Because $S \subseteq \mathrm{Seq}(n)$, we have $\aleph_0 = |S| \leqslant |\mathrm{Seq}(n)|$.

We then show that $\mathrm{Seq}(n) \leqslant \aleph_0$. This is simply because $\mathrm{Seq}(n) \subseteq \mathrm{Seq}(\mathbb{N})$ and $\mathrm{Seq}(\mathbb{N}) = \aleph_0$.

Now, by Cantor-Bernstein Theorem, $|\mathrm{Seq}(n)| = \aleph_0$. □

► EXERCISE 116 (4.3.5). *Let A be countable. The set $[A]^n = \left\{S \subseteq A : |S| = n\right\}$ is countable for all $n \in \mathbb{N}$, $n \neq 0$.*

PROOF. It is enough to prove the statement for $A = \mathbb{N}$. We use the Induction Principle in Exercise 69. $[A]^1$ is countable since $[A]^1 = \left\{\{a\} : a \in A\right\}$, and we can define a bijection $f : A \to [A]^1$ by letting $f(a) = \{a\}$ for all $a \in A$. Therefore, $\left|[A]^1\right| = |A| = \aleph_0$. Assume that $[A]^n$ is countable; particularly, we write $[A]^n$ as $[A]^n = \{S_1, S_2, \ldots\}$. We need to prove that $[A]^{n+1}$ is countable, too. For any $S_i \in [A]^n$, we construct a set

$$\mathcal{S}_i = \left\{S_i \cup \{j\} : j \in \mathbb{N} \smallsetminus S_i\right\}.$$

Notice that $J_i = \mathbb{N} \smallsetminus S_i$ is countable; in particular, there exists a bijection $g \colon \mathbb{N} \to J_i$. Define a bijection $h \colon J_i \to \mathcal{S}_i$ by letting $h(j) = S_i \cup \{j\}$, and we see that $|\mathcal{S}_i| = \aleph_0$.

Since $[A]^{n+1} = \bigcup_{i \in \mathbb{N}} \mathcal{S}_i$, the set $[A]^{n+1}$ is a countable union of countable sets. Now for each $i \in \mathbb{N}$, let $a_i = \langle a_i(n) \colon n \in \mathbb{N}\rangle$, where

$$a_i(n) = S_i \cup \{g(n)\}.$$

Then $\mathcal{S}_i = \{a_i(n) \colon n \in \mathbb{N}\}$. It follows from Theorem 4.3.9, $[A]^{n+1}$ is countable. □

► EXERCISE 117 (4.3.6). *A sequence* $\langle s_n \rangle_{n=0}^{\infty}$ *of natural numbers is* eventually constant *if there is* $n_0 \in \mathbb{N}$, $s \in \mathbb{N}$ *such that* $s_n = s$ *for all* $n \geqslant n_0$. *Show that the set of eventually constant sequences of natural numbers is countable.*

PROOF. Let $\mathcal{C}$ be the set of eventually constant sequences of natural numbers. A generic element of $\mathcal{C}$ is $\langle b_0, \ldots, b_{n_0-1}, s, s, \ldots \rangle$, where $\langle b_0, \ldots, b_{n_0-1} \rangle \in \mathbb{N}^{n_0}$, and $s \in \mathbb{N}$.

Let $\mathrm{Seq}(\mathbb{N})$ be the set of all finite sequences of elements of $\mathbb{N}$. Define $f_{n_0} \colon \mathcal{C} \to \mathrm{Seq}(\mathbb{N})$ as follows:

$$f\left(\langle b_0, \ldots, b_{n_0-1}, s, s, \ldots \rangle\right) = \langle b_0, \ldots, b_{n_0-1}, s \rangle .$$

Then f is bijective, and so $|\mathcal{C}| = \aleph_0$. □

► EXERCISE 118 (4.3.7). *A sequence* $\langle s_n \rangle_{n=0}^{\infty}$ *of natural numbers is (eventually)* periodic *if there are* $n_0, p \in \mathbb{N}$, $p \geqslant 1$, *such that for all* $n \geqslant n_0$, $s_{n+p} = s_n$. *Show that the set of all periodic sequences of natural numbers is countable.*

PROOF. Let $\mathcal{P}$ be the set of all eventually periodic sequences of natural numbers. A generic element of $\mathcal{P}$ is

$$\langle b_0, \ldots, b_{n_0-1}, a_{n_0}, a_{n_0+1}, \ldots, a_{n_0+p-1}, a_{n_0}, a_{n_0+1}, \ldots, a_{n_0+p-1}, a_{n_0}, \ldots \rangle .$$

Define $f : \mathcal{P} \to \mathrm{Seq}(\mathbb{N})$ by letting

$$\begin{aligned} f &\left(\langle b_0, \ldots, b_{n_0-1}, a_{n_0}, a_{n_0+1}, \ldots, a_{n_0+p-1}, a_{n_0}, a_{n_0+1}, \ldots, a_{n_0+p-1}, a_{n_0}, \ldots \rangle\right) \\ &= \langle b_0, \ldots, b_{n_0-1}, a_{n_0}, a_{n_0+1}, \ldots, a_{n_0+p-1} \rangle . \end{aligned}$$

f is bijective, and so $|\mathcal{P}| = \aleph_0$. □

► EXERCISE 119 (4.3.8). *A sequence* $\langle s_n \rangle_{n=0}^{\infty}$ *of natural numbers is called an* arithmetic progression *if there is* $d \in \mathbb{N}$ *such that* $s_{n+1} = s_n + d$ *for all* $n \in \mathbb{N}$. *Prove that the set of all arithmetic progressions is countable.*

PROOF. Let $\mathcal{A}$ be the set of all arithmetic progressions. A generic element of $\mathcal{A}$ is

$$\langle a, a+d, a+2d, a+3d, \ldots \rangle .$$

Now define a function $f : \mathcal{A} \to \mathbb{N} \times \mathbb{N}$ by letting

$$f\left(\langle a, a+d, a+2d, \ldots \rangle\right) = \langle a, d \rangle .$$

f is bijection and so $|\mathcal{A}| = \aleph_0$. □

► EXERCISE 120 (4.3.9). *For every* $\boldsymbol{s} = \langle s_0, \ldots, s_{n-1} \rangle \in \mathrm{Seq}(\mathbb{N} \smallsetminus \{0\})$, *let* $f(\boldsymbol{s}) = p_0^{s_0} \cdots p_{n-1}^{s_{n-1}}$, *where* p_i *is the* i*-th prime number. Show that* f *is one-to-one and use this fact to give another proof of* $|\mathrm{Seq}(\mathbb{N})| = \aleph_0$.

PROOF. (i) We use the Induction Principle on n to show that $f(\boldsymbol{s}) \neq f(\boldsymbol{s}')$, whereever $\boldsymbol{s}, \boldsymbol{s}' \in \mathrm{Seq}(\mathbb{N} \smallsetminus \{0\})$ and $\boldsymbol{s} \neq \boldsymbol{s}'$. It is clear that $p_0^{s_0} \neq p_0^{s_0'}$ if $s_0 \neq s_0'$, i.e., this claim holds for $|\boldsymbol{s}| = 1$. Assume which holds for $|\boldsymbol{s}| = n$. We need to show it holds for $|\boldsymbol{s}| = n + 1$.

Suppose $|\boldsymbol{s}| = |\boldsymbol{s}'| = n + 1$ and $\boldsymbol{s} \neq \boldsymbol{s}'$, but $f(\boldsymbol{s}) = f(\boldsymbol{s}')$; that is,

$$p_0^{s_0} \cdot p_{n-1}^{s_{n-1}} \cdot p_n^{s_n} = p_0^{s_0'} \cdots p_{n-1}^{s_{n-1}'} \cdot p_n^{s_n'}. \tag{4.4}$$

There are two cases make (4.4) hold:

- $s_n = s_n'$. In this case, $\langle s_0, \ldots, s_{n-1} \rangle \neq \langle s_0', \ldots, s_{n-1}' \rangle$, and by the inductive hypothesis, $p_0^{s_0} \cdots p_{n-1}^{s_{n-1}} \neq p_0^{s_0'} \cdots p_{n-1}^{s_{n-1}'}$. Therefore, (4.4) implies that

$$p_n^{s_n} \neq p_n^{s_n'}, \tag{4.5}$$

  but which means that $s_n \neq s_n'$. A contradiction.

- $s_n \neq s_n'$. In this case, (4.5) must hold. Under this case, there are two cases further:

  ◇ $\langle s_0, \ldots, s_{n-1} \rangle = \langle s_0', \ldots, s_{n-1}' \rangle$. Then,

$$p_0^{s_0} \cdots p_{n-1}^{s_{n-1}} = p_0^{s_0'} \cdots p_{n-1}^{s_{n-1}'}. \tag{4.6}$$

  However, (4.6) and (4.5) imply that (4.4) fails to hold.

  ◇ $\langle s_0, \ldots, s_{n-1} \rangle \neq \langle s_0', \ldots, s_{n-1}' \rangle$. In this case, we know by the inductive hypothesis that

$$p_0^{s_0} \cdots p_{n-1}^{s_{n-1}} \neq p_0^{s_0'} \cdots p_{n-1}^{s_{n-1}'}, \tag{4.7}$$

  Without loss of generality, we assume that $s_n < s_n'$. Then (4.4) implies that

$$p_0^{s_0} \cdots p_{n-1}^{s_{n-1}} = p_0^{s_0'} \cdots p_{n-1}^{s_{n-1}'} \cdot p_n^{s_n' - s_n}. \tag{4.8}$$

  But we know from the *Unique Factorization Theorem* [see, for example, Apostol 1974] that every natural number $n > 1$ can be represented as a product of prime factors in only one way, apart form the order of the factors. Therefore, (4.8) cannot hold since $p_n \neq p_i$, $\forall\ i \leqslant n - 1$, and $s_n' - s_n > 0$.

(ii) We now show f is indeed onto $\mathbb{N} \smallsetminus \{0, 1\}$. This is follows the Unique Factorization Theorem again; hence, $|\mathrm{Seq}(\mathbb{N} \smallsetminus 0)| = \aleph_0$. To prove $|\mathrm{Seq}(\mathbb{N})| = \aleph_0$, we consider the following function $g\colon \mathrm{Seq}(\mathbb{N}) \to \mathbb{N}$

$$g(\boldsymbol{s}^0) = f\left(\boldsymbol{s}^0 + \mathbf{1}\right), \quad \forall\ \boldsymbol{s}^0 \in \mathrm{Seq}(\mathbb{N}),$$

where $\mathbf{1}$ is the finite sequence $\langle 1, 1, \ldots \rangle$ which has the same length as $\boldsymbol{s}^0$. Then g is one-to-one and onto $\mathbb{N} \smallsetminus \{0, 1\}$, which mean that

$$|\mathrm{Seq}(\mathbb{N})| = \aleph_0.$$

□

▶ EXERCISE 121 (4.3.10). *Let $(S,<)$ be a linearly ordered set and let $\langle A_n : n \in \mathbb{N}\rangle$ be an infinite sequence of finite subsets of S. Then $\bigcup_{n=0}^{\infty} A_n$ is at most countable.*

PROOF. Because $(S,<)$ is a linearly ordered set, and $A_n \subseteq S$ is finite for all $n \in \mathbb{N}$, we can write A_n as

$$A_n = \left\{s_0, s_1, \dots, s_{|A_n|-1}\right\},$$

and rank the elements of A_n as

$$s_0 < s_1 < \dots < s_{|A_n|-1}.$$

Then we can construct $\langle a_n(k) : k < |A_n| - 1\rangle$, a unique enumeration of A_n, by letting $a_n(k) = s_k$. Therefore, $\bigcup_{n=0}^{\infty} A_n$ is at most countable. □

▶ EXERCISE 122 (4.3.11). *Any partition of an at most countable set has a set of representatives.*

PROOF. Let $\mathcal{P}$ be a partition of A. Then there exists an equivalence relation $\sim$ on A induced by $\mathcal{P}$. Since A is at most countable, the set of equivalence classes, $A/\sim = \{[a]_\sim : a \in A\}$, is at most countable. Hence,

$$A/\sim = \langle [a_1]_\sim, [a_2]_\sim, \dots\rangle,$$

and so there is a set of representatives: $\{a_1, a_2, \dots\}$. □

4.4 LINEAR ORDERINGS

▶ EXERCISE 123 (4.4.1). *Assume that $(A_1,<_1)$ is similar to $(B_1,\prec_1)$ and $(A_2,<_2)$ is similar to $(B_2,\prec_2)$.*

a. *The sum of $(A_1,<_1)$ and $(A_2,<_2)$ is similar to the sum of $(B_1,\prec_1)$ and $(B_2,\prec_2)$, assuming that $A_1 \cap A_2 = \varnothing = B_1 \cap B_2$.*

b. *The lexicographic product of $(A_1,<_1)$ and $(A_2,<_2)$ is similar to the lexicographic product of $(B_1,\prec_1)$ and $(B_2,\prec_2)$.*

PROOF. We use $(A,<) \cong (B,\prec)$ to denote that $(A,<)$ is similar to $(B,\prec)$.

(a) Let $(A,<)$ be the sum of $(A_1,<_1)$ and $(A_2,<_2)$, and let $(B,\prec)$ be the sum of $(B_1,\prec_1)$ and $(B_2,\prec_2)$. Then both $(A,<)$ and $(B,\prec)$ are linearly ordered sets (by Lemma 4.4.5 and Exercise 49). Because $(A_1,<_1) \cong (B_1,\prec_1)$, there is an isomorphism $f_1 : (A_1,<_1) \to (B_1,\prec_1)$; similarly, there is an isomorphism $f_2 : A_2 \to B_2$ since $(A_2,<_2) \cong (B_2,\prec_2)$. Define a bijection $g : A \to B$ by $g = f_1 \cup f_2$.

To see $a_1 < a_2$ iff $g(a_1) \prec g(a_2)$, notice that (i) If $a_1, a_2 \in A_1$, then $g(a_1) = f_1(a_1)$ and $g(a_2) = f_2(a_2)$; hence, $a_1 <_1 a_2$ iff $a_1 < a_2$ iff $f_1(a_1) \prec_1 f_1(a_2)$ iff $g(a_1) \prec g(a_2)$. (ii) If $a_1, a_2 \in A_2$ we get the similarly result. (iii) If $a_1 \in A_1$ and $a_2 \in A_2$, then $a_1 < a_2$ by the definition of $<$. Moreover, by the definition of g,

$g(a_1) \in B_1$ and $g(a_2) \in B_2$; then by the definition of $\prec$, we have $g(a_1) \prec g(a_2)$. For the inverse direction, suppose $g(a_1) \prec g(a_2)$. However, since $a_1 \in A_1$ and $a_2 \in A_2$, we know immediately that $a_1 < a_2$ by definition of $<$. We thus proved $(A, <) \cong (B, \prec)$.

(b) Let $A = A_1 \times A_2$ and $B = B_1 \times B_2$. We need to show that $(A, <) \cong (B, \prec)$, where $<$ and $\prec$ are the lexicographic orderings of A and B. First notice that both $(A, <)$ and $(B, \prec)$ are linearly ordered sets by Lemma 4.4.6. For any $(a_1, a_2) \in A$, let $f\colon A \to B$ be defined as

$$f(a_1, a_2) = \left(f_1(a_1), f_2(a_2)\right),$$

where $f_1\colon A_1 \to B_1$ and $f_2\colon A_2 \to B_2$ are isomorphisms. It is easy to see that f is bijective.

Now let $(a_1, a_2), (a_1', a_2') \in A$. Suppose $(a_1, a_2) < (a_1', a_2')$; then either $a_1 <_1 a_1'$, or $a_1 = a_1'$ and $a_2 <_2 a_2'$. In the first case, $f_1(a_1) \prec_1 f_1(a_1')$, and so $(f_1(a_1), f_2(a_2)) \prec (f_1(a_1'), f_2(a_2'))$; in the second case, $f_1(a_1) = f_1(a_1')$ and $f_2(a_2) \prec_2 f_2(a_2')$ and so $(f_1(a_1), f_2(a_2)) \prec (f_1(a_1'), f_2(a_2'))$.

To see the inverse direction, let $(f_1(a_1), f_2(a_2)) \prec (f_1(a_1'), f_2(a_2'))$. Then either $f_1(a_1) \prec_1 f_1(a_1')$ or $f_1(a_1) = f_1(a_1')$ and $f_2(a_2) \prec_2 f_2(a_2')$. In the first case, $f_1(a_1) \prec_1 f_1(a_1')$ and so $a_1 <_1 a_1'$ and so $(a_1, a_2) < (a_1', a_2')$; in the second case, $a_1 = a_1'$ and $a_2 <_2 a_2'$ and so $(a_1, a_2) < (a_1', a_2')$. □

► EXERCISE 124 (4.4.2). *Give an example of linear orderings $(A_1, <_1)$ and $(A_2, <_2)$ such that the sum of $(A_1, <_1)$ and $(A_2, <_2)$ does not have the same order type as the sum of $(A_2, <_2)$ and $(A_1, <_1)$ ("addition of order types is not commutative"). Do the same thing for lexicographic product.*

PROOF. (i) Let $(A_1, <_1) = \left(\mathbb{N} \smallsetminus \{0\}, <^{-1}\right)$, and $(A_2, <_2) = (\mathbb{N}, <)$, where $<$ denotes the usual ordering of numbers by size. Then the sum of $\left(\mathbb{N} \smallsetminus \{0\}, <^{-1}\right)$ and $(\mathbb{N}, <)$ is just $(\mathbb{Z}, <)$. Particularly, there is no greatest element in $(\mathbb{Z}, <)$. However, there is a greatest element in the sum of $(\mathbb{N}, <)$ and $\left(\mathbb{N} \smallsetminus \{0\}, <^{-1}\right)$, namely, -1.

(ii) This is just the case of lexicographic ordering and antilexicographic ordering. □

► EXERCISE 125 (4.4.3). *Prove that the sum and the lexicographic product of two well-orderings are well-orderings.*

PROOF. Let $(A_1, <_1)$ and $(A_2, <_2)$ be two well-ordered sets.

(i) Let $A_1 \cap A_2 = \varnothing$ and $(A, <)$ be the sum of $(A_1, <_1)$ and $(A_2, <_2)$. Let $B \subseteq A$ be nonempty. Write $B = (B \cap A_1) \cup (B \cap A_2)$. Let $B \cap A_1 = B_1$ and $B \cap A_2 = B_2$. Then $B_1 \subseteq A_1$, $B_2 \subseteq A_2$, and $B_1 \cap B_2 = \varnothing$. There are three cases:

- If $B_1 \neq \varnothing$ and $B_2 \neq \varnothing$, then B_1 has a least element b_1, and B_2 has a least element b_2. By definition, $b_1 < b_2$ and so b_1 is the least element of B.

- If $B_1 \neq \varnothing$ and $B_2 = \varnothing$, then B's least element is just b_1.

- If $B_1 = \varnothing$ and $B_2 \neq \varnothing$, then B's least element is just b_2.

(ii) Let $<$ be the lexicographic ordering on $A = A_1 \times A_2$. Take an arbitrary nonempty subset $C \subseteq A$. Let C_1 be the projection of C on A_1. Then $C_1 \neq \varnothing$ and so has a least element $\hat{c}_1$. Now take the set $\{c_2 \in A_2 : (\hat{c}_1, c_2) \in C\}$. This set is nonempty hence has a least element $\hat{c}_2$. We now show that $(\hat{c}_1, \hat{c}_2)$ is the least element of C: for every $(c_1, c_2) \in C$, either $\hat{c}_1 < c_1$, or $\hat{c}_1 = c_1$ and $\hat{c}_2 < c_2$. In both case, $(\hat{c}_1, \hat{c}_2) < (c_1, c_2)$. Thus, $(A, <)$ is well-ordered. □

► Exercise 126 (4.4.4). *If $\langle A_i : i \in \mathbb{N} \rangle$ is an infinite sequence of linearly ordered sets of natural numbers and $|A_i| \geqslant 2$ for all $i \in \mathbb{N}$, then the lexicographic ordering of $\bigtimes_{i \in \mathbb{N}} A_i$ is* not *a well-ordering.*

Proof. Because $|A_i| \geqslant 2$ for all $i \in \mathbb{N}$, we can pick $a_i^1 \in A_i$, $a_i^2 \in A_i$, and $a_i^1 < a_i^2$, where $<$ is the usual linear ordering on $\mathbb{N}$. Consider the infinite sequence $\langle \boldsymbol{a}_0, \boldsymbol{a}_1, \ldots \rangle$, where

$$\begin{aligned} \boldsymbol{a}_0 &= \left\langle a_0^2, a_1^2, a_2^2, a_3^2, a_4^2, \ldots \right\rangle, \\ \boldsymbol{a}_1 &= \left\langle a_0^1, a_1^2, a_2^2, a_3^2, a_4^2, \ldots \right\rangle, \\ \boldsymbol{a}_2 &= \left\langle a_0^1, a_1^1, a_2^2, a_3^2, a_4^2, \ldots \right\rangle, \\ &\cdots \end{aligned}$$

In this sequence, $\boldsymbol{a}_{n+1} \prec \boldsymbol{a}_n$ by the lexicographic ordering $\prec$. More explicitly, $\operatorname{diff}(\boldsymbol{a}_{n+1}, \boldsymbol{a}_n) = n$, and $\boldsymbol{a}_{n+1}(n) = a_n^1 < a_n^2 = \boldsymbol{a}_n(n)$. Then the set $\{\boldsymbol{a}_0, \boldsymbol{a}_1, \ldots\}$ does not have a least element, that is, the lexicographic ordering of $\bigtimes_{i \in \mathbb{N}} A_i$ is not well-ordering. □

► Exercise 127 (4.4.5). *Let $\langle (A_i, <_i) : i \in I \rangle$ be an indexed system of mutually disjoint linearly ordered sets, $I \subseteq \mathbb{N}$. The relation $\prec$ on $\bigcup_{i \in I} A_i$ defined by: $a \prec b$ iff either $a, b \in A_i$ and $a <_i b$ for some $i \in I$ or $a \in A_i$, $b \in A_j$ and $i < j$ (in the usual ordering of natural numbers) is a linear ordering. If all $<_i$ are well-orderings, so is $\prec$.*

Proof. We first show that $\prec$ is a linear ordering (compare with Exercise 49). (Transitivity) Let $a, b, c \in \bigcup_{i \in I} A_i$ with $a \prec b$ and $b \prec c$. If $a, b, c \in A_i$ for some $i \in I$, then $a <_i b$ and $b <_i c$ imply that $a \prec c$; if $a, b \in A_i$, $c \in A_j$, and $i < j$, then $a \prec c$; if $a \in A_i$, $b, c \in A_j$, and $i < j$, then $a \prec c$. (Asymmetry) Let $a, b \in \bigcup_{i \in I} A_i$ and $a \prec b$. If $a, b \in A_i$, then $a <_i b$, which implies that $a \not>_i b$, which implies that $a \not\succ b$; if $a \in A_i$, $b \in A_j$, and $i < j$, then, by definition, $a \not\succ b$. (Linearity) Given $a, b \in \bigcup_{i \in I} A_i$, one of the following cases has to occur: If $a, b \in A_i$ for some $i \in I$, then a, b is comparable since $<_i$ is; if $a \in A_i$, $b \in A_j$, and $i < j$, then $a \prec b$; if $a \in A_i$, $b \in A_j$, and $i > j$, then $b \prec a$.

Now suppose that all $<_i$ are well-orderings. Pick an arbitrary nonempty subset $A \subseteq \bigcup_{i \in I} A_i$. For each $a \in A$, there exists a unique $i_a \in I$ such that $a \in A_{i_a}$. Let

$$I_A = \{i \in I : a \in A_i \text{ for some } a \in A\}.$$

Notice that $I_A \neq \varnothing$. Then I_A has a least element i'. Since $A_{i'}$ is also nonempty, $A_{i'}$ has a least element $a_{i'}$. Hence, $a_{i'}$ is the least element of A. □

▶ EXERCISE 128 (4.4.6). *Let $(\mathbb{Z}, <)$ be the set of all integers with the usual linear ordering. Let $\prec$ be the lexicographic ordering of $\mathbb{Z}^{\mathbb{N}}$ as defined in Theorem 4.4.7. Finally, let $FS \subseteq \mathbb{Z}^{\mathbb{N}}$ be the set of all eventually constant elements of $\mathbb{Z}^{\mathbb{N}}$; i.e., $\langle a_i : i \in \mathbb{N}\rangle \in FS$ iff there exists $n_0 \in \mathbb{N}$, $a \in \mathbb{Z}$ such that $a_i = a$ for all $i \geqslant n_0$ (compare with Exercise 117). Prove that FS is countable and $(FS, \prec \cap FS^2)$ is a dense linear ordered set without endpoints.*

PROOF. The countability of FS is obtained by a similar proof as in Exercise 117. It is also easy to see that $(FS, \prec \cap FS^2)$ is a linear ordered set without endpoints. So we just show that it is dense.

Take two arbitrary elements $\boldsymbol{a} = \langle a_i : i \in \mathbb{N}\rangle$ and $\boldsymbol{b} = \langle b_i : i \in \mathbb{N}\rangle$ in FS, and assume that $\boldsymbol{a} \prec \boldsymbol{b}$. Then there exists $n_0 \in \mathbb{N}$ such that $a_{n_0} < b_{n_0}$, where n_0 is the least element of diff $(\boldsymbol{a}, \boldsymbol{b})$. Define $\boldsymbol{c} = \langle c_i : i \in \mathbb{N}\rangle$ by letting

$$c_i = \begin{cases} a_i & \text{if } i \leqslant n_0 \\ \max\{a_i, b_i\} & \text{if } i > n_0. \end{cases}$$

This infinite sequence $\boldsymbol{c}$ is well-defined since both $\boldsymbol{a}$ and $\boldsymbol{b}$ are eventually constant. Then $\boldsymbol{a} \prec \boldsymbol{c} \prec \boldsymbol{b}$. □

▶ EXERCISE 129 (4.4.7). *Let $\prec$ be the lexicographic ordering of $\mathbb{N}^{\mathbb{N}}$ (where $\mathbb{N}$ is assumed to be ordered in the usual way) and let $P \subseteq \mathbb{N}^{\mathbb{N}}$ be the set of all eventually periodic, but not eventually constant, sequences of natural numbers (see Exercises 117 and 118 for definitions of these concepts). Show that $(P, \prec \cap P^2)$ is a countable dense linearly ordered set without endpoints.*

PROOF. It is evident that $(P, \prec \cap P^2)$ is a countable linearly ordered set, so we focus on density. Take two arbitrary elements $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{N}^{\mathbb{N}}$ with $\boldsymbol{a} \prec \boldsymbol{b}$. Then there exists $n_0 \in \mathbb{N}$ such that $a_{n_0} < b_{n_0}$, where n_0 is defined as in the previous exercise. Define $\boldsymbol{c} \in P$ as in the previous exercise, we have $\boldsymbol{a} \prec \boldsymbol{c} \prec \boldsymbol{b}$. □

▶ EXERCISE 130 (4.4.8). *Let $(A, <)$ be linearly ordered. Define $\prec$ on* Seq(A) *by: $\langle a_0, \dots, a_{m-1}\rangle \prec \langle b_0, \dots, b_{n-1}\rangle$ iff there is $k < n$ such that $a_i = b_i$ for all $i < k$ and either $a_k < b_k$ or a_k is undefined (i.e., $k = m < n$). Prove that $\prec$ is a linear ordering. If $(A, <)$ is well-ordered,* $(\text{Seq}(A), \prec)$ *is also well-ordered.*

PROOF. *Transitivity*: Let $\langle a_0, \dots, a_{m-1}\rangle \prec \langle b_0, \dots, b_{n-1}\rangle \prec \langle c_0, \dots, c_{\ell-1}\rangle$. Then there exists $k_1 < n$ such that $a_i = b_i$ for all $i < k_1$ and either $a_{k_1} < b_{k_1}$ or a_{k_1} is undefined. Similarly, there exists $k_2 < \ell$ such that $b_i < c_i$ for all $i < k_2$ and either $b_{k_2} < c_{k_2}$ or b_{k_2} is undefined. Assume that $k_1 < k_2$.

- If $a_i = b_i$ for all $i < k_1$, $a_{k_1} < b_{k_1}$, $b_i = c_i$ for all $i < k_2$, and $b_{k_2} < c_{k_2}$, then $a_i = c_i$ for all $i < k_1$, and $a_{k_1} < c_{k_1}$, i.e., $\langle a_0, \dots, a_{m-1}\rangle \prec \langle c_0, \dots, c_{\ell-1}\rangle$.

- If $a_i = b_i$, $k_1 = m < n$, $b_i = c_i$ for all $i < k_2$, and $b_{k_2} < c_{k_2}$, then $a_i = b_i = c_i$ for all $i < k_1$, and a_{k_1} is undefined, i.e., $\langle a_0, \ldots, a_{m-1} \rangle \prec \langle c_0, \ldots, c_{\ell-1} \rangle$.
- If $a_i = b_i$ for all $i < k_1$, $a_{k_1} < b_{k_1}$, $b_i = c_i$ for all $i < k_2$, and $k_2 = n < \ell$, then $a_i = b_i = c_i$ for all $i < k_1$, and $a_{k_1} < b_{k_1} = c_{k_1}$, i.e., $\langle a_0, \ldots, a_{m-1} \rangle \prec \langle c_0, \ldots, c_{\ell-1} \rangle$.

We can see that $\langle a_0, \ldots, a_{m-1} \rangle \prec \langle c_0, \ldots, c_{\ell-1} \rangle$ also holds for $k_1 \geqslant k_2$.

Asymmetry: Follows from definition immediately.

Linearity: Given $\langle a_0, \ldots, a_{m-1} \rangle, \langle b_0, \ldots, b_{n-1} \rangle \in \mathrm{Seq}(A)$. If $m < n$, then either there exists $k < m$ such that $a_i = b_i$ for all $i < k$ and $a_k < b_k$ or $a_k > b_k$, which implies that $\langle a_0, \ldots, a_{m-1} \rangle \prec \langle b_0, \ldots, b_{n-1} \rangle$ or $\langle a_0, \ldots, a_{m-1} \rangle \succ \langle b_0, \ldots, b_{n-1} \rangle$; or $a_i = b_i$ for all $i < m$, which implies that $\langle a_0, \ldots, a_{m-1} \rangle \prec \langle b_0, \ldots, b_{n-1} \rangle$. All other cases can be analyzed similarly.

Well-ordering: Let $X \subseteq \mathrm{Seq}(A)$ be nonempty, and $(A, <)$ be well-ordered. Let

$$B_i = \{a_i \in A \colon \langle a_0, \ldots, a_i, \ldots, a_{n-1} \rangle \in X\}.$$

Then $B_i \subseteq A$ is nonempty and so has a least element b_i. The sequence $\langle b_0, \ldots, b_{\ell-1} \rangle$ is the least element of X and so $(\mathrm{Seq}(A), \prec)$ is well-ordered. □

► Exercise 131 (4.4.10). *Let $(A, <)$ be a linearly ordered set without endpoints, $A \neq \varnothing$. A* closed interval *$[a, b]$ is defined for $a, b \in A$ by $[a, b] = \{x \in A \colon a \leqslant x \leqslant b\}$. Assume that each closed interval $[a, b]$, $a, b \in A$, has a finite number of elements. Then $(A, <)$ is similar to the set $\mathbb{Z}$ of all integers in the usual ordering.*

Proof. Take arbitrary $a, b \in A$ with $a \leqslant b$. Denote $[a, b]$ as $\{a_{i_0}, a_{i_1}, \ldots, a_{i_k}\}$ (since it is finite), where $a_{i_0} = a$ and $a_{i_k} = k$, with $a_{i_0} < \cdots < a_{i_k}$. Let

$$h_{[a,b]} = \left\{(a_{i_0}, 0), (a_{i_1}, 1), \ldots, (a_{i_k}, k)\right\}.$$

Clearly, h is a partial isomorphism. Now for any $c \in A$, either $c < a$ or $c > b$. For example, assume that $c < a$. Let $[c, a] = \{c_{j_\ell}, \ldots, c_{j_0}\}$, where $c_{j_\ell} = c$ and $c_{j_0} = a$, with $c_{j_\ell} < \cdots < c_{j_0}$. Let

$$h_{[c,a]} = \left\{(c_{j_\ell}, -\ell), \ldots, (c_{j_1}, -1), (c_{j_0}, 0)\right\}.$$

Let $h = \bigcup_{a,b \in A} h_{[a,b]}$. Then h is an isomorphism and so $(A, <) \cong (\mathbb{Q}, <)$. □

► Exercise 132 (4.4.11). *Let $(A, <)$ be a dense linearly ordered set. Show that for all $a, b \in A$, $a < b$, the closed interval $[a, b]$, as defined in Exercise 131, has infinitely many elements.*

Proof. If $[a, b]$ has finitely element, then $(A, <) \cong (\mathbb{Z}, <)$. However, $(\mathbb{Z}, <)$ is not dense. □

► Exercise 133 (4.4.12). *Show that all countable dense linearly ordered sets with both endpoints are similar.*

PROOF. Let $(P, \prec)$ and $(Q, <)$ be such two sets. Let $\langle p_n : n \in \mathbb{N} \rangle$ be an injective sequence such that $P = \{p_n : n \in \mathbb{N}\}$, and let $\langle q_n : n \in \mathbb{N} \rangle$ be an injective sequence such that $Q = \{q_n : n \in \mathbb{N}\}$. We also assume that $p_0 \prec p_1 \prec \cdots \prec \overline{p}$ and $q_0 < q_1 < \cdots < \overline{q}$, where $\overline{p}$ is the greatest element of P and $\overline{q}$ is the greatest element of Q.

Let $h_0 : p_0 \mapsto q_0$. Having defined $h_n : \{p_0, \dots, p_n\} \to \{q_0, \dots, q_n\}$, we let $h_{n+1} : h_n \cup \{(p_{n+1}, q_{n+1})\}$. Now let $h = \bigcup_{=0}^{\infty} h_i$. Then h is an isomorphism and so $(P, \prec) \cong (Q, <)$. □

► EXERCISE 134 (4.4.13). *Let* $(\mathbb{Q}, <)$ *be the set of all rational numbers in the usual ordering. Find subsets of* $\mathbb{Q}$ *similar to*

a. *the sum of two copies of* $(\mathbb{N}, <)$*;*

b. *the sum of* $(\mathbb{N}, <)$ *and* $(\mathbb{N}, <^{-1})$*;*

c. *the lexicographic product of* $(\mathbb{N}, <)$ *and* $(\mathbb{N}, <)$*.*

PROOF. For (a) and (b), we take the subset as $\mathbb{Z}$. For (c), let $A = \{m - 1/(n + 1) : m, n \in \mathbb{N}, \text{irreducible}\}$. We show that $A \cong \mathbb{N} \times \mathbb{N}$. Let $h : A \to \mathbb{N} \times \mathbb{N}$ be defined as $h(m - 1/(n+1)) = (m, n)$. It is clear that h is bijective. First assume that $m_1 - 1/(n_1 + 1) < m_2 - 1/(n_2 + 1)$. Then it is impossible that $m_1 > m_2$; for otherwise,

$$\frac{1}{n_1 + 1} - \frac{1}{n_2 + 1} > m_2 - m_1 \geqslant 1,$$

which is impossible. If $m_1 < m_2$, there is nothing to prove. So assume that $m_1 = m_2$, but then $n_1 < n_2$ and hence $(m_1, n_1) < (m_2, n_2)$. The other hand can be proved similarly. □

4.5 Complete Linear Orderings

REMARK (p. 87). Let $(P, <)$ be a dense linearly ordered set. $(P, <)$ is complete iff it does not have any gaps.

PROOF. We first show that if $(P, <)$ does not have any gaps, then it is complete. Suppose $(P, <)$ is not complete, that is, there is a nonempty set $S \subseteq P$ bounded from above, and S does not have a supremum. Let

$$A = \left\{x \in P : x \leqslant s \text{ for some } s \in S\right\},$$
$$B = \left\{x \in P : x > s \text{ for every } s \in S\right\}.$$

Then (A, B) is a gap: $A \neq \varnothing$ since $S \subseteq A$, and $B \neq \varnothing$ since S is bounded from above. Next, for every $p \in P$, if $p > s$ for all $s \in S$ then $p \in B$; if $p \leqslant s$ for some $s \in S$ then $p \in A$, i.e., $A \cup B = P$. Finally, $A \cap B = \varnothing$, and if $a \in A$ and $b \in B$ then there exists $s \in S$ such that $a \leqslant s < b$, i.e., $a < b$.

If A has a greatest element, or B has a least element, then A has a supremum, but which means that S has a supremum, too. To see this, let $\sup A = \gamma$. Then $\gamma \geqslant a$ for all $a \in A$, and if $\gamma' < \gamma$, there exists $\widetilde{a} \in A$ such that $\gamma' < \widetilde{a} \leqslant \gamma$. Since $S \subseteq A$, we get $s \leqslant \gamma$ for all $s \in S$. So we need only to prove that there exists $\widetilde{s} \in S$ such that $\gamma' < s \leqslant \gamma$. By definition, there exists $\widetilde{s} \in S$ such that $\widetilde{s} \geqslant \widetilde{a}$; therefore, $\gamma' < \widetilde{a} \leqslant \widetilde{s} \leqslant \gamma$ implies that $\gamma' < \widetilde{s} \leqslant \gamma$ since $<$ is transitive.

For the other direction, assume that $(P, <)$ has a gap (A, B). Then $\varnothing \neq A \subseteq P$, A is bounded from above (since any element of B is an upper bound of A). But A does not have a supremum; hence $(P, <)$ is not complete. □

► EXERCISE 135 (4.5.1). *Prove that there is no $x \in \mathbb{Q}$ for which $x^2 = 2$.*

PROOF. (See Rudin, 1976, for this exercise and Exercise 136.) If there were such a $x \in \mathbb{Q}$, we could write $x = m/n$, where m and n are integers that are not both even. Let us assume this is done. Then $x^2 = 2$ implies

$$m^2 = 2n^2. \tag{4.9}$$

This shows that m^2 is even. Hence m is even (if m were odd, then $m = 2k + 1$, $k \in \mathbb{Z}$, then $m^2 = 2\left(2k^2 + 2k\right) + 1$ is odd), and so m^2 is divisible by 4. It follows that the right side of (4.9) is divisible by 4, so that n^2 is even, which implies that n is even.

The assumption that $x^2 = 2$ holds thus leads to the conclusion that both m and n are even, contrary to our choice of m and n. Thus, $x^2 \neq 2$ for all $x \in \mathbb{Q}$. □

► EXERCISE 136 (4.5.2). *Show that (A, B), where*

$$A = \left\{x \in \mathbb{Q} : x \leqslant 0 \textit{ or } (x > 0 \textit{ and } x^2 < 2)\right\}, B = \left\{x \in \mathbb{Q} : x > 0 \textit{ and } x^2 > 2\right\},$$

is a gap in $(\mathbb{Q}, <)$.

PROOF. To show that (A, B) is a gap in $(\mathbb{Q}, <)$, we need to show (a)–(c) of the definition hold. Since (a) and (b) are clear [note that $\sqrt{2} \notin \mathbb{Q}$ by Exercise 135], we need only to verify (c); that is, A does not have a greatest element, and B does not have a least element.

More explicitly, for every $p \in A$ we can find a rational $q \in A$ such that $p < q$, and for every $p \in B$ such that $q < p$. To to this, we associate with each rational $p > 0$ the number

$$q = p - \frac{p^2 - 2}{p + 2} = \frac{2p + 2}{p + 2}. \tag{4.10}$$

Then

$$q^2 - 2 = \frac{2(p^2 - 2)}{(p + 2)^2}. \tag{4.11}$$

- If $p \in A$ then $p^2 - 2 < 0$, (4.10) shows that $q > p$, and (4.11) shows that $q^2 < 2$. Thus $q \in A$.

- If $p \in B$ then $p^2 - 2 > 0$, (4.10) shows that $0 < q < p$, and (4.11) shows that $q^2 > 2$. Thus $q \in B$. □

► EXERCISE 137 (4.5.3). *Let* $0.a_1a_2a_3\cdots$ *be an infinite, but not periodic, decimal expansion. Let*

$$A = \left\{x \in \mathbb{Q} : x \leqslant 0.a_1a_2\cdots a_k \text{ for some } k \in \mathbb{N} \smallsetminus \{0\}\right\},$$
$$B = \left\{x \in \mathbb{Q} : x \geqslant 0.a_1a_2\cdots a_k \text{ for all } k \in \mathbb{N} \smallsetminus \{0\}\right\}.$$

Show that (A, B) *is a gap in* $(\mathbb{Q}, <)$.

PROOF. It is easy to see that A and B are nonempty, disjoint, and $A \cup B = \mathbb{Q}$. Further, if $a \in A$ and $b \in B$, then there exists $k \in \mathbb{N} \smallsetminus \{0\}$ such that $a \leqslant 0.a_1a_2\cdots a_k < b$.

If A has a greatest element α, then $\alpha = 0.a_1a_2\cdots a_k$ for some $k \in \mathbb{N} \smallsetminus \{0\}$. But $\alpha < 0.a_1a_2\cdots a_k1 \in A$. Similarly, B does not have a least element. □

► EXERCISE 138 (4.5.4). *Show that a dense linearly ordered set* $(P, <)$ *is complete iff every nonempty* $S \subseteq P$ *bounded from below has an infimum.*

PROOF. We first suppose $(P, <)$ is complete. Then by definition, every nonempty $S' \subseteq P$ bounded from above has a supremum. Now suppose $\varnothing \neq S \subseteq S$ is bounded from below. Let S' be the set of all lower bounds of S. Since S is bounded from below, $S' \neq \varnothing$, and since S' consists of exactly those $s' \in P$ which satisfy the inequality $s' \leqslant s$ for every $s \in S$, we see that every $s \in S$ is an upper bound of S'. Thus S' is bounded above and

$$\alpha = \sup S'$$

exists in P by definition of completion. We show that indeed $\alpha = \inf S$.

- If $\gamma < \alpha$ then γ is not an upper bound of S', hence $\gamma \notin S$. It follows that $\alpha \leqslant s$ for every $s \in S$ since s is an upper bound of S'. Thus α is an lower bound of S, i.e., $\alpha \in S'$.

- If $\alpha < \beta$ then $\beta \notin S'$, since α is an upper bound of S'.

We have shown that $\alpha \in S'$ but $\beta \notin S'$ if $\beta > \alpha$. In other words, α is a lower bound of S, but β is not if $\beta > \alpha$. This means that $\alpha = \inf S$.

With the same logic, we can prove the inverse direction. Suppose every nonempty $S \subseteq P$ bounded from below has an infimum. Let $\varnothing \neq S' \subseteq P$ is an arbitrary set bounded from above. We want to show that S' has a supremum. Let S be the set of all upper bounds of S'. Since S' is bounded above, $S \neq \varnothing$, and since S consists of exactly those $s \in P$ which satisfy the inequality $s \geqslant s'$ for every $s' \in S'$, we see that every $s' \in S'$ is an lower bound of S. Therefore, S is bounded from below and

$$\beta = \inf S$$

exists in P. We show that $\beta = \sup S'$, too.

- As before, we first show $\beta \in S$. If $\gamma > \beta$, then γ is not an lower bound of S, hence $\gamma \notin S'$. It follows that $s' \leqslant \beta$ for every $s' \in S'$; that is, β is an upper bound of S', so $\beta \in S$.
- If $\alpha < \beta$ then $\alpha \notin S$, since β is an upper bound of S'.

We have shown that $\beta \in S$ but $\alpha \notin S$ if $\alpha < \beta$. Therefore, $\beta = \sup S'$. □

► EXERCISE 139 (4.5.5). *Let D be dense in $(P, <)$, and let E be dense in $(D, <)$. Show that E is dense in $(P, <)$.*

PROOF. It seems that the definition of *denseness* in the Theorem 4.5.3(c) is wrong. We use the definition from Jech (2006):

DEFINITION 4.1. a. A linear ordering $(P, <)$ is *dense* if for all $a < b$ there exists a c such that $a < c < b$.

b. A set $D \subseteq P$ is a *dense subset* if for all $a < b$ in P there exists a $d \in D$ such that $a < d < b$.

Let $p_1, p_2 \in P$ and $p_1 < p_2$. Since D is dense in $(P, <)$, there exists $d_1 \in D$ such that

$$p_1 < d_1 < p_2. \tag{4.12}$$

Because $d_1 \in D \subseteq P$, we know there exists a $d_2 \in D$ such that

$$d_1 < d_2 < p_2. \tag{4.13}$$

Because E is dense in $(D, <)$, there exists $e \in E$ such that

$$d_1 < e < d_2. \tag{4.14}$$

Now combine (4.12)—(4.14) and adopt the fact that $<$ is linear, we conclude that for any $p_1 < p_2$, there exists $e \in E$ such that $p_1 < e < p_2$; that is, E is dense in $(P, <)$. □

► EXERCISE 140 (4.5.8). *Prove that the set $\mathbb{R} \smallsetminus \mathbb{Q}$ of all irrational numbers is dense in $\mathbb{R}$.*

PROOF. We want to show that for any $a, b \in \mathbb{R}$ and $a < b$, there is an $x \in \mathbb{R} \smallsetminus \mathbb{Q}$ such that $a < x < b$. We can chose such an x as follows:

$$x = \begin{cases} (a+b)/2 & \text{if } x \in \mathbb{R} \smallsetminus \mathbb{Q} \\ (a+b)/\sqrt{2} & \text{otherwise.} \end{cases}$$ □

4.6 Uncountable Sets

► Exercise 141 (4.6.1). *Use the diagonal argument to show that* $\mathbb{N}^{\mathbb{N}}$ *is uncountable.*

Proof. Consider any infinite sequence $\left\langle \boldsymbol{a}_n \in \mathbb{N}^{\mathbb{N}} : n \in \mathbb{N} \right\rangle$, we prove that there is some $\boldsymbol{d} \in \mathbb{N}^{\mathbb{N}}$, and $\boldsymbol{d} \neq \boldsymbol{a}_n$ for all $n \in \mathbb{N}$. This can be done by defining

$$\boldsymbol{d}(n) = \boldsymbol{a}_n(n) + 1.$$

Note that $\boldsymbol{a}_n(n) + 1 \in \mathbb{N}$, and $\boldsymbol{d} \neq \boldsymbol{a}_n$ for all $n \in \mathbb{N}$. □

► Exercise 142 (4.6.2). *Show that* $\left|\mathbb{N}^{\mathbb{N}}\right| = 2^{\aleph_0}$.

Proof. We first show that $\mathbb{N}^{\mathbb{N}} \subseteq \mathcal{P}(\mathbb{N} \times \mathbb{N})$. A generic element of $\mathbb{N}^{\mathbb{N}}$ can be written as $\{(1, a_1), (2, a_2), (3, a_3), \dots\}$. Since $(n, a_n) \in \mathbb{N} \times \mathbb{N}$ for all $n \in \mathbb{N}$, we have $\{(1, a_1), (2, a_2), \dots\} \subseteq \mathbb{N} \times \mathbb{N}$; that is, $\{(1, a_1), (2, a_2), \dots\} \in \mathcal{P}(\mathbb{N} \times \mathbb{N})$. Therefore,

$$2^{\mathbb{N}} \subseteq \mathbb{N}^{\mathbb{N}} \subseteq \mathcal{P}(\mathbb{N} \times \mathbb{N}).$$

Because $|\mathbb{N} \times \mathbb{N}| = |\mathbb{N}|$, we have $|\mathcal{P}(\mathbb{N} \times \mathbb{N})| = |\mathcal{P}(\mathbb{N})|$ (by Exercise 143); furthermore, $\left|2^{\mathbb{N}}\right| = |\mathcal{P}(\mathbb{N})|$, so $\left|2^{\mathbb{N}}\right| = |\mathcal{P}(\mathbb{N} \times \mathbb{N})|$. It follows from Cantor-Bernstein Theorem that $\left|\mathbb{N}^{\mathbb{N}}\right| = 2^{\aleph_0} = \mathfrak{c}$. □

► Exercise 143 (4.6.3). *Show that* $|A| = |B|$ *implies* $|\mathcal{P}(A)| = |\mathcal{P}(B)|$.

Proof. Let $f : A \to B$ be a bijection. For every subset $a \subseteq A$, we define a function $g : \mathcal{P}(A) \to \mathcal{P}(B)$ as follows:

$$g(a) = f[a],$$

where $f[a]$ is the image of a under f. Then it is easy to see that g is bijective. Hence, $|\mathcal{P}(A)| = |\mathcal{P}(B)|$. □

5

CARDINAL NUMBERS

5.1 Cardinal Arithmetic

► Exercise 144 (5.1.1). *Prove properties (a)–(n) of cardinal arithmetic stated in the text of this section.*

a. $\kappa + \lambda = \lambda + \kappa$.

b. $\kappa + (\lambda + \mu) = (\kappa + \lambda) + \mu$.

c. $\kappa \leqslant \kappa + \lambda$.

d. *If* $\kappa_1 \leqslant \kappa_2$ *and* $\lambda_1 \leqslant \lambda_2$, *then* $\kappa_1 + \lambda_1 \leqslant \kappa_2 + \lambda_2$.

e. $\kappa \cdot \lambda = \lambda \cdot \kappa$.

f. $\kappa \cdot \big(\lambda \cdot \mu\big) = (\kappa \cdot \lambda) \cdot \mu$.

g. $\kappa \cdot (\lambda + \mu) = \kappa \cdot \lambda + \kappa \cdot \mu$.

h. $\kappa \leqslant \kappa \cdot \lambda$ *if* $\lambda > 0$.

i. *If* $\kappa_1 \leqslant \kappa_2$ *and* $\lambda_1 \leqslant \lambda_2$, *then* $\kappa_1 \cdot \lambda_1 \leqslant \kappa_2 \cdot \lambda_2$.

j. $\kappa + \kappa = 2 \cdot \kappa$.

k. $\kappa + \kappa \leqslant \kappa \cdot \kappa$, *whenever* $\kappa \geqslant 2$.

l. $\kappa \leqslant \kappa^\lambda$ *if* $\lambda > 0$.

m. $\lambda \leqslant \kappa^\lambda$ *if* $\kappa > 1$.

n. *If* $\kappa_1 \leqslant \kappa_2$ *and* $\lambda_1 \leqslant \lambda_2$, *then* $\kappa_1^{\lambda_1} \leqslant \kappa_2^{\lambda_2}$.

Proof. We let $|A| = \kappa, |B| = \lambda$, and $|C| = \mu$ throughout this exercise.

(a & b) $A \cup B = B \cup A$, and $A \cup (B \cup C) = (A \cup B) \cup C$.

(c) Let $A \cap B = \varnothing$. Then $\kappa + \lambda = |A \cup B|$. Considering the embedding $\mathrm{Id}_A\colon A \to A \cup B$. Then $|A| \leqslant |A \cup B|$, i.e., $\kappa \leqslant \kappa + \lambda$.

(d) Let $|A_1| = \kappa_1, |A_2| = \kappa_2, |B_1| = \lambda_1, |B_2| = \lambda_2$, $A_1 \cap B_1 = \varnothing = A_2 \cap B_2$, $|A_1| \leqslant |A_2|$, and $|B_1| \leqslant |B_2|$. Let $f: A_1 \to A_2$ and $g: B_1 \to B_2$ be two injections. Define $h: A_1 \cup B_1 \to A_2 \cup B_2$ by letting

$$h(x) = \begin{cases} f(x) & \text{if } x \in A_1 \\ g(x) & \text{if } x \in B_1. \end{cases}$$

Then h is an injection, and so $\kappa_1 + \lambda_1 \leqslant \kappa_2 + \lambda_2$.

(e) Let $f: A \times B \to B \times A$ with $f((a,b)) = (b,a)$ for all $(a,b) \in A \times B$. Then f is bijective, and so $|A \times B| = |B \times A|$, i.e., $\kappa \cdot \lambda = \lambda \cdot \kappa$.

(f) By letting $f: (a,(b,c)) \mapsto ((a,b),c)$ for all $(a,(b,c)) \in A \times (B \times C)$, we see that $|A \times (B \times C)| = |(A \times B) \times C|$; hence, $\kappa \cdot (\lambda \cdot \mu) = (\kappa \cdot \lambda) \cdot \mu$.

(g) $A \times (B \cup C) = (A \times B) \cup (A \times C)$.

(h) Pick $b \in B$ (since $\lambda > 0$). Define $f: A \to A \times \{b\}$ by letting for all $a \in A$:

$$f(a) = (a,b).$$

Then f is bijective. Since $A \times \{b\} \subseteq A \times B$, we have (h).

(i) Let $|A_1| = \kappa_1, |A_2| = \kappa_2, |B_1| = \lambda_1, |B_2| = \lambda_2$, $\kappa_1 \leqslant \kappa_2$, and $\lambda_1 \leqslant \lambda_2$. Let $f: A_1 \to A_2$ and $g: B_1 \to B_2$ be two injections. By defining $h: A_1 \times B_1 \to A_2 \times B_2$ with

$$h(a,b) = (f(a), g(b)),$$

we see that h is injective. Therefore, $\kappa_1 \cdot \lambda_1 \leqslant \kappa_2 \cdot \lambda_2$.

(j) In the book.

(k) $\kappa + \kappa \leqslant 2 \cdot \kappa \leqslant \kappa \cdot \kappa$ if $\kappa \geqslant 2$, by part (j) and (i).

(l) For every $a \in A$, let $f_a \in A^B$ be defined as $f_a(b) \equiv a$ for all $b \in B$. Then we define a function $F: A \to A^B$ by letting $F(a) = f_a$. Then F is injective and so $\kappa \leqslant \kappa^\lambda$ if $\lambda > 0$.

(m) Take $a_1, a_2 \in A$ (since $\kappa > 1$). For every $b \in B$, we define a function $f_b: B \to A$ by letting

$$f_b(x) = \begin{cases} a_1 & \text{if } x = b \\ a_2 & \text{if } x \neq b. \end{cases}$$

Then define a function $F: B \to A^B$ as $F(b) = f_b$. This function F is injective, and so $|B| \leqslant \left|A^B\right|$.

(n) Let $|A_1| = \kappa_1, |A_2| = \kappa_2, |B_1| = \lambda_1, |B_2| = \lambda_2$, $\kappa_1 \leqslant \kappa_2$, and $\lambda_1 \leqslant \lambda_2$. Let $f: A_1 \to A_2$ and $g: B_1 \to B_2$ be two injections. For any $k \in A_1^{B_1}$, we can pick a $h_k \in A_2^{B_2}$ such that

$$h_k(x) = \begin{cases} (f \circ k \circ g^{-1})(x) & \text{if } x \in g[B_1] \\ \widehat{b}_2 & \text{if } x \in A_2 \smallsetminus g[B_1], \end{cases}$$

where $\widehat{b}_2 \in B_2$. Then the function $F: A_1^{B_1} \to A_2^{B_2}$ defined by $f(k) = h_k$ is injective, and so $\kappa_1^{\lambda_1} \leqslant \kappa_2^{\lambda_2}$. □

► EXERCISE 145 (5.1.2). *Show that $\kappa^0 = 1$ and $\kappa^1 = \kappa$ for all κ.*

PROOF. $\kappa^0 = 1$ because $A^{\varnothing} = \langle\,\rangle$ for all A.

Let $|A| = \kappa$ and $B = \{b\}$. Then $A^{\{b\}} = \{b\} \times A$; that is, $\left|A^{\{b\}}\right| = |A|$. □

► EXERCISE 146 (5.1.3). *Show that $1^\kappa = 1$ for all κ and $0^\kappa = 0$ for all $\kappa > 0$.*

PROOF. Let $A = \{a\}$ and $|B| = \kappa$. In this case, $\{a\}^B = \{f : B \to \{a\} : f(b) = a \text{ for all } b \in B\}$; that is, $\left|\{a\}^B\right| = 1 = |\{a\}|$.

Since $\varnothing^B = \varnothing$ for all B, we have $\left|\varnothing^B\right| = 0 = |\varnothing|$. □

► EXERCISE 147 (5.1.4). *Prove that $\kappa^\kappa \leqslant 2^{\kappa\cdot\kappa}$.*

PROOF. Let $|A| = \kappa$. We look for an injection $F : A^A \to \{0,1\}^{A\times A}$. For every element $f \in A^A$, let $F(f) : A \times A \to \{0,1\}$ be defined as

$$F(f)(a,b) = \begin{cases} 0 & \text{if } b \neq f(a) \\ 1 & \text{if } b = f(a). \end{cases}$$

To verify F is injective, take arbitrary $f, f' \in A^A$ with $f \neq f'$. Then there exists $a \in A$ such that $f(a) \neq f'(a)$. For the pair $(a, f(a)) \in A \times A$,

$$F(f)(a, f(a)) = 1 \neq 0 = F(f')(a, f(a)).$$

Hence, $F(f) \neq F(f')$ whenever $f \neq f'$. Thus, $\kappa^\kappa \leqslant 2^{\kappa\cdot\kappa}$. □

► EXERCISE 148 (5.1.5). *If $|A| \leqslant |B|$ and if $A \neq \varnothing$, then there is a mapping of B onto A.*

PROOF. Let $f : A \to B$ be an injection, and let $a \in A$. Define $g : B \to A$ as

$$g(b) = \begin{cases} f^{-1}(b) & \text{if } b \in f[A] \\ a & \text{if } b \in B \smallsetminus f[A]. \end{cases}$$

It is evident that g is surjective. □

► EXERCISE 149 (5.1.6). *If there is a mapping of B onto A, then $2^{|A|} \leqslant 2^{|B|}$.*

PROOF. Let $g : B \to A$ be surjective. Define $f : \mathcal{P}(A) \to \mathcal{P}(B)$ as $f(X) = g^{-1}[X]$. Then f is injective and so $2^{|A|} = |\mathcal{P}(A)| \leqslant |\mathcal{P}|(B) = 2^{|B|}$. □

► EXERCISE 150 (5.1.7). *Use Cantor's Theorem to show that the "set of all sets" does not exist.*

PROOF. Suppose $\mathcal{U}$ is the "set of all sets". Then $Y = \mathcal{P}(\bigcup \mathcal{U}) \subseteq \bigcup \mathcal{U}$, and so $|Y| \leqslant \left|\bigcup \mathcal{U}\right|$. But Cantors' Theorem says that $|Y| > \left|\bigcup \mathcal{U}\right|$. A contradiction. □

▶ EXERCISE 151 (5.1.8). *Let X be a set and let f be a one-to-one mapping of X into itself such that $f[X] \subset X$. Then X is infinite.*

PROOF. $f: X \to f[X]$ is bijective, and so $|X| = |f[X]|$. If X is finite, it contradicts Lemma 4.2.2. □

▶ EXERCISE 152 (5.1.9). *Every countable set is Dedekind infinite.*

PROOF. It suffices to consider $\mathbb{N}$. Let $f: \mathbb{N} \to \mathbb{N} \setminus \{0\}$ be defined as $f(n) = n+1$. Thus, $\mathbb{N}$ is Dedekind infinite. □

▶ EXERCISE 153 (5.1.10). *If X contains a countable subset, then X is Dedekind infinite.*

PROOF. Let $A \subseteq X$ be countable. Then there is an bijection $f: \mathbb{N} \to A$. Define a function $g: X \to X$ by

$$g(f(n)) = f(n+1) \qquad \text{for } n \in \mathbb{N}$$
$$g(x) = x \qquad \text{for } x \in X \setminus A$$

(see Figure 5.1). By this construction, $g: X \to X \setminus \{g(0)\}$ is bijective.

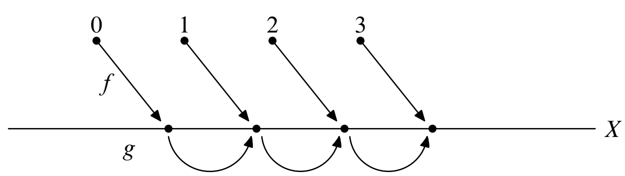


FIGURE 5.1. $f(0)$ is not in $\mathfrak{R}_g$.

□

▶ EXERCISE 154 (5.1.11). *If X is Dedekind infinite, then it contains a countable subset.*

PROOF. Let X be Dedekind infinite. Then there exists a bijection $f: X \to Y$, where $Y \subset X$. Pick $x \in X \setminus Y$. Let

$$x_0 = x, x_1 = f(x_0), \ldots, x_{n+1} = f(x_n), \ldots.$$

Then the set $\{x_n : n \in \mathbb{N}\}$ is countable. □

▶ EXERCISE 155 (5.1.12). *If A and B are Dedekind finite, then $A \cup B$ is Dedekind finite.*

PROOF. If A and B are Dedekind finite, then A and B does not contain a countable subset; hence, $A \cup B$ does not contain a countable subset, and so $A \cup B$ is Dedekind finite. □

► EXERCISE 156 (5.1.13). *If A and B are Dedekind finite, then $A \times B$ is Dedekind finite.*

PROOF. If A and B are Dedekind finite, then A and B does not contain a countable subset; hence, $A \times B$ does not contain a countable subset, and so $A \times B$ is Dedekind finite. □

► EXERCISE 157 (5.1.14). *If A is infinite, then $\mathcal{P}(\mathcal{P}(A))$ is Dedekind infinite.*

PROOF. For each $n \in \mathbb{N}$, let

$$S_n = \left\{X \subset A : |X| = n\right\}.$$

The set $\{S_n : n \in \mathbb{N}\}$ is a countable subset of $\mathcal{P}(\mathcal{P}(A))$, and hence $\mathcal{P}(\mathcal{P}(A))$ is Dedekind infinite. □

5.2 THE CARDINALITY OF THE CONTINUUM

► EXERCISE 158 (5.2.1). *Prove that the set of all finite sets of reals has cardinality $\mathfrak{c}$.*

PROOF. Every finite set of reals can be written as a finite union of open intervals with rational endpoints. For example, we can write $\{a, b, c\}$ as $(a, b) \cup (b, c)$. Thus, the cardinality of the set of all finite sets of reals is $\mathfrak{c}$. □

► EXERCISE 159 (5.2.2). *A real number x is* algebraic *if it is a solution of some equation*

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0, \qquad (*)$$

where $a_0, \ldots, a_n$ are integers. If x is not algebraic, it is called transcendental. *Show that the set of all algebraic numbers is countable and hence the set of all transcendental numbers has cardinality $\mathfrak{c}$.*

PROOF. Let $\mathcal{A}_n$ denote the set of algebraic numbers that satisfy polynomials of the form $a_k x^k + \cdots + a_1 x + a_0$ where $k < n$ and $\max\{|a_j|\} < n$. Note that there are at most n^n polynomials of this form, and each one has at most n roots. Hence, $\mathcal{A}_n$ is a finite set having at most $n^{n+1} < \aleph_0$ elements. Let $\mathcal{A}$ denote the set of all algebraic numbers. Then $|\mathcal{A}| = \left|\bigcup_{n \in \mathbb{N}} \mathcal{A}_n\right| \leqslant \aleph_0 \cdot \aleph_0 = \aleph_0$.

On the other hand, consider the following set of algebraic numbers:

$$\mathcal{A}' = \{x \in \mathbb{R} : a_0 + x = 0, a_0 \in \mathbb{Z}\}.$$

Obviously, $|\mathcal{A}'| = |\mathbb{Z}|$ and so $|\mathcal{A}| \geqslant |\mathbb{Z}| = \aleph_0$. It follows from Cantor-Benstein Theorem that $|\mathcal{A}| = \aleph_0$. □

► EXERCISE 160 (5.2.4). *The set of all closed subsets of reals has cardinality $\mathfrak{c}$.*

PROOF. Let $\mathscr{C}$ be the set of closed sets in $\mathbb{R}$, and $\mathscr{O}$ the set of open sets in $\mathbb{R}$. A set $E \in \mathscr{C}$ iff $\mathbb{R} \setminus E \in \mathscr{O}$; that is, there exists a bijection $f : \mathscr{C} \to \mathscr{O}$ defined by $f(E) = \mathbb{R} \setminus E$. Thus, $|\mathscr{C}| = |\mathscr{O}| = \mathfrak{c}$ by Theorem 5.2.6(b). □

► EXERCISE 161 (5.2.5). *Show that, for* $n > 0$, $n \cdot 2^{\mathfrak{c}} = \aleph_0 \cdot 2^{\mathfrak{c}} = \mathfrak{c} \cdot 2^{\mathfrak{c}} = 2^{\mathfrak{c}} \cdot 2^{\mathfrak{c}} = (2^{\mathfrak{c}})^n = (2^{\mathfrak{c}})^{\aleph_0} = (2^{\mathfrak{c}})^{\mathfrak{c}} = 2^{\mathfrak{c}}$.

PROOF. We have

$$2^{\mathfrak{c}} \leqslant n \cdot 2^{\mathfrak{c}} \leqslant \aleph_0 \cdot 2^{\mathfrak{c}} \leqslant \mathfrak{c} \cdot 2^{\mathfrak{c}} \leqslant 2^{\mathfrak{c}} \cdot 2^{\mathfrak{c}} = 2^{\mathfrak{c}+\mathfrak{c}} = 2^{\mathfrak{c}},$$

$$2^{\mathfrak{c}} \leqslant (2^{\mathfrak{c}})^n \leqslant (2^{\mathfrak{c}})^{\aleph_0} \leqslant (2^{\mathfrak{c}})^{\mathfrak{c}} \leqslant= 2^{\mathfrak{c}^2} = 2^{\mathfrak{c}},$$

and

$$2^{\mathfrak{c}} \leqslant n^{\mathfrak{c}} \leqslant \aleph_0^{\mathfrak{c}} \leqslant (2^{\aleph_0})^{\mathfrak{c}} = 2^{\aleph_0 \cdot \mathfrak{c}} = 2^{\mathfrak{c}}.$$

Thus, by the Cantor-Bernstein Theorem, we get the result. □

► EXERCISE 162 (5.2.6). *The cardinality of the set of all discontinuous functions is* $2^{\mathfrak{c}}$.

PROOF. Let $\mathscr{C}$ denote the set of all continuous functions, and $\mathscr{D}$ the set of all discontinuous functions. Suppose that $|\mathscr{D}| = \kappa < 2^{\mathfrak{c}}$. Then by Cantor's Theorem,

$$\left|\mathbb{R}^{\mathbb{R}}\right| = |\mathscr{D}| + |\mathscr{C}| = \kappa + \mathfrak{c} < 2^{\kappa+\mathfrak{c}} \leqslant 2^{2^{\mathfrak{c}}+\mathfrak{c}}.$$

Since

$$2^{\mathfrak{c}} + \mathfrak{c} \leqslant 2^{\mathfrak{c}} + 2^{\mathfrak{c}} = 2 \cdot 2^{\mathfrak{c}} = 2^{\mathfrak{c}}$$

by Exercise 161, we have

$$\left|\mathbb{R}^{\mathbb{R}}\right| < 2^{\mathfrak{c}} = \left|\mathbb{R}^{\mathbb{R}}\right|.$$

A contradiction. □

► EXERCISE 163 (5.2.7). *Construct a one-to-one mapping of* $\mathbb{R} \times \mathbb{R}$ *onto* $\mathbb{R}$.

PROOF. Using the hints. □

6

ORDINAL NUMBERS

6.1 Well-Ordered Sets

▶ Exercise 164 (6.1.1). *Give an example of a linearly ordered set* $(L, <)$ *and an initial segment* S *of* L *which is not of the form* $\{x : x < a\}$*, for any* $a \in L$.

Proof. We know from Lemma 6.1.2 that if L is a well-ordered set, then every initial segment is of the form $L[a]$ for some $a \in L$. Hence, we have to find a linear ordered set which is *not* well-ordered. We also know from Lemma 4.4.2 that every linear ordering on a finite set is a well-ordering. Therefore, our fist task is to find an infinite linear ordered $(L, <)$ which is not well-ordered.

As an example, let $L = \mathbb{R}$ and $S = (-\infty, 0]$. Then $(\mathbb{R}, <)$ is a linear ordered set, and S is an initial segment of L, but $S \neq \mathbb{R}[a]$ for any $a \in \mathbb{R}$. □

▶ Exercise 165 (6.1.2). $\omega + 1$ *is not isomorphic to* ω *(in the well-ordering by* $\in$*).*

Proof. We first show that $\omega = \mathbb{N}$ is an initial segment of $\omega + 1$. By definition, $\omega + 1 = \omega \cup \{\omega\}$, so $\omega \subset \omega + 1$. Choose any $\alpha \in \omega$, and let $\beta \in \alpha$. Both α and β are natural numbers, and so $\beta \in \omega$. Then, by Corollary 6.1.5 (a), $\omega + 1$ is not isomorphic to ω since $\omega + 1$ is a well-ordered sets. □

▶ Exercise 166 (6.1.3). *There exist* $2^{\aleph_0}$ *well-orderings of the set of all natural numbers.*

Proof. There are $\aleph_0^{\aleph_0} = \mathfrak{c}$ well-orderings on $\mathbb{N}$. □

▶ Exercise 167 (6.1.4). *For every infinite subset* A *of* $\mathbb{N}$*,* $(A, <)$ *is isomorphic to* $(\mathbb{N}, <)$.

Proof. Let $A \subseteq \mathbb{N}$ be infinite. Notice that $(A, <)$ is a well-ordered set, and A is not an initial segment of $\mathbb{N}$; for otherwise, $A = \mathbb{N}[n]$ for some $n \in \mathbb{N}$ and so A is finite.

A cannot be isomorphic to $\mathbb{N}[n]$ for all $n \in \mathbb{N}$ since $\mathbb{N}[n]$ is finite; similarly, $A[n]$ cannot be isomorphic to $\mathbb{N}$. Hence, by Theorem 6.1.3, A is isomorphic to $\mathbb{N}$. □

▶ EXERCISE 168 (6.1.5). *Let $(W_1, <_1)$ and $(W_2, <_2)$ be disjoint well-ordered sets, each isomorphic to $(\mathbb{N}, <)$. Show that the sum of the two linearly ordered sets is a well-ordering, and is isomorphic to the ordinal number $\omega + \omega = \{0, 1, 2, \ldots, \omega, \omega + 1, \omega + 2, \ldots\}$.*

PROOF. Let $(W, \prec)$ be the sum of $(W_1, <_1)$ and $(W_2, <_2)$. We have known that $(W, \prec)$ is a linearly ordered set. To see $(W, \prec)$ is well-ordered, take an arbitrary nonempty set $X \subset W$. Then $X = (W_1 \cap X) \cap (W_2 \cap X)$, and $(W_1 \cap X) \cap (W_2 \cap X) = \varnothing$. For $i = 1, 2$, if $W_i \cap X \neq \varnothing$, then it has a least element α_i. Let $\alpha = \min\{\alpha_1, \alpha_2\}$. Then α is the least element of X.

Let $f_i \colon W_i \to \mathbb{N}$, $i = 1, 2$, be two isomorphisms. To see $(W, \prec) \cong (\omega + \omega, <)$, let $f \colon W_1 \cup W_2 \to \omega + \omega$ be defined as

$$f(w) = \begin{cases} f_1(w) & \text{if } w \in W_1 \\ \omega + f_2(w) & \text{if } w \in W_2. \end{cases}$$

It is clear that f is an isomorphism and so $(W, <) \cong (\omega + \omega, <)$. □

▶ EXERCISE 169 (6.1.6). *Show that the lexicographic product $(\mathbb{N} \times \mathbb{N}, <)$ is isomorphic to $\omega \cdot \omega$.*

PROOF. Define a function $f \colon \mathbb{N} \times \mathbb{N} \to \omega \cdot \omega$ as follows: for an arbitrary $(m, n) \in \mathbb{N} \times \mathbb{N}$,

$$f(m, n) = \omega \cdot m + n.$$

Clearly, f is bijective. To see f is an isomorphism, let $(m, n) < (p, q)$. Then either $m < p$ or $m = p$ and $n < q$. For every case, $\omega \cdot m + n < \omega \cdot p + q$. □

▶ EXERCISE 170 (6.1.7). *Let $(W, <)$ be a well-ordered set, and let $a \notin W$. Extend $<$ to $W' = W \cup \{a\}$ by making a greater than all $x \in W$. Then W has smaller order type than W'.*

PROOF. We have $W'[a] = W$. Define a bijection $f \colon W \to W'[a]$ as $f(x) = x$ for all $x \in W$. Then f is an isomorphism. □

▶ EXERCISE 171 (6.1.8). *The sets $W = \mathbb{N} \times \{0, 1\}$ and $W' = \{0, 1\} \times \mathbb{N}$, ordered lexicographically, are nonisomorphic well-ordered sets.*

PROOF. See Figures 6.1 and 6.2. The first ordering is isomorphic to $(\omega, <)$, but the second ordering is isomorphic to $(\omega + \omega, <)$. Since $\omega + \omega$ is not isomorphic to ω (by Exercise 165, we get the result.

6.2 ORDINAL NUMBERS

REMARK. Let A be a nonempty set of ordinals. Take $\alpha \in A$, and consider the set $\alpha \cap A$.

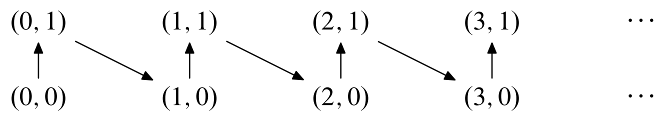


FIGURE 6.1. The lexicographic ordering on $\mathbb{N} \times \{0, 1\}$.

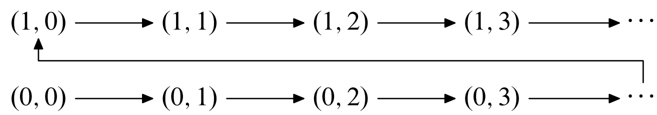


FIGURE 6.2. The lexicographic ordering on $\{0, 1\} \times \mathbb{N}$.

□

a. If $\alpha \cap A = \varnothing$, then α is the least element of A.

b. If $\alpha \cap A \neq \varnothing$, then γ, where γ is the least element of $\alpha \cap A$, is the least element of A.

PROOF. (a) If $\alpha \cap A = \varnothing$, then $\beta \notin \alpha$ for every $\beta \in A$. It follows from Theorem 6.2.6(c) that $\alpha \leqslant \beta$ for all $\beta \in A$. Hence, α is the least element of A.

(b) For every $\beta \in A$, if $\beta \notin \alpha$, then $\alpha \leqslant \beta$; if $\beta \in \alpha$, then $\beta < \alpha$. If $\alpha \cap A \neq \varnothing$, it has a least element γ in the ordering $\in_\alpha$; that is $\gamma \leqslant \beta$ for any $\beta \in \alpha \cap A$. Further, since $\gamma \in \alpha \cap A \subseteq \alpha$, we have $\gamma < \alpha$ and $\gamma \in A$. In sum,

$$\begin{cases} \gamma < \alpha \leqslant \beta & \text{if } \beta \in A \smallsetminus \alpha \\ \gamma \leqslant \beta & \text{if } \beta \in A \cap \alpha. \end{cases}$$

Hence, γ is the least element of A. □

► EXERCISE 172 (6.2.1). *A set X is transitive if and only if $X \subseteq \mathcal{P}(X)$.*

PROOF. Take an arbitrary $x \in X$. If X is transitive, then $x \subseteq X$, and so $x \in \mathcal{P}(X)$, i.e., $X \subseteq \mathcal{P}(X)$. On the other hand, if $X \subseteq \mathcal{P}(X)$, then $x \in X$ implies that $x \in \mathcal{P}(X)$, which is equivalent to $x \subseteq X$; hence X is transitive. □

► EXERCISE 173 (6.2.2). *A set X is transitive if and only if $\bigcup X \subseteq X$.*

PROOF. Take any $x \in \bigcup X$, then there exists $x_i \in X$ such that $x \in x_i$, that is, $x \in x_i \in X$; therefore, $x \in X$ if X is transitive and so $\bigcup X \subseteq X$. To see the converse direction, let $\bigcup X \subseteq X$. Take any $x \in \bigcup X$. There exists $x_i \in X$ such that $x \in x_i$; but $x \in X$ since $\bigcup X \subseteq X$, so X is transitive. □

► EXERCISE 174 (6.2.3). *Are the following sets transitive?*

a. $\{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}\}$,

b. $\{\varnothing, \{\varnothing\}, \{\{\varnothing\}\}, \{\varnothing, \{\varnothing\}\}\}$,

c. $\{\varnothing, \{\{\varnothing\}\}\}$.

Proof. (a) and (b) are transitive. However, (c) is not since $\{\varnothing\} \in \{\{\varnothing\}\}$, but $\{\varnothing\} \notin \{\varnothing, \{\{\varnothing\}\}\}$. □

► Exercise 175 (6.2.4). *Which of the following statements are true?*

a. *If X and Y are transitive, the $X \cup Y$ is transitive.*

b. *If X and Y are transitive, the $X \cap Y$ is transitive.*

c. *If $X \in Y$ and Y is transitive, then X is transitive.*

d. *If $X \subseteq Y$ and Y is transitive, then X is transitive.*

e. *If Y is transitive and $S \subseteq \mathcal{P}(Y)$, then $Y \cup S$ is transitive.*

Proof. (a), (b), and (e) are correct. □

► Exercise 176 (6.2.5). *If every $X \in S$ is transitive, then $\bigcup S$ is transitive.*

Proof. Let $u \in v \in \bigcup S$. Then there exists $X \in S$ such that $u \in v \in X$ and so $u \in X$ since X is transitive. Therefore, $u \in \bigcup S$, i.e., $\bigcup S$ is transitive. □

► Exercise 177 (6.2.7). *If a set of ordinals X does not have a greatest element, then* $\sup X$ *is a limit ordinal.*

Proof. If X does not have a greatest element, then $\sup X > \alpha$ for all $\alpha \in X$, and $\sup X$ is the least such ordinal. If there were β such that $\sup X = \beta + 1$, then β would be the greatest element of X. A contradiction. □

► Exercise 178 (6.2.8). *If X is a nonempty set of ordinals, then $\bigcap X$ is an ordinal. Moreover, $\bigcap X$ is the least element of X.*

Proof. If $u \in v \in \bigcap X$, then $u \in v \in \alpha$ for all $\alpha \in X$, and so $u \in \alpha$ for all $\alpha \in X$, i.e., $u \in \bigcap X$. Hence, $\bigcap X$ is transitive. It is evident to see that $\bigcap X$ is well-ordered. Thus, $\bigcap X$ is an ordinal. For every $\alpha \in X$, we have $\bigcap X \subseteq \alpha$; hence, $\bigcap X \leqslant \alpha$ for all $\alpha \in X$.

We finally show that $\bigcap X \in X$. If not, then $\bigcap X < \gamma$, where γ is the least element of X. It is impossible. □

References

[1] Apostol, Tom M. (1974) *Mathematical Analysis*: Pearson Education, 2nd edition. [58]

[2] Hrbacek, Karel and Thomas Jech (1999) *Introduction to Set Theory*, **220** of Pure and Applied Mathematics: A Series of Monographs and Textbooks, New York: Taylor & Francis Group, LLC, 3rd edition. [i]

[3] Jech, Thomas (2006) *Set Theory*, Springer Monographs in Mathematics, Berlin: Springer-Verlag, the third millennium edition. [67]

[4] Rudin, Walter (1976) *Principles of Mathematical Analysis*, New York: McGraw-Hill Companies, Inc. 3rd edition. [65]

Linear Algebra

A Solution Manual for
Axler (1997), Lax (2007), and Roman (2008)

Jianfei Shen

School of Economics, The University of New South Wales

Sydney, Australia 2009

I hear, I forget;
I see, I remember;
I do, I understand.
Old Proverb

Contents

Acronyms

Notation	Description
$U \trianglelefteq V$	U is a subspace of V
$\mathfrak{L}(X)$	The set of operators
$\mathfrak{R}_{\mathrm{T}}$	The range of T
$\mathfrak{N}_{\mathrm{T}}$	The null space of T
$\mathbb{F}$, K	The field on which a vector (linear) space is defined
$V \cong U$	V is isomorphic to U
$[\boldsymbol{x}] = \boldsymbol{x} + Y$	The coset Y in X and $\boldsymbol{x}$ is called a coset representative for $[\boldsymbol{x}]$
X/Y	The quotient space module Y
$\mathfrak{P}_n(\mathbb{F})$	The set of polynomials with degree $\leqslant n$, whose coefficients are in $\mathbb{F}$
$\mathrm{Sym}\,(X)$	The set of all permutations of the set X: the symmetric group on X
$\mathrm{sign}\,(\sigma)$	The signature of a permutation σ

Part I
Linear Algebra Done Right (Axler, 1997)

1

VECTOR SPACES

"As You Should Verify"

Remark. $P = \{p \in \mathfrak{P}(\mathbb{F}) : p(3) = 0\}$ is a subspace of $\mathfrak{P}(\mathbb{F})$.

Proof. The additive identity $0_{\mathfrak{P}(\mathbb{F})}$ is in the set; let $p, q \in P$, then $(p+q)(3) = p(3) + q(3) = 0$; for any $a \in \mathbb{F}$ and $p \in P$, we have $(ap)(3) = a \cdot 0 = 0$. □

Remark. If $U_1, \ldots, U_m$ are subspaces of V, then the sum $U_1 + \cdots + U_m$ is a subspace of V.

Proof. First, $\mathbf{0} \in U_i$ for all U_i implies that $\mathbf{0} = \mathbf{0} + \cdots + \mathbf{0} \in \sum_{i=1}^m U_i$. Now let $\boldsymbol{u}, \boldsymbol{v} \in \sum_{i=1}^m U_i$. Then $\boldsymbol{u} = \sum_{i=1}^m \boldsymbol{u}_i$ and $\boldsymbol{v} = \sum_{i=1}^m \boldsymbol{v}_i$, where $\boldsymbol{u}_i, \boldsymbol{v}_i \in U_i$, and so $\boldsymbol{u} + \boldsymbol{v} = \sum_{i=1}^m (\boldsymbol{u}_i + \boldsymbol{v}_i) \in \sum_{i=1}^m U_i$ since $\boldsymbol{u}_i + \boldsymbol{v}_i \in U_i$ for all i. Finally, let $\boldsymbol{u} = \sum_{i=1}^m \boldsymbol{u}_i \in \sum_{i=1}^m U_i$ and $a \in \mathbb{F}$. Then $a\boldsymbol{u} = \sum_{i=1}^m (a\boldsymbol{u}_i) \in \sum_{i=1}^m U_i$. □

Exercises

► Exercise 1 (1.1). *Suppose a and b are real numbers, not both 0. Find real numbers c and d such that $1/(a+bi) = c + di$.*

Solution. Note that for $\boldsymbol{z} \in \mathbb{C}$ with $\boldsymbol{z} \neq 0$, there exists a unique $\boldsymbol{w} \in \mathbb{C}$ such that $\boldsymbol{z}\boldsymbol{w} = 1$; that is, $\boldsymbol{w} = 1/\boldsymbol{z}$. Let $\boldsymbol{z} = a + bi$ and $\boldsymbol{w} = c + di$. Then

$$(a+bi)(c+di) = (ac - bd) + (ad + bc)i = 1 + 0i$$

yields

$$\begin{cases} ac - bd = 1, \\ ad + bc = 0 \end{cases} \implies \begin{cases} c = a/(a^2+b^2), \\ d = -b/(a^2+b^2). \end{cases}$$ □

► Exercise 2 (1.2). *Show that $(-1+\sqrt{3}i)/2$ is a cube root of 1.*

PROOF. We have

$$\begin{aligned}\left(\frac{-1+\sqrt{3}i}{2}\right)^3 &= \left(\frac{-1+\sqrt{3}i}{2}\right)^2 \cdot \left(\frac{-1+\sqrt{3}i}{2}\right)\\ &= \left(-\frac{1}{2}-\frac{\sqrt{3}}{2}i\right)\cdot\left(-\frac{1}{2}+\frac{\sqrt{3}}{2}i\right)\\ &= 1.\end{aligned}$$

□

► EXERCISE 3 (1.3). *Prove that* $-(-\boldsymbol{v}) = \boldsymbol{v}$ *for very* $\boldsymbol{v} \in V$.

PROOF. We have $\boldsymbol{v} + (-\boldsymbol{v}) = 0$, so by the uniqueness of additive inverse, the additive inverse of $-\boldsymbol{v}$, i.e., $-(-\boldsymbol{v})$, is $\boldsymbol{v}$. □

► EXERCISE 4 (1.4). *Prove that if* $a \in \mathbb{F}$, $\boldsymbol{v} \in V$, *and* $a\boldsymbol{v} = \boldsymbol{0}$, *then* $a = 0$ *or* $\boldsymbol{v} = \boldsymbol{0}$.

PROOF. Suppose that $\boldsymbol{v} \neq \boldsymbol{0}$ and $a \neq 0$. Then $\boldsymbol{v} = 1 \cdot \boldsymbol{v} = (a\boldsymbol{v})/a = \boldsymbol{0}/a = \boldsymbol{0}$. A contradiction. □

► EXERCISE 5 (1.5). *For each of the following subsets of* $\mathbb{F}^3$, *determine whether it is a subspace of* $\mathbb{F}^3$:

a. $U = \{(x_1, x_2, x_3) \in \mathbb{F}^3 : x_1 + 2x_2 + 3x_3 = 0\}$;

b. $U = \{(x_1, x_2, x_3) \in \mathbb{F}^3 : x_1 + 2x_2 + 3x_3 = 4\}$;

c. $U = \{(x_1, x_2, x_3) \in \mathbb{F}^3 : x_1 x_2 x_3 = 0\}$;

d. $U = \{(x_1, x_2, x_3) \in \mathbb{F}^3 : x_1 = 5x_3\}$.

SOLUTION. (a) Additive Identity: $\boldsymbol{0} \in U$; Closed under Addition: Let $\boldsymbol{x}, \boldsymbol{y} \in U$, then $\boldsymbol{x} + \boldsymbol{y} = (x_1 + y_1, x_2 + y_2, x_3 + y_3)$, and $(x_1 + y_1) + 2(x_2 + y_2) + 3(x_3 + y_3) = (x_1 + 2x_2 + 3x_3) + (y_1 + 2y_2 + 3y_3) = 0 + 0 = 0$; that is, $\boldsymbol{x} + \boldsymbol{y} \in U$. Closed under Scalar Multiplication: Pick any $a \in \mathbb{F}$ and $\boldsymbol{x} \in U$. Then $ax_1 + 2\cdot(ax_2) + 3\cdot(ax_3) = a \cdot (x_1 + 2x_2 + 3x_3) = 0$, i.e., $a\boldsymbol{x} \in U$. In sum, U is a subspace of $\mathbb{F}^3$, and actually, U is a hyperplane through the $\boldsymbol{0}$.

(b) U is not a subspace because $\boldsymbol{0} \notin U$.

(c) Let $\boldsymbol{x} = (1, 1, 0)$ and $\boldsymbol{y} = (0, 0, 1)$. Then $\boldsymbol{x}, \boldsymbol{y} \in U$, but $\boldsymbol{x} + \boldsymbol{y} = (1, 1, 1) \notin U$.

(d) $\boldsymbol{0} \in U$; Let $\boldsymbol{x}, \boldsymbol{y} \in U$. Then $x_1 + y_1 = 5(x_3 + y_3)$. Let $a \in \mathbb{F}$ and $\boldsymbol{x} \in U$. Then $ax_1 = a \cdot 5x_3$. □

► EXERCISE 6 (1.6). *Give an example of a nonempty subset* U *of* $\mathbb{R}^2$ *such that* U *is closed under addition and under taking additive inverses (meaning* $-\boldsymbol{u} \in U$ *whenever* $\boldsymbol{u} \in U$*), but* U *is not a subspace of* $\mathbb{R}^2$.

SOLUTION. Let $U = \mathbb{Z}^2$, which is not closed under scalar multiplication. □

► EXERCISE 7 (1.7). *Give an example of a nonempty subset* U *of* $\mathbb{R}^2$ *such that* U *is closed under scalar multiplication, but* U *is not a subspace of* $\mathbb{R}^2$.

SOLUTION. Let

$$U = \left\{(x,y) \in \mathbb{R}^2 \colon x = y\right\} \cup \left\{(x,y) \in \mathbb{R}^2 \colon x = -y\right\}.$$

In this case, $(x,x) + (x,-x) = (2x,0) \notin U$ unless $x = 0$. □

► EXERCISE 8 (1.8). *Prove that the intersection of any collection of subspaces of V is a subspace of V.*

PROOF. Let $\{U_i\}$ be a collection of subspaces of V. (i) Every U_i is a subspace, then $\mathbf{0} \in U_i$ for all i and so $\mathbf{0} \in \bigcap U_i$. (ii) Let $\boldsymbol{x}, \boldsymbol{y} \in \bigcap U_i$. Then $\boldsymbol{x}, \boldsymbol{y} \in U_i$ for all i and so $\boldsymbol{x} + \boldsymbol{y} \in U_i$, which implies that $\boldsymbol{x} + \boldsymbol{y} \in \bigcap U_i$. (iii) Let $a \in \mathbb{F}$ and $\boldsymbol{x} \in \bigcap U_i$. Then $a\boldsymbol{x} \in U_i$ for all i implies that $a\boldsymbol{x} \in \bigcap U_i$. □

► EXERCISE 9 (1.9). *Prove that the union of two subspaces of V is a subspace of V if and only if one of the subspaces is contained in the other.*

PROOF. Let U and W be two subspaces of V. The "If" part is trivial. So we focus on the "Only if" part. Let $U \cup W$ be a subspace. Suppose $U \not\subseteq W$ and $W \not\subseteq U$. Pick $\boldsymbol{x} \in U \setminus W$ and $\boldsymbol{y} \in W \setminus U$. Then $\boldsymbol{x} + \boldsymbol{y} \notin U$; for otherwise $\boldsymbol{y} = (\boldsymbol{x} + \boldsymbol{y}) - \boldsymbol{x} \in U$; similarly, $\boldsymbol{x} + \boldsymbol{y} \notin W$. But then $\boldsymbol{x} + \boldsymbol{y} \notin U \cup W$, which contradicts the fact that $\boldsymbol{x}, \boldsymbol{y} \in U \cup W$ and $U \cup W$ is a subspace.

A nontrivial vector space V over an infinite field $\mathbb{F}$ is not the union of a finite number of proper subspaces; see Roman (2008, Theorem 1.2). □

► EXERCISE 10 (1.10). *Suppose that U is a subspace of V. What is $U + U$?*

SOLUTION. Since $U \subset U$ and $U + U$ is the smallest subspace containing U, we have $U + U \subset U$; on the other hand, $U \subset U + U$ is clear. Hence, $U + U = U$. □

► EXERCISE 11 (1.11). *Is the operation of addition on the subspaces of V commutative? Associative?*

SOLUTION. Yes. Let U_1, U_2 and U_3 be subspaces of V.

$$\begin{aligned} U_1 + U_2 &= \{\boldsymbol{u}_1 + \boldsymbol{u}_2 \colon \boldsymbol{u}_1 \in U_1, \boldsymbol{u}_2 \in U_2\} \\ &= \{\boldsymbol{u}_2 + \boldsymbol{u}_1 \colon \boldsymbol{u}_2 \in U_2, \boldsymbol{u}_1 \in U_1\} \\ &= U_2 + U_1. \end{aligned}$$

Similarly for associativity. □

► EXERCISE 12 (1.12). *Does the operation of addition on the subspaces of V have an additive identity? Which subspaces have additive inverses?*

SOLUTION. The set $\{\mathbf{0}\}$ is the additive identity: $U + \{\mathbf{0}\} = \{\boldsymbol{u} + \mathbf{0} \colon \boldsymbol{u} \in U\} = U$.

Only the set $\{\mathbf{0}\}$ has additive inverse. Suppose that U is a subspace, and its additive inverse is W, i.e., $U + W = \{\boldsymbol{u} + \boldsymbol{w} \colon \boldsymbol{u} \in U \text{ and } \boldsymbol{w} \in W\} = \{\mathbf{0}\}$. Since

$\mathbf{0} \in U$, we have $\mathbf{0} + \boldsymbol{w} = \mathbf{0}$ for all $\boldsymbol{w} \in W$, which means that $W = \{\mathbf{0}\}$. But it is clearly that $U + \{\mathbf{0}\} = \{\mathbf{0}\}$ iff $U = \{\mathbf{0}\}$. □

► EXERCISE 13 (1.13). *Prove or give a counterexample: if U_1, U_2, W are subspaces of V such that $U_1 + W = U_2 + W$, then $U_1 = U_2$.*

SOLUTION. Suppose $U_1, U_2 \subseteq W$. Then $U_1 + W = U_2 + W$ for any U_1 and U_2. Hence, the statement is false in general. □

► EXERCISE 14 (1.14). *Suppose U is the subspace of $\mathfrak{P}(\mathbb{F})$ consisting of all polynomials p of the form $p(z) = az^2 + bz^5$, where $a, b \in \mathbb{F}$. Find a subspace W of $\mathfrak{P}(\mathbb{F})$ such that $\mathfrak{P}(\mathbb{F}) = U \oplus W$.*

SOLUTION. Let

$$W = \left\{ p \in \mathfrak{P}(\mathbb{F}) \colon p(z) = a_0 + a_1 z + a_3 z^3 + a_4 z^4 \right\}.$$ □

► EXERCISE 15 (1.15). *Prove or give a counterexample: if U_1, U_2, W are subspaces of V such that $V = U_1 \oplus W$ and $V = U_2 \oplus W$, then $U_1 = U_2$.*

SOLUTION. Let $V = \mathbb{R}^2$, $W = \{(x, 0) \in \mathbb{R}^2 \colon x \in \mathbb{R}\}$, $U_1 = \{(x, x) \in \mathbb{R}^2 \colon x \in \mathbb{R}\}$, and $U_2 = \{(x, -x) \in \mathbb{R}^2 \colon x \in \mathbb{R}\}$. Then

$$\begin{aligned} U_1 + W &= \left\{ (x + y, x) \in \mathbb{R}^2 \colon x, y \in \mathbb{R} \right\} = \mathbb{R}^2 = V, \\ U_2 + W &= \left\{ (x + y, -x) \in \mathbb{R}^2 \colon x, y \in \mathbb{R} \right\} = \mathbb{R}^2 = V, \\ U_i \cap W &= \{(0, 0)\}, i = 1, 2. \end{aligned}$$

Therefore, $V = U_i \oplus W$ for $i = 1, 2$, but $U_1 \neq U_2$. □

2

FINITE-DIMENSIONAL VECTOR SPACES

"As You Should Verify"

Remark (p.22). The span of any list of vectors in V is a subspace of V.

Proof. If $U = ()$, define $\operatorname{span}(U) = \{\mathbf{0}\}$, which is a subspace of V. Now let $U = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ be a list of vectors in V. Then $\operatorname{span}(U) = \{\sum_{i=1}^{n} a_i \boldsymbol{v}_i : a_i \in \mathbb{F}\}$. (i) $\mathbf{0} = \sum_{i=1}^{n} 0\boldsymbol{v}_i \in \operatorname{span}(U)$. (ii) Let $\boldsymbol{u} = \sum_{i=1}^{n} a_i \boldsymbol{v}_i$ and $\boldsymbol{v} = \sum_{i=1}^{n} b_i \boldsymbol{v}_i$. Then $\boldsymbol{u} + \boldsymbol{v} = \sum_{i=1}^{n} (a_i + b_i)\boldsymbol{v}_i \in \operatorname{span}(U)$. (iii) For every $\boldsymbol{u} = \sum_{i=1}^{n} a_i \boldsymbol{v}_i$, we have $a\boldsymbol{u} = \sum_{i=1}^{n} (aa_i)\boldsymbol{v}_i \in \operatorname{span}(U)$. □

Remark (p.23). $\mathfrak{P}_m(\mathbb{F})$ is a subspace of $\mathfrak{P}(\mathbb{F})$.

Proof. (i) $0_{\mathfrak{P}(\mathbb{F})} \in \mathfrak{P}_m(\mathbb{F})$ since its degree is $-\infty < m$ by definition. (ii) Let $p = \sum_{i=0}^{\ell} a_\ell z^\ell$ and $q = \sum_{j=0}^{n} b_j z^j$, where $\ell, n \leqslant m$ and $a_\ell, b_n \neq 0$. Without loss of generality, suppose $\ell \geqslant n$. Then $p + q = \sum_{i=0}^{n} (a_i + b_i) z^i + \sum_{j=n+1}^{\ell} a_j z^j \in \mathfrak{P}_m(\mathbb{F})$. (iii) It is easy to see that if $p \in \mathfrak{P}_m(\mathbb{F})$ then $ap \in \mathfrak{P}_m(\mathbb{F})$. □

Exercises

► Exercise 16 (2.1). *Prove that if* $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ *spans* V*, then so does the list* $(\boldsymbol{v}_1 - \boldsymbol{v}_2, \boldsymbol{v}_2 - \boldsymbol{v}_3, \dots, \boldsymbol{v}_{n-1} - \boldsymbol{v}_n, \boldsymbol{v}_n)$ *obtained by subtracting from each vector (except the last one) the following vector.*

Proof. We first show that $\operatorname{span}(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n) \subseteq \operatorname{span}(\boldsymbol{v}_1 - \boldsymbol{v}_2, \dots, \boldsymbol{v}_{n-1} - \boldsymbol{v}_n, \boldsymbol{v}_n)$. Suppose that $V = \operatorname{span}(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$. Then, for any $\boldsymbol{v} \in V$, there exist $a_1, \dots, a_n \in \mathbb{F}$ such that

$$\begin{aligned}
\boldsymbol{v} &= a_1\boldsymbol{v}_1 + a_2\boldsymbol{v}_2 + \cdots + a_n\boldsymbol{v}_n \\
&= a_1(\boldsymbol{v}_1 - \boldsymbol{v}_2) + (a_1 + a_2)\boldsymbol{v}_2 + a_3\boldsymbol{v}_3 + \cdots + a_n\boldsymbol{v}_n \\
&= a_1(\boldsymbol{v}_1 - \boldsymbol{v}_2) + (a_1 + a_2)(\boldsymbol{v}_2 - \boldsymbol{v}_3) + (a_1 + a_2 + a_3)\boldsymbol{v}_3 + a_4\boldsymbol{v}_4 + \cdots + a_n\boldsymbol{v}_n \\
&= \sum_{i=1}^{n-1}\left[\left(\sum_{j=1}^{i} a_j\right)(\boldsymbol{v}_i - \boldsymbol{v}_{i+1})\right] + a_n\boldsymbol{v}_n \\
&\in \operatorname{span}(\boldsymbol{v}_1 - \boldsymbol{v}_2, \boldsymbol{v}_2 - \boldsymbol{v}_3, \ldots, \boldsymbol{v}_{n-1} - \boldsymbol{v}_n, \boldsymbol{v}_n).
\end{aligned}$$

For the converse direction, let $\boldsymbol{u} \in \operatorname{span}(\boldsymbol{v}_1 - \boldsymbol{v}_2, \boldsymbol{v}_2 - \boldsymbol{v}_3, \ldots, \boldsymbol{v}_{n-1} - \boldsymbol{v}_n, \boldsymbol{v}_n)$. Then there exist $b_1, \ldots, b_n \in \mathbb{F}$ such that

$$\begin{aligned}
\boldsymbol{u} &= b_1(\boldsymbol{v}_1 - \boldsymbol{v}_2) + b_2(\boldsymbol{v}_2 - \boldsymbol{v}_3) + \cdots + b_{n-1}(\boldsymbol{v}_{n-1} - \boldsymbol{v}_n) + b_n \cdot \boldsymbol{v}_n \\
&= b_1\boldsymbol{v}_1 + (b_2 - b_1)\boldsymbol{v}_2 + (b_3 - b_2)\boldsymbol{v}_3 + \cdots + (b_n - b_{n-1})\boldsymbol{v}_n \\
&\in \operatorname{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n).
\end{aligned}$$

□

► EXERCISE 17 (2.2). *Prove that if $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ is linearly independent in V, then so is the list $(\boldsymbol{v}_1 - \boldsymbol{v}_2, \boldsymbol{v}_2 - \boldsymbol{v}_3, \ldots, \boldsymbol{v}_{n-1} - \boldsymbol{v}_n, \boldsymbol{v}_n)$ obtained by subtracting from each vector (except the last one) the following vector.*

PROOF. Let

$$0 = \sum_{i=1}^{n-1} a_i(\boldsymbol{v}_i - \boldsymbol{v}_{i+1}) + a_n\boldsymbol{v}_n = a_1\boldsymbol{v}_1 + (a_2 - a_1)\boldsymbol{v}_2 + \cdots + (a_n - a_{n-1})\boldsymbol{v}_n.$$

Since $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ is linear independent, we have $a_1 = a_2 - a_1 = \cdots = a_n - a_{n-1} = 0$, i.e., $a_1 = a_2 = \cdots = a_n = 0$. □

► EXERCISE 18 (2.3). *Suppose $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ is linearly independent in V and $\boldsymbol{w} \in V$. Prove that if $(\boldsymbol{v}_1 + \boldsymbol{w}, \ldots, \boldsymbol{v}_n + \boldsymbol{w})$ is linearly dependent, then $\boldsymbol{w} \in \operatorname{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$.*

PROOF. If $(\boldsymbol{v}_1 + \boldsymbol{w}, \ldots, \boldsymbol{v}_n + \boldsymbol{w})$ is linearly dependent, then there exists a list $(a_1, \ldots, a_n) \neq \mathbf{0}$ such that

$$\sum_{i=1}^{n} a_i\,(\boldsymbol{v}_i + \boldsymbol{w}) = \sum_{i=1}^{n} a_i\boldsymbol{v}_i + \left(\sum_{i=1}^{n} a_i\right)\boldsymbol{w} = 0. \tag{2.1}$$

Since $(a_1, \ldots, a_n) \neq \mathbf{0}$, we know that $\sum_{i=1}^{n} a_i \neq 0$. It follows from (2.1) that

$$\boldsymbol{w} = \sum_{i=1}^{n}\left(-a_i \bigg/ \sum_{j=1}^{n} a_j\right)\boldsymbol{v}_i \in \operatorname{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n).$$

□

► EXERCISE 19 (2.4). *Suppose m is a positive integer. Is the set consisting of $\mathbf{0}$ and all polynomials with coefficients in $\mathbb{F}$ and with degree equal to m a subspace of $\mathfrak{P}(\mathbb{F})$?*

SOLUTION. No. Consider p, q with

$$p(z) = a_0 + a_1 z + \cdots + a_m z^m,$$
$$q(z) = b_0 + b_1 z + \cdots - a_m z^m,$$

where $a_m \neq 0$. Then $p(z)+q(z) = (a_0+b_0)+(a_1+b_1)z+\cdots+(a_{m-1}+b_{m-1})z^{m-1}$, whose degree is less than or equal to $m-1$. Hence, this set of polynomials with degree equal to m is not closed under addition. □

► EXERCISE 20 (2.5). *Prove that $\mathbb{F}^\infty$ is infinite dimensional.*

PROOF. Suppose that $\mathbb{F}^\infty$ is finite dimensional. Then every linearly independent list of vectors in a finite dimensional vector space can be extended to a basis of the vector space. Consider the following list

$$((1,0,0,0,\ldots),(0,1,0,0,\ldots),(0,0,1,0,\ldots),\ldots,(0,\ldots,1,0,\ldots)),$$

where each vector is in $\mathbb{F}^\infty$, and the length of the above list is n. It is easy to show that this list is linearly independent, but it can not be expanded to a basis of $\mathbb{F}^\infty$. □

► EXERCISE 21 (2.6). *Prove that the real vector space consisting of all continuous real-valued functions on the interval* $[0,1]$ *is infinite dimensional.*

PROOF. Consider the following set $\{p(z) \in \mathfrak{P}(\mathbb{F}) : z \in [0,1]\}$, which is a subspace, but is infinite dimensional. □

► EXERCISE 22 (2.7). *Prove that V is infinite dimensional if and only if there is a sequence $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots$ of vectors in V such that $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ is linearly independent for every positive integer n.*

PROOF. Let V be infinite dimensional. Clearly, there exists a nonzero vector $\boldsymbol{v}_1 \in V$; for otherwise, $V = \{\mathbf{0}\}$ and so V is finite dimensional. Since V is infinite dimensional, $\operatorname{span}(\boldsymbol{v}_1) \neq V$; hence there exists $\boldsymbol{v}_2 \in V \smallsetminus \operatorname{span}(\boldsymbol{v}_1)$; similarly, $\operatorname{span}(\boldsymbol{v}_1, \boldsymbol{v}_2) \neq V$; thus we can choose $\boldsymbol{v}_3 \in V \smallsetminus \operatorname{span}(\boldsymbol{v}_1, \boldsymbol{v}_2)$. We thus construct an infinite sequence $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots$

We then use the Induction Principle to prove that for every positive integer n, the list $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ is linearly independent. Obviously, $\boldsymbol{v}_1$ is linear independent since $\boldsymbol{v}_1 \neq \mathbf{0}$. Let us assume that $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ is linear independent for some positive integer n. We now show that $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n, \boldsymbol{v}_{n+1})$ is linear independent. If not, then there exist $a_1, \ldots, a_n, a_{n+1} \in \mathbb{F}$, not all 0, such that $\sum_{i=1}^{n+1} a_i \boldsymbol{v}_i = \mathbf{0}$. We must have $a_{n+1} \neq 0$: if $a_{n+1} = 0$, then $\sum_{i=1}^{n} a_i \boldsymbol{v}_i = 0$ implies that $a_1 = \cdots = a_n = a_{n+1} = 0$ since $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ is linear independent by the induction hypothesis. Hence,

$$\boldsymbol{v}_{n+1} = \sum_{i=1}^{n} (-a_i / a_{n+1}) \boldsymbol{v}_i,$$

i.e. $\boldsymbol{v}_{n+1} \in \operatorname{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$, which contradicts the construction of $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{n+1})$.

Conversely, assume that there exists an infinite sequence $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots$ of vectors in V, and $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ is linearly independent for any positive integer n. Suppose V is finite dimensional; that is, there is a spanning list of vectors $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m)$ of V, and such that the length of every linearly independent list of vectors is less than or equal to m (by Theorem 2.6). A contradiction. □

► EXERCISE 23 (2.8). *Let U be the subspace of $\mathbb{R}^5$ defined by*

$$U = \left\{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5 : x_1 = 3x_2 \text{ and } x_3 = 7x_4\right\}.$$

Find a basis of U.

PROOF. A particular basis of U can be $((3, 1, 0, 0, 0), (0, 0, 7, 1, 0), (0, 0, 0, 0, 1))$. □

► EXERCISE 24 (2.9). *Prove or disprove: there exists a basis (p_0, p_1, p_2, p_3) of $\mathfrak{P}_3(\mathbb{F})$ such that none of the polynomials p_0, p_1, p_2, p_3 has degree 2.*

PROOF. Notice that $p_0 = 1$, $p_1 = z$, $p_2' = z^2$, and $p_3 = z^3$ is a standard basis of $\mathfrak{P}_3(\mathbb{F})$, but p_2' has degree 2. So we can let $p_2 = p_2' + p_3 = z^2 + z^3$. Then $\operatorname{span}(p_0, p_1, p_2, p_3) = \mathfrak{P}_3(\mathbb{F})$ and so (p_0, p_1, p_2, p_3) is a basis of $\mathfrak{P}_3(\mathbb{F})$ by Theorem 2.16. □

► EXERCISE 25 (2.10). *Suppose that V is finite dimensional, with $\dim V = n$. Prove that there exist one-dimensional subspaces $U_1, \ldots, U_n$ of V such that*

$$V = U_1 \oplus \cdots \oplus U_n.$$

PROOF. Let $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ be a basis of V. For each $i = 1, \ldots, n$, let $U_i = \operatorname{span}(\boldsymbol{v}_i)$. Then each U_i is a subspace of V and so $U_1 + \cdots + U_n \subseteq V$. Clearly, $\dim V = \sum_{i=1}^n \dim U_i = n$. By Proposition 2.19, it suffices to show that $V \subseteq U_1 + \cdots + U_n$. It follows because for every $\boldsymbol{v} \in V$,

$$\boldsymbol{v} = \sum_{i=1}^n a_i \boldsymbol{v}_i \in U_1 + \cdots + U_n.$$ □

► EXERCISE 26 (2.11). *Suppose that V is finite dimensional and U is a subspace of V such that $\dim U = \dim V$. Prove that $U = V$.*

PROOF. Let $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ be a basis of U. Since $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ is linearly independent in V and the length of $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ is equal to $\dim V$, it is a basis of V. Therefore, $V = \operatorname{span}(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) = U$. □

► EXERCISE 27 (2.12). *Suppose that $p_0, p_1, \ldots, p_m$ are polynomials in $\mathcal{P}_m(\mathbb{F})$ such that $p_j(2) = 0$ for each j. Prove that $(p_0, p_1, \ldots, p_m)$ is not linearly independent in $\mathfrak{P}_m(\mathbb{F})$.*

PROOF. $\dim \mathfrak{P}_m(\mathbb{F}) = m+1$ since $(1, z, \ldots, z^m)$ is a basis of $\mathfrak{P}_m(\mathbb{F})$. If $(p_0, \ldots, p_m)$ is linear independent, then it is a basis of $\mathfrak{P}_m(\mathbb{F})$ by Proposition 2.17. Then $p = \sum_{i=0}^{m} p_i$ for every $p \in \mathfrak{P}_m(\mathbb{F})$. Take an arbitrary $p \in \mathfrak{P}_m(\mathbb{F})$ with $p(2) \neq 0$ and we get a contradiction. □

► EXERCISE 28 (2.13). *Suppose U and W are subspaces of $\mathbb{R}^8$ such that $\dim U = 3$, $\dim W = 5$, and $U + W = \mathbb{R}^8$. Prove that $U \cap W = \{\mathbf{0}\}$.*

PROOF. Since $\mathbb{R}^8 = U + W$ and $\dim \mathbb{R}^8 = \dim U + \dim V$, we have $\mathbb{R}^8 = U \oplus V$ by Proposition 2.19; then Proposition 1.9 implies that $U \cap W = \{\mathbf{0}\}$. □

► EXERCISE 29 (2.14). *Suppose that U and W are both five-dimensional subspaces of $\mathbb{R}^9$. Prove that $U \cap W \neq \{\mathbf{0}\}$.*

PROOF. If $U \cap W = \{\mathbf{0}\}$, then $\dim U + W = \dim U + \dim W - \dim U \cap W = 5 + 5 - 0 = 10 > 9$; but $U + W \subseteq \mathbb{R}^9$. A contradiction. □

► EXERCISE 30 (2.15). *Prove or give a counterexample that*

$$\begin{aligned}\dim U_1 + U_2 + U_3 = {} & \dim U_1 + \dim U_2 + \dim U_3 \\ & - \dim U_1 \cap U_2 - \dim U_1 \cap U_3 - \dim U_2 \cap U_3 \\ & + \dim U_1 \cap U_2 \cap U_3.\end{aligned}$$

SOLUTION. We construct a counterexample to show the proposition is faulse. Let

$$\begin{aligned} U_1 &= \left\{(x, 0) \in \mathbb{R}^2 \colon x \in \mathbb{R}\right\}, \\ U_2 &= \left\{(0, x) \in \mathbb{R}^2 \colon x \in \mathbb{R}\right\}, \\ U_3 &= \left\{(x, x) \in \mathbb{R}^2 \colon x \in \mathbb{R}\right\}. \end{aligned}$$

Then $U_1 \cap U_2 = U_1 \cap U_3 = U_2 \cap U_3 = U_1 \cap U_2 \cap U_3 = \{(0, 0)\}$; hence

$$\dim U_1 \cap U_2 = \dim U_1 \cap U_3 = \dim U_2 \cap U_3 = \dim U_1 \cap U_2 \cap U_3 = 0.$$

But $\dim U_1 + U_2 + U_3 = 2$ since $U_1 + U_2 + U_3 = \mathbb{R}^2$. □

► EXERCISE 31 (2.16). *Prove that if V is finite dimensional and $U_1, \ldots, U_m$ are subspaces of V, then $\dim U_1 + \cdots + U_m \leqslant \sum_{i=1}^{m} \dim U_i$.*

PROOF. Let $(\boldsymbol{u}_i^1, \ldots, \boldsymbol{u}_i^{n_i})$ be a basis of U_i for each $i = 1, \ldots, m$. Then

$$\sum_{i=1}^{m} \dim U_i = \sum_{i=1}^{m} n_i.$$

Let

$$(\boldsymbol{u}_1^1, \ldots, \boldsymbol{u}_1^{n_1}, \ldots, \boldsymbol{u}_m^1, \ldots, \boldsymbol{u}_m^{n_m}) = B.$$

Clearly, $U_1 + \cdots + U_m = \operatorname{span}(B)$, and $\dim \operatorname{span}(B) \leqslant \sum_{i=1}^{m} n_i$ by Theorem 2.10. Therefore, $\dim U_1 + \cdots + U_m \leqslant \sum_{i=1}^{m} \dim U_i$. □

► EXERCISE 32 (2.17). *Suppose V is finite dimensional. Prove that if $U_1, \ldots, U_m$ are subspaces of V such that $V = U_1 \oplus \cdots \oplus U_m$, then $\dim V = \sum_{i=1}^{m} \dim U_i$.*

PROOF. Let the list $(\boldsymbol{u}_i^1, \ldots, \boldsymbol{u}_i^{n_i})$ be a basis of U_i for all $i = 1, \ldots, m$. Then $\sum_{i=1}^{m} \dim U_i = \sum_{i=1}^{m} n_i$. Let

$$(\boldsymbol{u}_1^1, \ldots, \boldsymbol{u}_1^{n_1}, \ldots, \boldsymbol{u}_m^1, \ldots, \boldsymbol{u}_m^{n_m}) = U.$$

Then $\operatorname{span}(U) = V$. We show that $(\boldsymbol{u}_1^1, \ldots, \boldsymbol{u}_1^{n_1}, \ldots, \boldsymbol{u}_m^1, \ldots, \boldsymbol{u}_m^{n_m})$ is linear independent. Let

$$\mathbf{0} = \underbrace{(a_1^1 \boldsymbol{u}_1^1 + \cdots + a_1^{n_1} \boldsymbol{u}_1^{n_1})}_{\boldsymbol{u}_1} + \cdots + \underbrace{(a_m^1 \boldsymbol{u}_m^1 + \cdots + a_m^{n_m} \boldsymbol{u}_m^{n_m})}_{\boldsymbol{u}_m}.$$

Then $\sum_{i=1}^{m} \boldsymbol{u}_i = \mathbf{0}$ and so $\boldsymbol{u}_i = \mathbf{0}$ for each $i = 1, \ldots, m$ (since $V = \bigoplus_{i=1}^{m} U_i$). But then $a_1^1 = \cdots = a_m^{n_m} = 0$. Thus, $(\boldsymbol{u}_1^1, \ldots, \boldsymbol{u}_m^{n_m})$ is linear independent and spans V, i.e. it is a basis of V. □

3

LINEAR MAPS

"As You Should Verify"

Remark (p. 40). Given a basis $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ of V and any choice of vectors $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n \in W$, we can construct a linear map $\mathrm{T}\colon V \to W$ such that

$$\mathrm{T}(a_1\boldsymbol{v}_1 + \cdots + a_n\boldsymbol{v}_n) = a_1\boldsymbol{w}_1 + \cdots + a_n\boldsymbol{w}_n,$$

where $a_1, \ldots, a_n$ are arbitrary elements of $\mathbb{F}$. Then T is linear.

Proof. Let $\boldsymbol{u}, \boldsymbol{w} \in V$ with $\boldsymbol{u} = \sum_{i=1}^{n} a_i\boldsymbol{v}_i$ and $\boldsymbol{v} = \sum_{i=1}^{n} b_i\boldsymbol{v}_i$; let $a \in \mathbb{F}$. Then

$$\begin{aligned}\mathrm{T}(\boldsymbol{u}+\boldsymbol{v}) = \mathrm{T}\left(\sum_{i=1}^{n}(a_i+b_i)\boldsymbol{v}_i\right) &= \sum_{i=1}^{n}(a_i+b_i)\boldsymbol{w}_i \\ &= \sum_{i=1}^{n}a_i\boldsymbol{w}_i + \sum_{i=1}^{n}b_i\boldsymbol{w}_i \\ &= \mathrm{T}\boldsymbol{u} + \mathrm{T}\boldsymbol{v},\end{aligned}$$

and

$$\mathrm{T}(a\boldsymbol{u}) = \mathrm{T}\left(\sum_{i=1}^{n}(aa_i)\boldsymbol{v}_i\right) = \sum_{i=1}^{n}(aa_i)\boldsymbol{w}_i = a\left(\sum_{i=1}^{n}a_i\boldsymbol{w}_i\right) = a\mathrm{T}\boldsymbol{u}. \qquad \square$$

Remark (p. 40-41). Let $\mathrm{S}, \mathrm{T} \in \mathfrak{L}(V, W)$. Then $\mathrm{S}+\mathrm{T}, a\mathrm{T} \in \mathfrak{L}(V, W)$.

Proof. As for S + T, we have $(\mathrm{S}+\mathrm{T})(\boldsymbol{u}+\boldsymbol{v}) = \mathrm{S}(\boldsymbol{u}+\boldsymbol{v}) + \mathrm{T}(\boldsymbol{u}+\boldsymbol{v}) = \mathrm{S}\boldsymbol{u} + \mathrm{S}\boldsymbol{v} + \mathrm{T}\boldsymbol{u} + \mathrm{T}\boldsymbol{v} = (\mathrm{S}+\mathrm{T})(\boldsymbol{u}) + (\mathrm{S}+\mathrm{T})(\boldsymbol{v})$, and $(\mathrm{S}+\mathrm{T})(a\boldsymbol{v}) = \mathrm{S}(a\boldsymbol{v}) + \mathrm{T}(a\boldsymbol{v}) = a(\mathrm{S}+\mathrm{T})(\boldsymbol{v})$.

As for $a\mathrm{T}$, we have $(a\mathrm{T})(\boldsymbol{u}+\boldsymbol{v}) = a\left[\mathrm{T}(\boldsymbol{u}+\boldsymbol{v})\right] = a\left[\mathrm{T}\boldsymbol{u}+\mathrm{T}\boldsymbol{v}\right] = a\mathrm{T}\boldsymbol{u} + a\mathrm{T}\boldsymbol{v} = (a\mathrm{T})\boldsymbol{u} + (a\mathrm{T})\boldsymbol{v}$, and $(a\mathrm{T})(b\boldsymbol{v}) = a\left[\mathrm{T}(b\boldsymbol{v})\right] = ab\mathrm{T}\boldsymbol{v} = b(a\mathrm{T})\boldsymbol{v}$. $\square$

Exercises

► Exercise 33 (3.1). *Show that every linear map from a one-dimensional vector space to itself is multiplication by some scalar. More precisely, prove that if* $\dim V = 1$ *and* $\mathrm{T} \in \mathfrak{L}(V, V)$*, then there exists* $a \in \mathbb{F}$ *such that* $\mathrm{T}\boldsymbol{v} = a\boldsymbol{v}$ *for all* $\boldsymbol{v} \in V$.

Proof. Let $\boldsymbol{w} \in V$ be a basis of V. Then $\mathrm{T}\boldsymbol{w} = a\boldsymbol{w}$ for some $a \in \mathbb{F}$. For an arbitrary $\boldsymbol{v} \in V$, there exists $b \in \mathbb{F}$ such that $\boldsymbol{v} = b\boldsymbol{w}$. Then

$$\mathrm{T}\boldsymbol{v} = \mathrm{T}(b\boldsymbol{w}) = b(\mathrm{T}\boldsymbol{w}) = b(a\boldsymbol{w}) = a(b\boldsymbol{w}) = a\boldsymbol{v}. \qquad \square$$

► Exercise 34 (3.2). *Give an example of a function* $f: \mathbb{R}^2 \to \mathbb{R}$ *such that* $f(a\boldsymbol{v}) = af(\boldsymbol{v})$ *for all* $a \in \mathbb{R}$ *and all* $\boldsymbol{v} \in \mathbb{R}^2$ *but* f *is not linear.*

Proof. For any $\boldsymbol{v} = (v_1, v_2) \in \mathbb{R}^2$, let

$$f(v_1, v_2) = \begin{cases} v_1 & \text{if } v_1 = v_2 \\ 0 & \text{if } v_1 \neq v_2. \end{cases}$$

Now consider $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^2$ with $u_1 \neq u_2$, $v_1 \neq v_2$, but $u_1 + v_1 = u_2 + v_2 > 0$. Notice that

$$f(\boldsymbol{u} + \boldsymbol{v}) = u_1 + v_1 > 0 = f(\boldsymbol{u}) + f(\boldsymbol{v}).$$

Hence, f is not linear. $\square$

► Exercise 35 (3.3). *Suppose that* V *is finite dimensional. Prove that any linear map on a subspace of* V *can be extended to a linear map on* V*. In other words, show that if* U *is a subspace of* V *and* $\mathrm{S} \in \mathfrak{L}(U, W)$*, then there exists* $\mathrm{T} \in \mathfrak{L}(V, W)$ *such that* $\mathrm{T}\boldsymbol{u} = \mathrm{S}\boldsymbol{u}$ *for all* $\boldsymbol{u} \in U$.

Proof. Let $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_m)$ be a basis of U, and extend it to a basis of V:

$$(\boldsymbol{u}_1, \dots, \boldsymbol{u}_m, \boldsymbol{v}_1, \dots, \boldsymbol{v}_n).$$

Choose n vectors $\boldsymbol{w}_1, \dots, \boldsymbol{w}_n$ from W. Define a map $\mathrm{T}: V \to W$ by letting

$$\mathrm{T}\left(\sum_{i=1}^{m} a_i \boldsymbol{u}_i + \sum_{j=1}^{n} a_j \boldsymbol{v}_j\right) = \sum_{i=1}^{m} a_i \mathrm{S}\boldsymbol{u}_i + \sum_{j=1}^{n} a_j \boldsymbol{w}_j.$$

It is trivial to see that $\mathrm{S}\boldsymbol{u} = \mathrm{T}\boldsymbol{u}$ for all $\boldsymbol{u} \in U$. So we only show that T is a linear map. Let $\boldsymbol{u}, \boldsymbol{v} \in V$ with $\boldsymbol{u} = \sum_{i=1}^{m} a_i \boldsymbol{u}_i + \sum_{j=1}^{n} a_j \boldsymbol{v}_j$ and $\boldsymbol{v} = \sum_{i=1}^{m} b_i \boldsymbol{u}_i + \sum_{j=1}^{n} b_j \boldsymbol{v}_j$; let $a \in \mathbb{F}$. Then

$$
\begin{aligned}
\mathrm{T}(\boldsymbol{u}+\boldsymbol{v}) &= \mathrm{T}\left(\sum_{i=1}^{m}(a_i+b_i)\boldsymbol{u}_i + \sum_{j=1}^{n}\left(a_j+b_j\right)\boldsymbol{v}_j\right) \\
&= \sum_{i=1}^{m}(a_i+b_i)\mathrm{S}\boldsymbol{u}_i + \sum_{j=1}^{n}(a_j+b_j)\boldsymbol{w}_j \\
&= \left[\sum_{i=1}^{m}a_i\mathrm{S}\boldsymbol{u}_i + \sum_{j=1}^{n}a_j\boldsymbol{w}_j\right] + \left[\sum_{i=1}^{m}b_i\mathrm{S}\boldsymbol{u}_i + \sum_{j=1}^{n}b_j\boldsymbol{w}_j\right] \\
&= \mathrm{T}\boldsymbol{u} + \mathrm{T}\boldsymbol{v},
\end{aligned}
$$

and

$$
\begin{aligned}
\mathrm{T}a\boldsymbol{u} = \mathrm{T}\left(a\left(\sum_{i=1}^{m}a_i\boldsymbol{u}_i + \sum_{j=1}^{n}a_j\boldsymbol{v}_j\right)\right) &= \mathrm{T}\left(\sum_{i=1}^{m}aa_i\boldsymbol{u}_i + \sum_{j=1}^{n}aa_j\boldsymbol{v}_j\right) \\
&= \sum_{i=1}^{m}aa_i\mathrm{S}\boldsymbol{u}_i + \sum_{j=1}^{n}aa_j\boldsymbol{w}_j \\
&= a\left[\sum_{i=1}^{m}a_i\mathrm{S}\boldsymbol{u}_i + \sum_{j=1}^{n}a_j\boldsymbol{w}_j\right] \\
&= a\mathrm{T}\boldsymbol{u}.
\end{aligned}
$$
□

► EXERCISE 36 (3.4). *Suppose that* T *is a linear map from* V *to* $\mathbb{F}$. *Prove that if* $\boldsymbol{u} \in V$ *is not in* $\mathcal{N}_{\mathrm{T}}$, *then*

$$V = \mathcal{N}_{\mathrm{T}} \oplus \{a\boldsymbol{u} : a \in \mathbb{F}\}.$$

PROOF. Let $\mathrm{T} \in \mathfrak{L}(V, \mathbb{F})$. Since $\boldsymbol{u} \in V \smallsetminus \mathcal{N}_{\mathrm{T}}$, we get $\boldsymbol{u} \neq \boldsymbol{0}$ and $\mathrm{T}\boldsymbol{u} \neq 0$. Thus, $\dim \mathcal{R}_{\mathrm{T}} \geqslant 1$. Since $\dim \mathcal{R}_{\mathrm{T}} \leqslant \dim \mathbb{F} = 1$, we get $\dim \mathcal{R}_{\mathrm{T}} = 1$. It follows from Theorem 3.4 that

$$\dim V = \dim \mathcal{N}_{\mathrm{T}} + 1 = \dim \mathcal{N}_{\mathrm{T}} + \dim \{a\boldsymbol{u} : a \in \mathbb{F}\}. \tag{3.1}$$

Let $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)$ be a basis of $\mathcal{N}_{\mathrm{T}}$. Then $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m, \boldsymbol{u})$ is linear independent since $\boldsymbol{u} \notin \operatorname{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m) = \mathcal{N}_{\mathrm{T}}$. It follows from (3.1) that $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m, \boldsymbol{u})$ is a basis of V (by Proposition 2.17). Therefore

$$
\begin{aligned}
V = \operatorname{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m, \boldsymbol{u}) &= \left\{\sum_{i=1}^{m}a_i\boldsymbol{v}_i + a\boldsymbol{u} : a_1, \ldots, a_m, a \in \mathbb{F}\right\} \\
&= \left\{\sum_{i=1}^{m}a_i\boldsymbol{v}_i : a_1, \ldots, a_m \in \mathbb{F}\right\} + \{a\boldsymbol{u} : a \in \mathbb{F}\} \\
&= \mathcal{N}_{\mathrm{T}} + \{a\boldsymbol{u} : a \in \mathbb{F}\}.
\end{aligned} \tag{3.2}
$$

It follows from (3.1) and (3.2) that $V = \mathcal{N}_{\mathrm{T}} \oplus \{a\boldsymbol{u} : a \in \mathbb{F}\}$ by Proposition 2.19. □

► EXERCISE 37 (3.5). *Suppose that* $\mathrm{T} \in \mathfrak{L}(V, W)$ *is injective and* $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ *is linearly independent in* V*. Prove that* $(\mathrm{T}\boldsymbol{v}_1, \ldots, \mathrm{T}\boldsymbol{v}_n)$ *is linearly independent in* W*.*

PROOF. Let

$$\mathbf{0} = \sum_{i=1}^{n} a_i \cdot \mathrm{T}\boldsymbol{v}_i = \mathrm{T}\left(\sum_{i=1}^{n} a_i \boldsymbol{v}_i\right).$$

Then $\sum_{i=1}^{n} a_i \boldsymbol{v}_i = \mathbf{0}$ since $\mathcal{N}_{\mathrm{T}} = \{\mathbf{0}\}$. The linear independence of $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ implies that $a_1 = \cdots = a_n = 0$. □

► EXERCISE 38 (3.6). *Prove that if* $\mathrm{S}_1, \ldots, \mathrm{S}_n$ *are injective linear maps such that* $\mathrm{S}_1 \cdots \mathrm{S}_n$ *makes sense, then* $\mathrm{S}_1 \cdots \mathrm{S}_n$ *is injective.*

PROOF. We use mathematical induction to prove this claim. It holds for $n = 1$ trivially. Let us suppose that $\mathrm{S}_1 \cdots \mathrm{S}_n$ is injective if $\mathrm{S}_1, \ldots, \mathrm{S}_n$ are. Now assume that $\mathrm{S}_1, \ldots, \mathrm{S}_{n+1}$ are all injective linear maps. Let $\mathrm{T} = \mathrm{S}_1 \cdots \mathrm{S}_{n+1}$. For every $\boldsymbol{v} \in \mathcal{N}_{\mathrm{T}}$ we have

$$\mathbf{0} = \mathrm{T}\boldsymbol{v} = (\mathrm{S}_1 \cdots \mathrm{S}_n)(\mathrm{S}_{n+1}\boldsymbol{v}).$$

But the above display implies that $\mathrm{S}_{n+1}\boldsymbol{v} = \mathbf{0}$ since $(\mathrm{S}_1 \cdots \mathrm{S}_n)$ is injective by the induction hypothesis, which implies further that $\boldsymbol{v} = \mathbf{0}$ since S_{n+1} is injective. This proves that $\mathcal{N}_{\mathrm{T}} = \{\mathbf{0}\}$ and so T is injective. □

► EXERCISE 39 (3.7). *Prove that if* $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ *spans* V *and* $\mathrm{T} \in \mathfrak{L}(V, W)$ *is surjective, then* $(\mathrm{T}\boldsymbol{v}_1, \ldots, \mathrm{T}\boldsymbol{v}_n)$ *spans* W*.*

PROOF. Since T is surjective, for any $\boldsymbol{w} \in W$, there exists $\boldsymbol{v} \in V$ such that $\mathrm{T}\boldsymbol{v} = \boldsymbol{w}$; since $V = \operatorname{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$, there exists $(a_1, \ldots, a_n) \in \mathbb{F}^n$ such that $\boldsymbol{v} = \sum_{i=1}^{n} a_i \boldsymbol{v}_i$. Hence,

$$\boldsymbol{w} = \mathrm{T}\left(\sum_{i=1}^{n} a_i \boldsymbol{v}_i\right) = \sum_{i=1}^{n} a_i \mathrm{T}\boldsymbol{v}_i,$$

that is, $W = \operatorname{span}(\mathrm{T}\boldsymbol{v}_1, \ldots, \mathrm{T}\boldsymbol{v}_n)$. □

► EXERCISE 40 (3.8). *Suppose that* V *is finite dimensional and that* $\mathrm{T} \in \mathfrak{L}(V, W)$*. Prove that there exists a subspace* U *of* V *such that* $U \cap \mathcal{N}_{\mathrm{T}} = \{\mathbf{0}\}$ *and* $\mathcal{R}_{\mathrm{T}} = \{\mathrm{T}\boldsymbol{u} : \boldsymbol{u} \in U\}$*.*

PROOF. Let $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m)$ be a basis of $\mathcal{N}_{\mathrm{T}}$, which can be extended to a basis $(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_m, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ of V. Let $U = \operatorname{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$. Then $U \cap \mathcal{N}_{\mathrm{T}} = \{\mathbf{0}\}$ (see the proof of Proposition 2.13).

To see $\mathcal{R}_{\mathrm{T}} = \{\mathrm{T}\boldsymbol{u} : \boldsymbol{u} \in U\}$, take an arbitrary $\boldsymbol{v} \in V$. Then

$$\mathrm{T}\boldsymbol{v} = \mathrm{T}\left(\sum_{i=1}^{m} a_i \boldsymbol{u}_i + \sum_{j=1}^{n} a_j \boldsymbol{v}_j\right) = \mathrm{T}\left(\sum_{j=1}^{n} a_j \boldsymbol{v}_j\right) = \mathrm{T}\boldsymbol{u}$$

for some $\boldsymbol{u} = \sum_{j=1}^{n} a_j \boldsymbol{v}_j \in U$. □

► EXERCISE 41 (3.9). *Prove that if* T *is a linear map from* $\mathbb{F}^4$ *to* $\mathbb{F}^2$ *such that*

$$\mathcal{N}_{\mathrm{T}} = \left\{(x_1, x_2, x_3, x_4) \in \mathbb{F}^4 : x_1 = 5x_2 \text{ and } x_3 = 7x_4\right\},$$

then T *is surjective.*

PROOF. Let

$$\boldsymbol{v}_1 = \begin{pmatrix} 5 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 7 \\ 1 \end{pmatrix}.$$

It is easy to see that $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ is a basis of $\mathcal{N}_{\mathrm{T}}$; that is, $\dim \mathcal{N}_{\mathrm{T}} = 2$. Then

$$\dim \mathcal{R}_{\mathrm{T}} = \dim \mathbb{F}^4 - \dim \mathcal{N}_{\mathrm{T}} = 4 - 2 = 2 = \dim \mathbb{F}^2,$$

and so T is surjective. □

► EXERCISE 42 (3.10). *Prove that there does not exist a linear map from* $\mathbb{F}^5$ *to* $\mathbb{F}^2$ *whose null space equals*

$$\left\{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{F}^5 : x_1 = 3x_2 \text{ and } x_3 = x_4 = x_5\right\}.$$

PROOF. It is easy to see that the following two vectors consist of a basis of $\mathcal{N}_{\mathrm{T}}$ if $\mathrm{T} \in \mathcal{L}(\mathbb{F}^5, \mathbb{F}^2)$:

$$\boldsymbol{v}_1 = \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}.$$

Then, $\dim \mathcal{N}_{\mathrm{T}} = 2$ and so $\dim \mathcal{R}_{\mathrm{T}} = 5 - 2 = 3 > \dim \mathbb{F}^2$, which is impossible. □

► EXERCISE 43 (3.11). *Prove that if there exists a linear map on V whose null space and range are both finite dimensional, then V is finite dimensional.*

PROOF. If $\dim \mathcal{N}_{\mathrm{T}} < \infty$ and $\dim \mathcal{R}_{\mathrm{T}} < \infty$, then $\dim V = \dim \mathcal{N}_{\mathrm{T}} + \dim \mathcal{R}_{\mathrm{T}} < \infty$. □

► EXERCISE 44 (3.12). *Suppose that V and W are both finite dimensional. Prove that there exists a surjective linear map from V onto W if and only if* $\dim W \leqslant \dim V$.

PROOF. If there exists a surjective linear map $\mathrm{T} \in \mathcal{L}(V, W)$, then $\dim W = \dim \mathcal{R}_{\mathrm{T}} = \dim V - \dim \mathcal{N}_{\mathrm{T}} \geqslant \dim V$.

Now let $\dim W \leqslant \dim V$. Let $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ be a basis of V, and let $(\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m)$ be a basis of W, with $m \leqslant n$. Define $\mathrm{T} \in \mathcal{L}(V, W)$ by letting

$$\mathrm{T}\left(\sum_{i=1}^{m} a_i \boldsymbol{v}_i + \sum_{j=m+1}^{n} a_j \boldsymbol{v}_j\right) = \sum_{i=1}^{m} a_i \boldsymbol{w}_i .$$

Then for every $\boldsymbol{w} = \sum_{i=1}^{m} a_i \boldsymbol{w}_i \in W$, there exists $\boldsymbol{v} = \sum_{j=1}^{n} a_j \boldsymbol{v}_j$ such that $\mathrm{T}\boldsymbol{v} = \boldsymbol{w}$, i.e. T is surjective. □

► EXERCISE 45 (3.13). *Suppose that V and W are finite dimensional and that U is a subspace of V. Prove that there exists $\mathrm{T} \in \mathfrak{L}(V, W)$ such that $\mathcal{N}_\mathrm{T} = U$ if and only if* $\dim U \geqslant \dim V - \dim W$.

PROOF. For every $\mathrm{T} \in \mathfrak{L}(V, W)$, if $\mathcal{N}_\mathrm{T} = U$, then $\dim U = \dim V - \dim \mathcal{R}_\mathrm{T} \geqslant \dim V - \dim W$.

Now let $\dim U \geqslant \dim V - \dim W$. Let $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_m)$ be a basis of U, which can be extended to a basis $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_m, \boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ of V. Let $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_p)$ be a basis of W. Then $m \geqslant (m+n) - p$ implies that $n \leqslant p$. Define $\mathrm{T} \in \mathfrak{L}(V, W)$ by letting

$$\mathrm{T}\left(\sum_{i=1}^{m} a_i \boldsymbol{u}_i + \sum_{j=1}^{n} a_j \boldsymbol{v}_j\right) = \sum_{j=1}^{n} a_j \boldsymbol{w}_j .$$

Then $\mathcal{N}_\mathrm{T} = U$. □

► EXERCISE 46 (3.14). *Suppose that W is finite dimensional and $\mathrm{T} \in \mathfrak{L}(V, W)$. Prove that* T *is injective if and only if there exists* $\mathrm{S} \in \mathfrak{L}(W, V)$ *such that* ST *is the identity map on V.*

PROOF. Suppose first that $\mathrm{ST} = \mathrm{Id}_V$. Then for any $\boldsymbol{u}, \boldsymbol{v} \in V$ with $\boldsymbol{u} \neq \boldsymbol{v}$, we have $\boldsymbol{u} = (\mathrm{ST})\boldsymbol{u} \neq (\mathrm{ST})\boldsymbol{v} = \boldsymbol{v}$; that is, $\mathrm{S}(\mathrm{T}\boldsymbol{u}) \neq \mathrm{S}(\mathrm{T}\boldsymbol{v})$, and so $\mathrm{T}\boldsymbol{u} \neq \mathrm{T}\boldsymbol{v}$.

For the inverse direction, let T be injective. Then $\dim V \leqslant \dim W$ by Corollary 3.5. Also, $\dim W < +\infty$. Let $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ be a basis of V. It follows from Exercise 37 that $(\mathrm{T}\boldsymbol{v}_1, \dots, \mathrm{T}\boldsymbol{v}_n)$ is linearly independent, and so can be extended to a basis $(\mathrm{T}\boldsymbol{v}_1, \dots, \mathrm{T}\boldsymbol{v}_n, \boldsymbol{w}_1, \dots, \boldsymbol{w}_m)$ of W. Define $\mathrm{S} \in \mathfrak{L}(W, V)$ by letting

$$\mathrm{S}(\mathrm{T}\boldsymbol{v}_i) = (\mathrm{ST})\,\boldsymbol{v}_i = \boldsymbol{v}_i, \quad \text{and} \quad \mathrm{S}(\boldsymbol{w}_i) = \boldsymbol{0}_V .$$ □

► EXERCISE 47 (3.15). *Suppose that V is finite dimensional and $\mathrm{T} \in \mathfrak{L}(V, W)$. Prove that* T *is surjective if and only if there exists* $\mathrm{S} \in \mathfrak{L}(W, V)$ *such that* TS *is the identity map on W.*

PROOF. If $\mathrm{TS} = \mathrm{Id}_W$, then for any $\boldsymbol{w} \in W$, we have $\mathrm{T}(\mathrm{S}\boldsymbol{w}) = \mathrm{Id}_W(\boldsymbol{w}) = \boldsymbol{w}$, that is, there exists $\mathrm{S}\boldsymbol{w} \in V$ such that $\mathrm{T}(\mathrm{S}\boldsymbol{w}) = \boldsymbol{w}$, and so T is surjective.

If T is surjective, then $\dim W \leqslant \dim V$. Let $(\boldsymbol{w}_1, \dots, \boldsymbol{w}_m)$ be a basis of W, and let $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ be a basis of V with $n \geqslant m$. Define $\mathrm{S} \in \mathfrak{L}(W, V)$ by letting

$$\mathrm{S}\boldsymbol{w}_i = \boldsymbol{v}_i, \quad \text{with} \quad \mathrm{T}\boldsymbol{v}_i = \boldsymbol{w}_i .$$ □

► EXERCISE 48 (3.16[1]). *Suppose that U and V are finite-dimensional vector spaces and that* $\mathrm{S} \in \mathfrak{L}(V, W)$, $\mathrm{T} \in \mathfrak{L}(U, V)$. *Prove that*

$$\dim \mathcal{N}_{\mathrm{ST}} \leqslant \dim \mathcal{N}_{\mathrm{S}} + \dim \mathcal{N}_{\mathrm{T}}.$$

PROOF. We have $W \xleftarrow{\mathrm{S}} V \xleftarrow{\mathrm{T}} U$. Since

$$\mathcal{R}_{\mathrm{ST}} = (\mathrm{ST})[U] = \mathrm{S}\big[\mathrm{T}[U]\big] = \mathrm{S}[\mathcal{R}_{\mathrm{T}}],$$

we have

$$\dim \mathcal{R}_{\mathrm{ST}} = \dim \mathrm{S}[\mathcal{R}_{\mathrm{T}}].$$

Let N be the complement of $\mathcal{R}_{\mathrm{T}}$ so that $V = \mathcal{R}_{\mathrm{T}} \oplus N$; then

$$\dim V = \dim \mathcal{R}_{\mathrm{T}} + \dim N, \tag{3.3}$$

and

$$\mathcal{R}_{\mathrm{S}} = \mathrm{S}[V] = \mathrm{S}[\mathcal{R}_{\mathrm{T}}] + \mathrm{S}[N].$$

It follows from Theorem 2.18 that

$$\begin{aligned}\dim \mathcal{R}_{\mathrm{S}} &= \dim \mathrm{S}[\mathcal{R}_{\mathrm{T}}] + \dim \mathrm{S}[N] - \dim \mathrm{S}[\mathcal{R}_{\mathrm{T}}] \cap \mathrm{S}[N] \\ &\leqslant \dim \mathrm{S}[\mathcal{R}_{\mathrm{T}}] + \dim \mathrm{S}[N] \\ &\leqslant \dim \mathrm{S}[\mathcal{R}_{\mathrm{T}}] + \dim N \\ &= \dim \mathcal{R}_{\mathrm{ST}} + \dim N,\end{aligned}$$

and hence that

$$\begin{aligned}\dim V - \dim \mathcal{N}_{\mathrm{S}} &= \dim \mathcal{R}_{\mathrm{S}} \\ &\leqslant \dim \mathcal{R}_{\mathrm{ST}} + \dim N \\ &= \dim \mathcal{R}_{\mathrm{ST}} + \dim V - \dim \mathcal{R}_{\mathrm{T}},\end{aligned} \tag{3.4}$$

where the last equality is from (3.3). Hence, (3.4) becomes

$$\dim \mathcal{R}_{\mathrm{T}} - \dim \mathcal{N}_{\mathrm{S}} \leqslant \dim \mathcal{R}_{\mathrm{ST}},$$

or equivalently,

$$\dim U - \dim \mathcal{N}_{\mathrm{T}} - \dim \mathcal{N}_{\mathrm{S}} \leqslant \dim U - \dim \mathcal{N}_{\mathrm{ST}};$$

that is,

$$\dim \mathcal{N}_{\mathrm{ST}} \leqslant \dim \mathcal{N}_{\mathrm{S}} + \dim \mathcal{N}_{\mathrm{T}}. \qquad \square$$

► EXERCISE 49 (3.17). *Prove that the distributive property holds for matrix addition and matrix multiplication.*

PROOF. Let $\mathbf{A} = [a_{ij}] \in \mathrm{Mat}(m, n, \mathbb{F})$, $\mathbf{B} = [b_{ij}] \in \mathrm{Mat}\big(n, p, \mathbb{F}\big)$, and $\mathbf{C} = [c_{ij}] \in \mathrm{Mat}\big(n, p, \mathbb{F}\big)$. Then $\mathbf{B} + \mathbf{C} = [b_{ij} + c_{ij}] \in \mathrm{Mat}\big(n, p, \mathbb{F}\big)$. It is evident that $\mathbf{AB}$ and

[1] See Halmos (1995, Problem 95, p.270).

AC are $m \times p$ matrices. Further,

$$\begin{aligned}\mathbf{A}(\mathbf{B}+\mathbf{C}) &= \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11}+c_{11} & \cdots & b_{1p}+c_{1p} \\ \vdots & \ddots & \vdots \\ b_{n1}+c_{n1} & \cdots & b_{np}+c_{np} \end{pmatrix} \\ &= \begin{pmatrix} \sum_{i=1}^{n} a_{1i}b_{i1} + \sum_{i=1}^{n} a_{1i}c_{i1} & \cdots & \sum_{i=1}^{n} a_{1i}b_{ip} + \sum_{i=1}^{n} a_{1i}c_{ip} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{n} a_{mi}b_{i1} + \sum_{i=1}^{n} a_{mi}c_{i1} & \cdots & \sum_{i=1}^{n} a_{mi}b_{ip} + \sum_{i=1}^{n} a_{mi}c_{ip} \end{pmatrix} \\ &= \mathbf{AB}+\mathbf{AC}. \end{aligned}$$

□

► EXERCISE 50 (3.18). *Prove that matrix multiplication is associative.*

PROOF. Similar to Exercise 49. □

► EXERCISE 51 (3.19). *Suppose* $\mathrm{T} \in \mathfrak{L}(\mathbb{F}^n, \mathbb{F}^m)$ *and that*

$$\mathcal{M}(\mathrm{T}) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix},$$

where we are using the standard bases. Prove that

$$\mathrm{T}(x_1, \ldots, x_n) = \left(\sum_{i=1}^{n} a_{1i}x_i, \ldots, \sum_{i=1}^{n} a_{mi}x_i \right)$$

for every $(x_1, \ldots, x_n) \in \mathbb{F}^n$*.*

PROOF. We need to prove that $\mathrm{T}\boldsymbol{x} = \mathcal{M}(\mathrm{T}) \cdot \boldsymbol{x}$ for any $\boldsymbol{x} \in \mathbb{F}^n$. Let $(\boldsymbol{e}_1^n, \ldots, \boldsymbol{e}_n^n)$ be the standard basis for $\mathbb{F}^n$, and let $(\boldsymbol{e}_1^m, \ldots, \boldsymbol{e}_m^m)$ be the standard basis for $\mathbb{F}^m$. Then

$$\begin{aligned}\mathrm{T}(x_1, \ldots, x_n) = \mathrm{T}\left(\sum_{i=1}^{n} x_i \boldsymbol{e}_i^n \right) = \sum_{i=1}^{n} x_i \mathrm{T}\boldsymbol{e}_i^n &= \sum_{i=1}^{n} x_i \sum_{j=1}^{m} a_{ji} \boldsymbol{e}_j^m \\ &= \left(\sum_{i=1}^{n} a_{1i}x_i, \ldots, \sum_{i=1}^{n} a_{mi}x_i \right). \end{aligned}$$ □

► EXERCISE 52 (3.20). *Suppose* $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ *is a basis of* V*. Prove that the function* $\mathrm{T}\colon V \to \mathrm{Mat}(n, 1, \mathbb{F})$ *defined by* $\mathrm{T}\boldsymbol{v} = \mathcal{M}(\boldsymbol{v})$ *is an invertible linear map of* V *onto* $\mathrm{Mat}(n, 1, \mathbb{F})$*; here* $\mathcal{M}(\boldsymbol{v})$ *is the matrix of* $\boldsymbol{v} \in V$ *with respect to the basis* $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$*.*

PROOF. For every $\boldsymbol{v} = \sum_{i=1}^{n} b_i \boldsymbol{v}_i \in V$, we have

$$\mathcal{M}(\boldsymbol{v}) = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}.$$

Since $a\boldsymbol{v} = \sum_{i=1}^{n} (ab_i)\, \boldsymbol{v}_i$ for any $a \in \mathbb{F}$, we have $\mathcal{M}(a\boldsymbol{v}) = a\mathcal{M}(\boldsymbol{v})$. Further, for any $\boldsymbol{u} = \sum_{i=1}^{n} a_i \boldsymbol{v}_i \in V$, and any $\boldsymbol{v} = \sum_{i=1}^{n} b_i \boldsymbol{v}_i \in V$, we have $\boldsymbol{u} + \boldsymbol{v} = \sum_{i=1}^{n} (a_i + b_i)\boldsymbol{v}_i$; hence, $\mathcal{M}(\boldsymbol{u} + \boldsymbol{v}) = \mathcal{M}(\boldsymbol{u}) + \mathcal{M}(\boldsymbol{v})$. Therefore, T is a liner map.

We now show that T is invertible by proving T is bijective. (i) If $\mathrm{T}\boldsymbol{v} = (0, \dots, 0)'$, then $\boldsymbol{v} = \sum_{i=1}^{n} 0\boldsymbol{v}_i = \mathbf{0}_V$; that is, $\mathcal{N}_{\mathrm{T}} = \{\mathbf{0}_T\}$. Hence, T is injective. (ii) Take any $M = (a_1, \dots, a_n)' \in \mathrm{Mat}\,(n, 1, \mathbb{F})$. Let $\boldsymbol{v} = \sum_{i=1}^{n} a_i \boldsymbol{v}_i$. Then $\mathrm{T}\boldsymbol{v} = M$; that is, T is surjective. □

► EXERCISE 53 (3.21). *Prove that every linear map from* $\mathrm{Mat}\,(n, 1, \mathbb{F})$ *to* $\mathrm{Mat}\,(m, 1, \mathbb{F})$ *is given by a matrix multiplication. In other words, prove that if*

$$\mathrm{T} \in \mathfrak{L}(\mathrm{Mat}\,(n, 1, \mathbb{F})\,, \mathrm{Mat}\,(m, 1, \mathbb{F})),$$

then there exists an $m \times n$ *matrix* $\mathbf{A}$ *such that* $\mathrm{T}\mathbf{B} = \mathbf{A}\mathbf{B}$ *for every* $\mathbf{B} \in \mathrm{Mat}\,(n, 1, \mathbb{F})$*.*

PROOF. A basis of $\mathrm{Mat}\,(m, n, \mathbb{F})$ consists of those $m \times n$ matrices that have 0 in all entries except for a 1 in one entry. Therefore, a basis for $\mathrm{Mat}\,(n, 1, \mathbb{F})$ consists of the standard basis of $\mathbb{F}^n$, $\left(\boldsymbol{e}_1^n, \dots, \boldsymbol{e}_n^n\right)$, where, for example, $\boldsymbol{e}_1^n = (1, 0, \dots, 0)'$. For any $\mathrm{T} \in \mathfrak{L}(\mathrm{Mat}\,(n, 1, \mathbb{F})\,, \mathrm{Mat}\,(m, 1, \mathbb{F}))$, let

$$\underset{(m \times n)}{\mathbf{A}} := \begin{pmatrix} \mathrm{T}\boldsymbol{e}_1^n & \cdots & \mathrm{T}\boldsymbol{e}_n^n \end{pmatrix}.$$

Then for any $\mathbf{B} = \sum_{i=1}^{n} a_i \boldsymbol{e}_i^n \in \mathrm{Mat}\,(n, 1, \mathbb{F})$, we have

$$\mathrm{T}\mathbf{B} = \mathrm{T}\left(\sum_{i=1}^{n} a_i \boldsymbol{e}_i^n\right) = \sum_{i=1}^{n} a_i \mathrm{T}\boldsymbol{e}_i^n = \mathbf{A}\mathbf{B}. \qquad \square$$

► EXERCISE 54 (3.22). *Suppose that* V *is finite dimensional and* $\mathrm{S}, \mathrm{T} \in \mathfrak{L}(V)$*. Prove that* ST *is invertible if and only if both* S *and* T *are invertible.*

PROOF. First assume that both S and T are invertible. Then $(\mathrm{ST})\left(\mathrm{T}^{-1}\mathrm{S}^{-1}\right) = \mathrm{S}\mathrm{Id}\mathrm{S}^{-1} = \mathrm{Id}$ and $\left(\mathrm{T}^{-1}\mathrm{S}^{-1}\right)(\mathrm{ST}) = \mathrm{Id}$. Hence, ST is invertible and $(\mathrm{ST})^{-1} = \mathrm{T}^{-1}\mathrm{S}^{-1}$.

Now suppose that ST is invertible, so it is injective. Take any $\boldsymbol{u}, \boldsymbol{v} \in V$ with $\boldsymbol{u} \neq \boldsymbol{v}$; then $(\mathrm{ST})\,\boldsymbol{u} \neq (\mathrm{ST})\,\boldsymbol{v}$; that is,

$$\boldsymbol{u} \neq \boldsymbol{v} \Longrightarrow \mathrm{S}(\mathrm{T}\boldsymbol{u}) \neq \mathrm{S}(\mathrm{T}\boldsymbol{v}). \tag{3.5}$$

But then $\mathrm{T}\boldsymbol{u} \neq \mathrm{T}\boldsymbol{v}$, which implies that T is invertible by Theorem 3.21. Finally, for any $\boldsymbol{u}, \boldsymbol{v} \in V$ with $\boldsymbol{u} \neq \boldsymbol{v}$, there exist $\boldsymbol{u}', \boldsymbol{v}' \in V$ with $\boldsymbol{u}' \neq \boldsymbol{v}'$ such that $\boldsymbol{u} = \mathrm{T}\boldsymbol{u}'$ and $\boldsymbol{v} = \mathrm{T}\boldsymbol{v}'$. Hence, by (3.5), $\boldsymbol{u} \neq \boldsymbol{v}$ implies that

$$\mathrm{S}\boldsymbol{u} = \mathrm{S}(\mathrm{T}\boldsymbol{u}') \neq \mathrm{S}(\mathrm{T}\boldsymbol{v}') = \mathrm{S}\boldsymbol{v};$$

that is, S is injective, too. Applying Theorem 3.21 once again, we know that S is invertible. □

► EXERCISE 55 (3.23). *Suppose that V is finite dimensional and* $\mathrm{S}, \mathrm{T} \in \mathfrak{L}(V)$. *Prove that* $\mathrm{ST} = \mathrm{Id}$ *if and only if* $\mathrm{TS} = \mathrm{Id}$.

PROOF. We only prove the *only if* part; the *if* part can be proved similarly. If $\mathrm{ST} = \mathrm{Id}$, then ST is bijective and so invertible. Then by Exercise 54, both S and T are invertible. Therefore,

$$\mathrm{ST} = \mathrm{Id} \iff \mathrm{S}^{-1}\mathrm{ST} = \mathrm{S}^{-1}\mathrm{Id} \iff \mathrm{T} = \mathrm{S}^{-1} \iff \mathrm{TS} = \mathrm{S}^{-1}\mathrm{S} = \mathrm{Id}.$$ □

► EXERCISE 56 (3.24). *Suppose that V is finite dimensional and* $\mathrm{T} \in \mathfrak{L}(V)$. *Prove that* T *is a scalar multiple of the identity if and only if* $\mathrm{ST} = \mathrm{TS}$ *for every* $\mathrm{S} \in \mathfrak{L}(V)$.

PROOF. If $\mathrm{T} = a\mathrm{Id}$ for some $a \in \mathbb{F}$, then for any $\mathrm{S} \in \mathfrak{L}(V)$, we have

$$\mathrm{ST} = a\mathrm{SId} = a\mathrm{S} = a\mathrm{IdS} = \mathrm{TS}.$$

For the converse direction, assume that $\mathrm{ST} = \mathrm{TS}$ for all $\mathrm{S} \in \mathfrak{L}(V)$. □

► EXERCISE 57 (3.25). *Prove that if V is finite dimensional with* $\dim V > 1$, *then the set of noninvertible operators on V is not a subspace of* $\mathfrak{L}(V)$.

PROOF. Since every finite-dimensional vector space is isomorphic to some $\mathbb{F}^n$, we just focus on $\mathbb{F}^n$. For simplicity, consider $\mathbb{F}^2$. Let $\mathrm{S}, \mathrm{T} \in \mathbb{F}^2$ with

$$\mathrm{S}(a, b) = (a, 0) \quad \text{and} \quad \mathrm{T}(a, b) = (0, b).$$

Obviously, both S and T are noninvertible since they are not injective; however, $\mathrm{S} + \mathrm{T} = \mathrm{Id}$ is invertible. □

► EXERCISE 58 (3.26). *Suppose n is a positive integer and $a_{ij} \in \mathbb{F}$ for $i, j = 1, \dots, n$. Prove that the following are equivalent:*

a. *The trivial solution $x_1 = \cdots = x_n = 0$ is the only solution to the homogeneous system of equations*

$$\sum_{k=1}^{n} a_{1k} x_k = 0$$
$$\vdots$$
$$\sum_{k=1}^{n} a_{nk} x_k = 0.$$

b. *For every $c_1, \dots, c_n \in \mathbb{F}$, there exists a solution to the system of equations*

$$\sum_{k=1}^{n} a_{1k} x_k = c_1$$
$$\vdots$$
$$\sum_{k=1}^{n} a_{nk} x_k = c_n.$$

PROOF. Let

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

If we let $\mathrm{T}\boldsymbol{x} = A\boldsymbol{x}$, then by Exercise 52, $\mathrm{T} \in \mathfrak{L}(\mathbb{F}^n, \mathbb{F}^n)$. (a) implies that $\mathcal{N}_\mathrm{T} = \{\mathbf{0}\}$; hence

$$\dim \mathcal{R}_\mathrm{T} = n - 0 = n.$$

Since $\mathcal{R}_\mathrm{T}$ is a subspace of $\mathbb{F}^n$, we have $\mathcal{R}_\mathrm{T} = \mathbb{F}^n$, that is, T is surjective: for any $(c_1, \ldots, c_n)$, there is a unique solution $(x_1, \ldots, x_n)$. □

4

POLYNOMIALS

► EXERCISE 59 (4.1). *Suppose m and n are positive integers with $m \leqslant n$. Prove that there exists a polynomial $p \in \mathfrak{P}_n(\mathbb{F})$ with exactly m distinct roots.*

PROOF. Let
$$p(z) = \prod_{i=1}^{m} (z - \lambda_i)^{m_i},$$
where $\lambda_1, \dots, \lambda_m \in \mathbb{F}$ are distinct and $\sum_{i=1}^{m} m_i = n$. □

► EXERCISE 60 (4.2). *Suppose that $z_1, \dots, z_{m+1}$ are distinct elements of $\mathbb{F}$ and that $w_1, \dots, w_{m+1} \in \mathbb{F}$. Prove that there exists a unique polynomial $p \in \mathfrak{P}_m(\mathbb{F})$ such that $p(z_j) = w_j$ for $j = 1, \dots, m+1$.*

PROOF. Let $p_i(x) = \prod_{j \neq i} (x - z_j)$. Then $\deg p_i = m$ and $p_i(z_j) \neq 0$ if and only if $i = j$. Define
$$p(x) = \sum_{i=1}^{m+1} \frac{w_i}{p_i(z_i)} p_i(x).$$
Then $\deg p = m$ and
$$\begin{aligned} p(z_j) &= \frac{w_1}{p_1(z_1)} p_1(z_j) + \cdots + \frac{w_j}{p_j(z_j)} p_j(z_j) + \cdots + \frac{w_{m+1}}{p_{m+1}(z_{m+1})} p_{m+1}(z_j) \\ &= w_j. \end{aligned}$$
□

► EXERCISE 61 (4.3). *Prove that if $p, q \in \mathfrak{P}(\mathbb{F})$, with $p \neq 0$, then there exist unique polynomials $s, r \in \mathfrak{P}(\mathbb{F})$ such that $q = sp + r$ and $\deg r < \deg p$.*

PROOF. Assume that there also exist $s', r' \in \mathfrak{P}(\mathbb{F})$ such that $q = s'p + r'$ and $\deg r' < \deg p$. Then
$$(s - s')p + (r - r') = 0.$$

If $s \neq s'$, then $\deg (s - s')p + \deg (r - r') = \deg (s - s') + \deg p + \deg (r - r') \geqslant 0$; but $\deg 0 = -\infty$. Hence, $s = s'$ and so $r = r'$. □

► EXERCISE 62 (4.4). *Suppose $p \in \mathfrak{P}(\mathbb{C})$ has degree m. Prove that p has m distinct roots if and only if p and its derivative p' have no roots in common.*

PROOF. If λ is a root of p, then we can write p as $p(z) = (z - \lambda)q(z)$. Then

$$p'(z) = q(z) + (z - \lambda)q'(z).$$

So λ is also a root for p' if and only if λ is a root of q; that is, λ is a multiple root. A contradiction. □

► EXERCISE 63 (4.5). *Prove that every polynomial with odd degree and real coefficients has a real root.*

PROOF. If $p \in \mathfrak{P}(\mathbb{R})$ with $\deg p$ is odd, then $p(-\infty) < 0$ and $p(+\infty) > 0$. Then there exists $x^* \in \mathbb{R}$ such that $p(x^*) = 0$. □

5

EIGENVALUES AND EIGENVECTORS

"As You Should Verify"

Remark (p.80). Fix an operator $\mathrm{T} \in \mathfrak{L}(V)$, then the function from $\mathfrak{P}(\mathbb{F})$ to $\mathfrak{L}(V)$ given by $p \mapsto p(\mathrm{T})$ is linear.

Proof. Let the mapping be $A\colon \mathfrak{P}(\mathbb{F}) \to \mathfrak{L}(V)$ with $A(p) = p(\mathrm{T})$. For any $p, q \in \mathfrak{P}(\mathbb{F})$, we have $A(p+q) = (p+q)(\mathrm{T}) = p(\mathrm{T}) + q(\mathrm{T}) = A(p) + A(q)$. For any $a \in \mathbb{F}$, we have $A(ap) = (ap)(\mathrm{T}) = ap(\mathrm{T}) = aA(p)$. □

Exercises

► Exercise 64 (5.1). *Suppose* $\mathrm{T} \in \mathfrak{L}(V)$*. Prove that if* $U_1, \dots, U_m$ *are subspaces of* V *invariant under* T*, then* $U_1 + \cdots + U_m$ *is invariant under* T*.*

Proof. Take an arbitrary $\boldsymbol{u} \in U_1 + \cdots + U_m$; then $\boldsymbol{u} = \boldsymbol{u}_1 + \cdots + \boldsymbol{u}_m$, where $\boldsymbol{u}_i \in U_i$ for every $i = 1, \dots, m$. Therefore, $\mathrm{T}\boldsymbol{u} = \mathrm{T}\boldsymbol{u}_1 + \cdots + \mathrm{T}\boldsymbol{u}_m \in U_1 + \cdots + U_m$ since $\mathrm{T}\boldsymbol{u}_i \in U_i$. □

► Exercise 65 (5.2). *Suppose* $\mathrm{T} \in \mathfrak{L}(V)$*. Prove that the intersection of any collection of subspaces of* V *invariant under* T *is invariant under* T*.*

Proof. Let the collection $\{U_i \trianglelefteq V : i \in I\}$ of subspaces of V invariant under T, where I is an index set. Let $U = \bigcap_{i \in I} U_i$. Then $\boldsymbol{u} \in U_i$ for every $i \in I$ if $\boldsymbol{u} \in U$, and so $\mathrm{T}\boldsymbol{u} \in U_i$ for every $i \in I$. Then $\mathrm{T}\boldsymbol{u} \in U$; that is, U is invariant under T. □

► Exercise 66 (5.3). *Prove or give a counterexample: if* U *is a subspace of* V *that is invariant under every operator on* V*, then* $U = \{\mathbf{0}\}$ *or* $U = V$*.*

Proof. Assume that $U \neq \{\mathbf{0}\}$ and $U \neq V$. Let $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_m)$ be a basis of U, which then can be extended to a basis $(\boldsymbol{u}_1, \dots, \boldsymbol{u}_m, \boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ of V, where $n \geqslant 1$ since $U \neq V$. Define an operator $\mathrm{T} \in \mathfrak{L}(V)$ by letting $\mathrm{T}(a_1\boldsymbol{u}_1 + \cdots + a_m\boldsymbol{u}_m +$

$b_1\boldsymbol{v}_1 + \cdots + b_n\boldsymbol{v}_n) = (a_1 + \cdots + a_m + b_1 + \cdots + b_n)\boldsymbol{v}_1$. Then U fails be invariant clearly. ◻

▶ EXERCISE 67 (5.4). *Suppose that* $\mathrm{S}, \mathrm{T} \in \mathfrak{L}(V)$ *are such that* $\mathrm{ST} = \mathrm{TS}$. *Prove that* $\mathcal{N}_{\mathrm{T}-\lambda\mathrm{Id}}$ *is invariant under* S *for every* $\lambda \in \mathbb{F}$.

PROOF. If $\boldsymbol{u} \in \mathcal{N}_{\mathrm{T}-\lambda\mathrm{Id}}$, then $(\mathrm{T} - \lambda\mathrm{Id})(\boldsymbol{u}) = \mathrm{T}\boldsymbol{u} - \lambda\boldsymbol{u} = \boldsymbol{0}$; hence

$$
\begin{aligned}
\mathrm{S}(\mathrm{T}\boldsymbol{u} - \lambda\boldsymbol{u}) = \mathrm{S}\boldsymbol{0} &\iff \mathrm{ST}\boldsymbol{u} - \lambda\mathrm{S}\boldsymbol{u} = \boldsymbol{0} \\
&\iff \mathrm{TS}\boldsymbol{u} - \lambda\mathrm{S}\boldsymbol{u} = \boldsymbol{0} \\
&\iff (\mathrm{T} - \lambda\mathrm{Id})(\mathrm{S}\boldsymbol{u}) = \boldsymbol{0};
\end{aligned}
$$

that is, $\mathrm{S}\boldsymbol{u} \in \mathcal{N}_{\mathrm{T}-\lambda\mathrm{Id}}$. ◻

▶ EXERCISE 68 (5.5). *Define* $\mathrm{T} \in \mathfrak{L}(\mathbb{F}^2)$ *by* $\mathrm{T}(w, z) = (z, w)$. *Find all eigenvalues and eigenvectors of* T.

PROOF. $\mathrm{T}\boldsymbol{u} = \lambda\boldsymbol{u}$ implies that $(z, w) = (\lambda w, \lambda z)$. Hence, $\lambda_1 = 1$, $\lambda_2 = -1$, and the corresponding eigenvectors are $(1, 1)$ and $(1, -1)$. Since $\dim \mathbb{F}^2 = 2$, they are the all eigenvalues and eigenvectors of T. ◻

▶ EXERCISE 69 (5.6). *Define* $\mathrm{T} \in \mathfrak{L}(\mathbb{F}^3)$ *by* $\mathrm{T}(z_1, z_2, z_3) = (2z_2, 0, 5z_3)$. *Find all eigenvalues and eigenvectors of* T.

PROOF. If $\lambda \in \mathbb{F}$ is a eigenvalue of T and $(z_1, z_2, z_3) \neq \boldsymbol{0}$ is a corresponding eigenvector, then $\mathrm{T}(z_1, z_2, z_3) = \lambda(z_1, z_2, z_3)$, that is,

$$
\begin{cases}
2z_2 = \lambda z_1 & \text{(i)} \\
0 = \lambda z_2 & \text{(ii)}, \\
5z_3 = \lambda z_3 & \text{(iii)}.
\end{cases}
\tag{5.1}
$$

- If $z_2 \neq 0$, then $\lambda = 0$ from (ii); but then $z_2 = 0$ from (i). A contradiction. Hence, $z_2 = 0$ and (5.1) becomes

$$
\begin{cases}
0 = \lambda z_1 & \text{(i}'\text{)} \\
5z_3 = \lambda z_3 & \text{(iii}'\text{)}.
\end{cases}
\tag{5.2}
$$

- If $z_3 \neq 0$, then $\lambda = 5$ from (iii′); then (i′) implies that $z_1 = 0$. Hence, $\lambda = 5$ is an eigenvalue, and the corresponding eigenvector is $(0, 0, 1)$.
- If $z_1 \neq 0$, then $\lambda = 0$ from (i′); then (iii′) implies that $z_3 = 0$. Hence, $\lambda = 0$ is an eigenvalue, and the corresponding eigenvector is $(1, 0, 0)$. ◻

▶ EXERCISE 70 (5.7). *Suppose* n *is a positive integer and* $\mathrm{T} \in \mathfrak{L}(\mathbb{F}^n)$ *is defined by*

$$
\mathrm{T}(x_1, \ldots, x_n) = \left(\sum_{i=1}^{n} x_i, \ldots, \sum_{i=1}^{n} x_i \right);
$$

in other words, T *is the operator whose matrix (with respect to the standard basis) consists of all* 1*'s. Find all eigenvalues and eigenvectors of* T*.*

Proof. If $\lambda \in \mathbb{F}$ is an eigenvalue of T and $(x_1, \dots, x_n) \neq \mathbf{0}$ is a corresponding eigenvector, then $\sum_{i=1}^n x_i \neq 0$ and

$$\begin{pmatrix} \sum_{i=1}^n x_i \\ \vdots \\ \sum_{i=1}^n x_i \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix}.$$

Hence, $\lambda \neq 0$, $x_i \neq 0$ for all $i = 1, \dots, n$, and $\lambda x_1 = \cdots = \lambda x_n$ implies that $x_1 = \cdots = x_n$, and so the unique eigenvalue of T is $(\sum_{i=1}^n x_i)/x_i = n$. Then an eigenvector to n is $(1, \dots, 1)$. □

► Exercise 71 (5.8). *Find all eigenvalues and eigenvectors of the backward shift operator* $\mathrm{T} \in \mathfrak{L}(\mathbb{F}^\infty)$ *defined by* $\mathrm{T}(z_1, z_2, z_3, \dots) = (z_2, z_3, \dots)$*.*

Proof. For any $\lambda \in \mathbb{F}$ with $\lambda \neq 0$, we have $\mathrm{T}(\lambda, \lambda^2, \lambda^3, \dots) = (\lambda^2, \lambda^3, \dots) = \lambda \cdot (\lambda, \lambda^2, \dots)$; hence, every $\lambda \neq 0$ is an eigenvalue of T. We now show that $\lambda = 0$ is also an eigenvalue: let $\boldsymbol{z} = (z_1, 0, \dots)$ with $z_1 \neq 0$. Then $\mathrm{T}\boldsymbol{z} = (0, 0, \dots) = 0 \cdot \boldsymbol{z}$. □

► Exercise 72 (5.9). *Suppose* $\mathrm{T} \in \mathfrak{L}(V)$ *and* $\dim \mathcal{R}_\mathrm{T} = k$*. Prove that* T *has at most* $k + 1$ *distinct eigenvalues.*

Proof. Suppose that T has more than or equal to $k+2$ distinct eigenvalues. We take the first $k + 2$ eigenvalues: $\lambda_1, \dots, \lambda_{k+2}$. Then there are $k + 2$ corresponding nonzero eigenvectors, $\boldsymbol{u}_1, \dots, \boldsymbol{u}_{k+2}$, satisfying $\mathrm{T}\boldsymbol{u}_1 = \lambda_1 \boldsymbol{u}_1, \dots, \mathrm{T}\boldsymbol{u}_{k+2} = \lambda_{k+2}\boldsymbol{u}_{k+2}$. Since the $k + 2$ eigenvectors are linearly independent, the list $(\lambda_1 \boldsymbol{u}_1, \dots, \lambda_{k+2}\boldsymbol{u}_{k+2})$ is linearly independent, too (there are $n + 1$ vectors if one λ is zero). Obviously, the above list is in $\mathcal{R}_\mathrm{T}$, which means that $\dim \mathcal{R}_\mathrm{T} \geqslant k + 1$. A contradiction. □

► Exercise 73 (5.10). *Suppose* $\mathrm{T} \in \mathfrak{L}(V)$ *is invertible and* $\lambda \in \mathbb{F} \setminus \{0\}$*. Prove that* λ *is an eigenvalue of* T *if and only if* $1/\lambda$ *is an eigenvalue of* T^{-1}*.*

Proof. If $\lambda \neq 0$ be an eigenvalue of T, then there exists a nonzero $\boldsymbol{u} \in V$ such that $\mathrm{T}\boldsymbol{u} = \lambda \boldsymbol{u}$. Therefore,

$$\mathrm{T}^{-1}(\mathrm{T}\boldsymbol{u}) = \mathrm{T}^{-1}(\lambda \boldsymbol{u}) \iff \boldsymbol{u} = \lambda \mathrm{T}^{-1}\boldsymbol{u} \iff \mathrm{T}^{-1}\boldsymbol{u} = \boldsymbol{u}/\lambda;$$

that is, $1/\lambda$ is an eigenvalue of T^{-1}. The other direction can be proved with the same way. □

► Exercise 74 (5.11). *Suppose* $\mathrm{S}, \mathrm{T} \in \mathfrak{L}(V)$*. Prove that* ST *and* TS *have the same eigenvalues.*

Proof. Let λ be an eigenvalue of ST, and $\boldsymbol{u} \neq \mathbf{0}$ be the corresponding eigenvector. Then $(\mathrm{ST})\boldsymbol{u} = \lambda \boldsymbol{u}$. Therefore,

$$\mathrm{T}(\mathrm{ST}\boldsymbol{u}) = \mathrm{T}(\lambda \boldsymbol{u}) \iff (\mathrm{TS})(\mathrm{T}\boldsymbol{u}) = \lambda(\mathrm{T}\boldsymbol{u}).$$

Hence, if $\mathrm{T}\boldsymbol{u} \neq \mathbf{0}$, then λ is an eigenvalue of TS, and the corresponding eigenvector is $\mathrm{T}\boldsymbol{u}$; if $\mathrm{T}\boldsymbol{u} = \mathbf{0}$, then $(\mathrm{ST})\boldsymbol{u} = \mathrm{S}(\mathrm{T}\boldsymbol{u}) = \mathbf{0}$ implies that $\lambda = 0$ (since $\boldsymbol{u} \neq \mathbf{0}$). In this case, T is not injective, and so TS is not injective (by Exercise 54). But this means that there exists $\boldsymbol{v} \neq \mathbf{0}$ such that $(\mathrm{TS})\boldsymbol{v} = \mathbf{0} = 0\boldsymbol{v}$; that is, 0 is an eigenvalue of TS. The other direction can be proved with the same way. □

► EXERCISE 75 (5.12). *Suppose* $\mathrm{T} \in \mathfrak{L}(V)$ *is such that every vector in* V *is an eigenvector of* T. *Prove that* T *is a scalar multiple of the identity operator.*

PROOF. Let $B = (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ be a basis of V and take arbitrary $\boldsymbol{v}_i$ and $\boldsymbol{v}_j$ from B. Then there are λ_i and λ_j such that $\mathrm{T}\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i$ and $\mathrm{T}\boldsymbol{v}_j = \lambda_j \boldsymbol{v}_j$. Since $\boldsymbol{v}_i + \boldsymbol{v}_j$ is also an eigenvector, there is $\lambda \in \mathbb{F}$ such that $\mathrm{T}(\boldsymbol{v}_i + \boldsymbol{v}_j) = \lambda(\boldsymbol{v}_i + \boldsymbol{v}_j)$. Therefore,

$$\lambda_i \boldsymbol{v}_i + \lambda_j \boldsymbol{v}_j = \lambda \boldsymbol{v}_i + \lambda \boldsymbol{v}_j;$$

that is, $(\lambda_i - \lambda)\,\boldsymbol{v}_i + \left(\lambda_j - \lambda\right)\boldsymbol{v}_j = \mathbf{0}$. Since $(\boldsymbol{v}_i, \boldsymbol{v}_j)$ is linearly independent, we have $\lambda_i = \lambda_j = \lambda$. Hence, for any $\boldsymbol{v} = \sum_{i=1}^n a_i \boldsymbol{v}_i \in V$, we have

$$\mathrm{T}\boldsymbol{v} = \mathrm{T}\left(\sum_{i=1}^n a_i \boldsymbol{v}_i\right) = \sum_{i=1}^n a_i \lambda \boldsymbol{v}_i = \lambda \left(\sum_{i=1}^n a_i \boldsymbol{v}_i\right) = \lambda \boldsymbol{v},$$

i.e., $\mathrm{T} = \lambda \mathrm{Id}$. □

► EXERCISE 76 (5.13). *Suppose* $\mathrm{T} \in \mathfrak{L}(V)$ *is such that every subspace of* V *with dimension* $\dim V - 1$ *is invariant under* T. *Prove that* T *is a scalar multiple of the identity operator.*

PROOF. Let $\dim V = n$ and $(\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n)$ be a basis of V. We first show that there exists $\lambda_1 \in \mathbb{F}$ such that $\mathrm{T}\boldsymbol{v}_1 = \lambda_1 \boldsymbol{v}_1$.

Let $V_1 = \{a\boldsymbol{v}_1 : a \in \mathbb{F}\}$ and $U_1 = \mathrm{span}(\boldsymbol{v}_2, \dots, \boldsymbol{v}_n)$. Then for any $\boldsymbol{v} = \sum_{i=1}^n a_i \boldsymbol{v}_i \in V$, we have

$$\begin{aligned}
\mathrm{T}\boldsymbol{v} = \mathrm{T}\left(a_1 \boldsymbol{v}_1 + \sum_{i=2}^n a_i \boldsymbol{v}_i\right) &= a_1 \mathrm{T}\boldsymbol{v}_1 + \mathrm{T}\left(\sum_{i=2}^n a_i \boldsymbol{v}_i\right) \\
&= a_1 \left(\sum_{j=1}^n b_j \boldsymbol{v}_j\right) + \mathrm{T}\left(\sum_{i=2}^n a_i \boldsymbol{v}_i\right) \\
&= (a_1 b_1)\,\boldsymbol{v}_1 + \left[\sum_{i=2}^n (a_1 b_i)\,\boldsymbol{v}_i + \mathrm{T}\left(\sum_{i=2}^n a_i \boldsymbol{v}_i\right)\right] \\
&\in V_1 + U_1,
\end{aligned}$$

where $\mathrm{T}\left(\sum_{i=2}^n a_i \boldsymbol{v}_i\right) \in U_1$ since U_1 is invariant under T.

Since $V = V_1 + U_1$ and $\dim V = \dim V_1 + \dim U_1$, we have $V = V_1 \oplus U_1$ by Proposition 2.19, which implies that $V_1 \cap U_1 = \{\mathbf{0}\}$ by Proposition 1.9. If $\mathrm{T}\boldsymbol{v}_1 \notin V_1$, then $\mathrm{T}\boldsymbol{v}_1 \neq \mathbf{0}$ and $\mathrm{T}\boldsymbol{v}_1 \in U_1$; hence, there exist $c_2, \ldots, c_n \in \mathbb{F}$ not all zero such that

$$\mathrm{T}\boldsymbol{v}_1 = \sum_{i=2}^{n} c_i \boldsymbol{v}_i .$$

Without loss of generality, we suppose that $c_n \neq 0$.

Let $V_n = \{a\boldsymbol{v}_n : a \in \mathbb{F}\}$ and $U_n = \operatorname{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{n-1})$. Similarly, $V = V_n \oplus U_n$ and $V_n \cap U_n = \{\mathbf{0}\}$. Since U_n is invariant under T, we have $\mathrm{T}\boldsymbol{v}_1 \in U_n$, that is, $\mathrm{T}\boldsymbol{v}_1 = \sum_{j=1}^{n-1} d_j \boldsymbol{v}_j$, but which means that $c_n = 0$ A contradiction. We thus proved that $\mathrm{T}\boldsymbol{v}_1 \in V_1$, i.e., there is $\lambda_1 \in \mathbb{F}$ such that $\mathrm{T}\boldsymbol{v}_1 = \lambda_1 \boldsymbol{v}_1$. But this way can be applied to every $\boldsymbol{v}_i$. Therefore, every $\boldsymbol{v}_i$ is an eigenvector of T. By Exercise 75, T is a scalar multiple of the identity operator. □

► Exercise 77 (5.14). *Suppose* $\mathrm{S}, \mathrm{T} \in \mathfrak{L}(V)$ *and* S *is invertible. Prove that if* $p \in \mathfrak{P}(\mathbb{F})$ *is a polynomial, then* $p(\mathrm{STS}^{-1}) = \mathrm{S}p(\mathrm{T})\mathrm{S}^{-1}$.

Proof. Let $p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_m z^m$. Then

$$p(\mathrm{STS}^{-1}) = a_0\mathrm{Id} + a_1 \cdot (\mathrm{STS}^{-1}) + a_2 \cdot (\mathrm{STS}^{-1})^2 + \cdots + a_m \cdot (\mathrm{STS}^{-1})^m .$$

We also have

$$\begin{aligned}(\mathrm{STS}^{-1})^n &= (\mathrm{STS}^{-1}) \cdot (\mathrm{STS}^{-1}) \cdot (\mathrm{STS}^{-1})^{n-2} \\ &= (\mathrm{ST}^2\mathrm{S}^{-1}) \cdot (\mathrm{STS}^{-1})^{n-2} \\ &= \cdots \\ &= \mathrm{ST}^n\mathrm{S}^{-1} .\end{aligned}$$

Therefore,

$$\begin{aligned}\mathrm{S}p(\mathrm{T})\mathrm{S}^{-1} &= \mathrm{S}(a_0\mathrm{Id} + a_1\mathrm{T} + a_2\mathrm{T}^2 + \cdots + a_m\mathrm{T}^m)\mathrm{S}^{-1} \\ &= a_0\mathrm{Id} + a_1 \cdot (\mathrm{STS}^{-1}) + a_2 \cdot (\mathrm{ST}^2\mathrm{S}^{-1}) + \cdots + a_m \cdot (\mathrm{ST}^m\mathrm{S}^{-1}) \\ &= p(\mathrm{STS}^{-1}) .\end{aligned}$$

□

► Exercise 78 (5.15). *Suppose* $\mathbb{F} = \mathbb{C}$, $\mathrm{T} \in \mathfrak{L}(V)$, $p \in \mathfrak{P}(\mathbb{C})$, *and* $a \in \mathbb{C}$. *Prove that* a *is an eigenvalue of* $p(\mathrm{T})$ *if and only if* $a = p(\lambda)$ *for some eigenvalue* λ *of* T.

Proof. If λ is an eigenvalue of T, then there exists $\boldsymbol{v} \neq \mathbf{0}$ such that $\mathrm{T}\boldsymbol{v} = \lambda\boldsymbol{v}$. Thus,

$$\begin{aligned}[p(\mathrm{T})](\boldsymbol{v}) &= (a_0\mathrm{Id} + a_1\mathrm{T} + a_2\mathrm{T}^2 + \cdots + a_m\mathrm{T}^m)\boldsymbol{v} \\ &= a_0\boldsymbol{v} + a_1\mathrm{T}\boldsymbol{v} + a_2\mathrm{T}^2\boldsymbol{v} + \cdots a_m\mathrm{T}^m\boldsymbol{v} \\ &= a_0\boldsymbol{v} + a_1\lambda\boldsymbol{v} + a_2\mathrm{T}(\lambda\boldsymbol{v}) + \cdots + a_m\mathrm{T}^{m-1}(\lambda\boldsymbol{v}) \\ &= a_0\boldsymbol{v} + (a_1\lambda)\boldsymbol{v} + (a_2\lambda^2)\boldsymbol{v} + \cdots + (a_m\lambda^m)\boldsymbol{v} \\ &= p(\lambda)\boldsymbol{v};\end{aligned}$$

that is, $p(\lambda)$ is an eigenvalue of $p(\mathrm{T})$.

Conversely, let $a \in \mathbb{C}$ be an eigenvalue of $p(\mathrm{T}) = a_0\mathrm{Id} + a_1\mathrm{T} + \cdots + a_m\mathrm{T}^m$, and $\boldsymbol{v}$ be the corresponding eigenvector. Then $p(\mathrm{T})(\boldsymbol{v}) = a\boldsymbol{v}$; that is,

$$\left[(a_0 - a)\mathrm{Id} + a_1\mathrm{T} + \cdots + a_m\mathrm{T}^m\right]\boldsymbol{v} = \mathbf{0}.$$

It follows from Corollary 4.8 that the above display can be rewritten as follows:

$$\left[c(\mathrm{T} - \lambda_1\mathrm{Id}) \cdots (\mathrm{T} - \lambda_m\mathrm{Id})\right]\boldsymbol{v} = \mathbf{0}, \tag{5.3}$$

where $c, \lambda_1, \ldots, \lambda_m \in \mathbb{C}$ and $c \neq 0$. Hence, for some $i = 1, \ldots, m$, we have $(\mathrm{T} - \lambda_i\mathrm{Id})\boldsymbol{v} = \mathbf{0}$; that is, λ_i is an eigenvalue of T. □

► EXERCISE 79 (5.16). *Show that the result in the previous exercise does not hold if* $\mathbb{C}$ *is replaced with* $\mathbb{R}$.

PROOF. Let $\mathrm{T} \in \mathfrak{L}(\mathbb{R}^2)$ defined by $\mathrm{T}(w, z) = (-z, w)$. Then T has no eigenvalue (see p. 78). But $\mathrm{T}^2(w, z) = \mathrm{T}(-z, w) = (-w, -z)$ has an eigenvalue: let $(-w, -z) = \lambda(w, z)$; then

$$\begin{cases} -w = \lambda w \\ -z = \lambda z. \end{cases}$$

Hence, $\lambda = -1$. □

► EXERCISE 80 (5.17). *Suppose* V *is a complex vector space and* $\mathrm{T} \in \mathfrak{L}(V)$. *Prove that* T *has an invariant subspace of dimension* j *for each* $j = 1, \ldots, \dim V$.

PROOF. Suppose that $\dim V = n$. Let $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ be a basis of V with respect to which T has an upper-triangular matrix (by Theorem 5.13)

$$\mathcal{M}\left(\mathrm{T}, (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)\right) = \begin{pmatrix} \lambda_1 & & & * \\ & \lambda_2 & & \\ & & \ddots & \\ \mathbf{0} & & & \lambda_n \end{pmatrix}.$$

Then it follows from Proposition 5.12 that the claim holds. □

► EXERCISE 81 (5.18). *Give an example of an operator whose matrix with respect to some basis contains only* 0*'s on the diagonal, but the operator is invertible.*

PROOF. Let $\mathrm{T} \in \mathfrak{L}(\mathbb{R}^2)$. Take the standard basis $\left((0,1), (1,0)\right)$ of $\mathbb{R}^2$, with respect to which T has the following matrix

$$\mathcal{M}(\mathrm{T}) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Then $\mathrm{T}(x, y) = \mathcal{M}(\mathrm{T}) \cdot (x, y)' = \left(y, x\right)$ is invertible. □

► EXERCISE 82 (5.19). *Give an example of an operator whose matrix with respect to some basis contains only nonzero numbers on the diagonal, but the operator is not invertible.*

PROOF. Consider the standard basis $\big((1,0),(0,1)\big)$ of $\mathbb{R}^2$. Let $\mathrm{T} \in \mathcal{L}(\mathbb{R}^2)$ be defined as $\mathrm{T}(x,y) = (x,0)$. Then T is not injective and so is not invertible. Its matrix is

$$\mathcal{M}(\mathrm{T}) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \qquad \square$$

► EXERCISE 83 (5.20). *Suppose that* $\mathrm{T} \in \mathcal{L}(V)$ *has* $\dim V$ *distinct eigenvalues and that* $\mathrm{S} \in \mathcal{L}(V)$ *has the same eigenvectors as* T *(not necessarily with the same eigenvalues). Prove that* $\mathrm{ST} = \mathrm{TS}$*.*

PROOF. Let $\dim V = n$. Let $\lambda_1, \dots, \lambda_n$ be n distinct eigenvalues of T and $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ be n eigenvector corresponding to the eigenvalues. Then $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ is independent and so is a basis of V. Further, the matrix of T with respect to $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ is given by

$$\mathcal{M}\big(\mathrm{T}, (\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)\big) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Since S has the same eigenvectors as T, so for any $\boldsymbol{v}_i$, there is some $\widehat{\lambda}_i$ such that $\mathrm{S}\boldsymbol{v}_i = \widehat{\lambda}_i \boldsymbol{v}_i$. For every $\boldsymbol{v} = \sum_{i=1}^n a_i \boldsymbol{v}_i \in V$ we have

$$\begin{aligned}
(\mathrm{ST})(\boldsymbol{v}) = \mathrm{S}\left[\mathrm{T}\left(\sum_{i=1}^n a_i \boldsymbol{v}_i\right)\right] = \mathrm{S}\left(\sum_{i=1}^n a_i \mathrm{T}\boldsymbol{v}_i\right) &= \mathrm{S}\left(\sum_{i=1}^n a_i \lambda_i \boldsymbol{v}_i\right) \\
&= \sum_{i=1}^n (a_i \lambda_i)\, \mathrm{S}\boldsymbol{v}_i \\
&= \sum_{i=1}^n \left(a_i \lambda_i \widehat{\lambda}_i\right) \boldsymbol{v}_i,
\end{aligned}$$

and

$$\begin{aligned}
(\mathrm{TS})(\boldsymbol{v}) = \mathrm{T}\left[\mathrm{S}\left(\sum_{i=1}^n a_i \boldsymbol{v}_i\right)\right] = \mathrm{T}\left(\sum_{i=1}^n a_i \mathrm{S}\boldsymbol{v}_i\right) &= \mathrm{T}\left(\sum_{i=1}^n \left(a_i \widehat{\lambda}_i\right) \boldsymbol{v}_i\right) \\
&= \sum_{i=1}^n \left(a_i \widehat{\lambda}_i\right) \mathrm{T}\boldsymbol{v}_i \\
&= \sum_{i=1}^n \left(a_i \lambda_i \widehat{\lambda}_i\right) \boldsymbol{v}_i.
\end{aligned}$$

Hence, ST = TS. □

▶ EXERCISE 84 (5.21). *Suppose* $\mathbf{P} \in \mathcal{L}(V)$ *and* $\mathbf{P}^2 = \mathbf{P}$. *Prove that* $V = \mathcal{N}_{\mathbf{P}} \oplus \mathcal{R}_{\mathbf{P}}$.

PROOF. By Theorem 3.4, $\dim V = \dim \mathcal{N}_{\mathbf{P}} + \dim \mathcal{R}_{\mathbf{P}}$, so it suffices to show that $V = \mathcal{N}_{\mathbf{P}} + \mathcal{R}_{\mathbf{P}}$ by Proposition 2.19. Take an arbitrary $\boldsymbol{v} \in V$. Since $\mathbf{P}^2 = \mathbf{P}$, we have

$$\mathbf{P}^2\boldsymbol{v} = \mathbf{P}\boldsymbol{v} \iff \mathbf{P}(\mathbf{P}\boldsymbol{v} - \boldsymbol{v}) = \mathbf{0} \iff \mathbf{P}\boldsymbol{v} - \boldsymbol{v} \in \mathcal{N}_{\mathbf{P}};$$

that is, there exists $\boldsymbol{u} \in \mathcal{N}_{\mathbf{P}}$ such that $\mathbf{P}\boldsymbol{v} - \boldsymbol{v} = \boldsymbol{u}$. Therefore,

$$\boldsymbol{v} = -\boldsymbol{u} + \mathbf{P}\boldsymbol{v} \in \mathcal{N}_{\mathbf{P}} + \mathcal{R}_{\mathbf{P}}.$$ □

▶ EXERCISE 85 (5.22). *Suppose* $V = U \oplus W$, *where* U *and* W *are nonzero subspaces of* V. *Find all eigenvalues and eigenvectors of* $\mathbf{P}_{U,W}$.

PROOF. We first show that $\lambda = 0$ is an eigenvalue of $\mathbf{P}_{U,W}$. Since $W \neq \{\mathbf{0}\}$, we can take $\boldsymbol{w} \in W$ with $\boldsymbol{w} \neq \mathbf{0}$. Obviously, $\boldsymbol{w} \in V$ and $\boldsymbol{w}$ can be written as $\boldsymbol{w} = \mathbf{0} + \boldsymbol{w}$ uniquely. Then

$$\mathbf{P}_{U,W}(\boldsymbol{w}) = \mathbf{0} = 0\boldsymbol{w};$$

that is, 0 is an eigenvalue of $\mathbf{P}_{U,W}$ and any $\boldsymbol{w} \in W$ with $\boldsymbol{w} \neq \mathbf{0}$ is an eigenvector corresponding to 0.

Now let us check whether there is eigenvalue $\lambda \neq 0$. If there is an eigenvalue $\lambda \neq 0$ under $\mathbf{P}_{U,W}$, then there exists $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w} \neq \mathbf{0}$, where $\boldsymbol{u} \in U$ and $\boldsymbol{w} \in W$, such that $\mathbf{P}_{U,W}(\boldsymbol{v}) = \lambda\boldsymbol{v}$, but which means that

$$\boldsymbol{u} = \lambda(\boldsymbol{u} + \boldsymbol{w}).$$

Then $\boldsymbol{w} = (1-\lambda)\boldsymbol{u}/\lambda \in U$ since $\lambda \neq 0$, and which implies that $\boldsymbol{w} = \mathbf{0}$ since $V = U \oplus W$ forces $U \cap W = \{\mathbf{0}\}$. Therefore, $\boldsymbol{v} = \boldsymbol{u} \neq \mathbf{0}$ and

$$\mathbf{P}_{U,W}(\boldsymbol{v})\mathbf{P}_{U,W}(\boldsymbol{u}) = \boldsymbol{u} = 1 \cdot \boldsymbol{u},$$

that is, $\lambda = 1$ is the unique nonzero eigenvalue of $\mathbf{P}_{U,W}$. □

▶ EXERCISE 86 (5.23). *Give an example of an operator* $\mathrm{T} \in \mathcal{L}(\mathbb{R}^4)$ *such that* T *has no (real) eigenvalues.*

PROOF. Our example is based on (5.4). Let $\mathrm{T} \in \mathcal{L}(\mathbb{R}^4)$ be defined by

$$\mathrm{T}(x_1, x_2, x_3, x_4) = (-x_2, x_1, -x_4, x_3).$$

Suppose that λ is a (real) eigenvalue of T; then

$$\begin{cases} \lambda x_1 = -x_2 \\ \lambda x_2 = x_1 \\ \lambda x_3 = -x_4 \\ \lambda x_4 = x_3. \end{cases}$$

If $\lambda = 0$, then $(x_1, x_2, x_3, x_4) = \mathbf{0}$. So $\lambda \neq 0$. It is evident that

$$x_1 \neq 0 \iff x_2 \neq 0, \quad \text{and} \quad x_3 \neq 0 \iff x_4 \neq 0.$$

Suppose that $x_1 \neq 0$. Then from the first two equations we have

$$\lambda^2 x_2 = -x_2 \implies \lambda^2 = -1,$$

which has no solution in $\mathbb{R}$. Hence, $x_1 = x_2 = 0$ when $\lambda \neq 0$. Similarly, we can show that $x_3 = x_4 = 0$ if $\lambda \neq 0$. □

► EXERCISE 87 (5.24). *Suppose V is a real vector space and $\mathrm{T} \in \mathfrak{L}(V)$ has no eigenvalues. Proves that every subspace of V invariant under T has even dimension.*

PROOF. If U is invariant under T and $\dim U$ is odd, then $\mathrm{T}\restriction_U \in \mathfrak{L}(U)$ has an eigenvalue. But this implies that T has an eigenvalue. A contradiction. □

6

INNER-PRODUCT SPACES

"AS YOU SHOULD VERIFY"

REMARK (p. 113). The orthogonal projection $\mathbf{P}_U$ has the following properties:

a. $\mathbf{P}_U \in \mathfrak{L}(V)$;

b. $\mathcal{R}_{\mathbf{P}_U} = U$;

c. $\mathcal{N}_{\mathbf{P}_U} = U^\perp$;

d. $\boldsymbol{v} - \mathbf{P}_U \boldsymbol{v} \in U^\perp$ for every $\boldsymbol{v} \in V$;

e. $\mathbf{P}_U^2 = \mathbf{P}_U$;

f. $\|\mathbf{P}_U \boldsymbol{v}\| \leqslant \|\boldsymbol{v}\|$ for every $\boldsymbol{v} \in V$;

g. $\mathbf{P}_U \boldsymbol{v} = \sum_{i=1}^m \langle \boldsymbol{v}, \boldsymbol{e}_i \rangle \boldsymbol{e}_i$ for every $\boldsymbol{v} \in V$, where $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m)$ is a basis of U.

PROOF. (a) For any $\boldsymbol{v}, \boldsymbol{v}' \in V$, we have

$$\begin{aligned}\mathbf{P}_U(\boldsymbol{v} + \boldsymbol{v}') = \mathbf{P}_U\left[(\boldsymbol{u} + \boldsymbol{w}) + (\boldsymbol{u}' + \boldsymbol{w}')\right] &= \mathbf{P}_U\left[(\boldsymbol{u} + \boldsymbol{u}') + (\boldsymbol{w} + \boldsymbol{w}')\right] \\ &= \boldsymbol{u} + \boldsymbol{u}' \\ &= \mathbf{P}_U \boldsymbol{v} + \mathbf{P}_U \boldsymbol{v}',\end{aligned}$$

where $\boldsymbol{u}, \boldsymbol{u}' \in U$ and $\boldsymbol{w}, \boldsymbol{w}' \in U^\perp$. Also it is true that $\mathbf{P}_U(a\boldsymbol{v}) = a\mathbf{P}_U \boldsymbol{v}$. Therefore, $\mathbf{P}_U \in \mathfrak{L}(V)$.

(b) Write every $\boldsymbol{v} \in V$ as $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$, where $\boldsymbol{u} \in U$ and $\boldsymbol{w} \in U^\perp$. Since $\mathbf{P}_U \boldsymbol{v} = \boldsymbol{u}$, we have one direction that $\mathcal{R}_{\mathbf{P}_U} \subseteq U$. For the other direction, notice that $U = \mathbf{P}_U[U] \subseteq \mathcal{R}_{\mathbf{P}_U}$.

(c) If $\boldsymbol{v} \in \mathcal{N}_{\mathbf{P}_U}$, then $\mathbf{0} = \mathbf{P}_U \boldsymbol{v} = \boldsymbol{u}$; that is, $\boldsymbol{v} = \mathbf{0} + \boldsymbol{w}$ with $\boldsymbol{w} \in U^\perp$. This proves that $\mathcal{N}_{\mathbf{P}_U} \subseteq U^\perp$. The other inclusion direction is clear.

(d) For every $\boldsymbol{v} \in V$, we have $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$, where $\boldsymbol{u} \in U$ and $\boldsymbol{w} \in U^\perp$. Hence, $\boldsymbol{v} - \mathbf{P}_U \boldsymbol{v} = (\boldsymbol{u} + \boldsymbol{w}) - \boldsymbol{u} = \boldsymbol{w} \in U^\perp$.

(e) For every $\boldsymbol{v} \in V$, we have $\mathbf{P}_U^2 \boldsymbol{v} = \mathbf{P}_U(\mathbf{P}_U \boldsymbol{v}) = \mathbf{P}_U \boldsymbol{u} = \boldsymbol{u} = \mathbf{P}_U \boldsymbol{v}$.

(f) We can write every $\boldsymbol{v} \in V$ as $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$ with $\boldsymbol{u} \in U$ and $\boldsymbol{w} \in U^{\perp}$; therefore, $\|\boldsymbol{v}\| = \|\boldsymbol{u} + \boldsymbol{w}\| \overset{*}{=} \|\boldsymbol{u}\| + \|\boldsymbol{w}\| \geqslant \|\boldsymbol{u}\| = \|\mathbf{P}_U \boldsymbol{v}\|$, where $(*)$ holds since $U \perp U^{\perp}$.

(g) It follows from Axler (1997, 6.31, p.112). □

REMARK (p. 119-120). Verify that the function $\mathrm{T} \mapsto \mathrm{T}^*$ has the following properties:

a. $(\mathrm{S} + \mathrm{T})^* = \mathrm{S}^* + \mathrm{T}^*$ for all $\mathrm{S}, \mathrm{T} \in \mathfrak{L}(V, W)$;

b. $(a\mathrm{T})^* = \overline{a}\,\mathrm{T}^*$ for all $a \in \mathbb{F}$ and $\mathrm{T} \in \mathfrak{L}(V, W)$;

c. $(\mathrm{T}^*)^* = \mathrm{T}$ for all $\mathrm{T} \in \mathfrak{L}(V, W)$;

d. $\mathrm{Id}^* = \mathrm{Id}$, where Id is the identity operator on V;

e. $(\mathrm{ST})^* = \mathrm{T}^*\mathrm{S}^*$ for all $\mathrm{T} \in \mathfrak{L}(V, W)$ and $\mathrm{S} \in \mathfrak{L}(W, U)$.

PROOF. (a) $\langle (\mathrm{S} + \mathrm{T})\boldsymbol{v}, \boldsymbol{w}\rangle = \langle \mathrm{S}\boldsymbol{v}, \boldsymbol{w}\rangle + \langle \mathrm{T}\boldsymbol{v}, \boldsymbol{w}\rangle = \langle \boldsymbol{v}, \mathrm{S}^*\boldsymbol{w}\rangle + \langle \boldsymbol{v}, \mathrm{T}^*\boldsymbol{w}\rangle = \langle \boldsymbol{v}, (\mathrm{S}^* + \mathrm{T}^*)\boldsymbol{w}\rangle$.

(b) $\langle (a\mathrm{T})\boldsymbol{v}, \boldsymbol{w}\rangle = a\langle \mathrm{T}\boldsymbol{v}, \boldsymbol{w}\rangle = a\langle \boldsymbol{v}, \mathrm{T}^*\boldsymbol{w}\rangle = \langle \boldsymbol{v}, (\overline{a}\,\mathrm{T}^*)(\boldsymbol{w})\rangle$.

(c) $\langle \mathrm{T}^*\boldsymbol{w}, \boldsymbol{v}\rangle = \overline{\langle \boldsymbol{v}, \mathrm{T}^*\boldsymbol{w}\rangle} = \overline{\langle \mathrm{T}\boldsymbol{v}, \boldsymbol{w}\rangle} = \langle \boldsymbol{w}, \mathrm{T}\boldsymbol{v}\rangle$.

(d) $\langle \mathrm{Id}\boldsymbol{v}, \boldsymbol{w}\rangle = \langle \boldsymbol{v}, \boldsymbol{w}\rangle = \langle \boldsymbol{v}, \mathrm{Id}\boldsymbol{w}\rangle$.

(e) $\langle (\mathrm{ST})\boldsymbol{v}, \boldsymbol{w}\rangle = \langle \mathrm{S}(\mathrm{T}\boldsymbol{v}), \boldsymbol{w}\rangle = \langle \mathrm{T}\boldsymbol{v}, \mathrm{S}^*\boldsymbol{w}\rangle = \langle \boldsymbol{v}, (\mathrm{T}^*\mathrm{S}^*)\boldsymbol{w}\rangle$. □

EXERCISES

► EXERCISE 88 (6.1). *Prove that if $\boldsymbol{x}$, $\boldsymbol{y}$ are nonzero vectors in $\mathbb{R}^2$, then $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \|\boldsymbol{x}\|\,\|\boldsymbol{y}\|\cos\theta$, where θ is the angle between $\boldsymbol{x}$ and $\boldsymbol{y}$.*

PROOF. Using notation as in Figure 6.1, the law of cosines states that

$$\|\boldsymbol{x} - \boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 - 2\|\boldsymbol{x}\|\,\|\boldsymbol{y}\|\cos\theta. \tag{6.1}$$

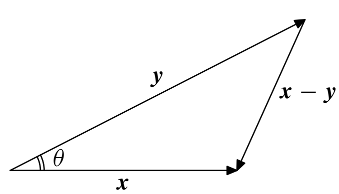


FIGURE 6.1. The law of cosines

After inserting $\|\boldsymbol{x} - \boldsymbol{y}\|^2 = \langle \boldsymbol{x} - \boldsymbol{y}, \boldsymbol{x} - \boldsymbol{y}\rangle = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2 - 2\langle \boldsymbol{x}, \boldsymbol{y}\rangle$ into (6.1), we get the conclusion. □

► EXERCISE 89 (6.2). *Suppose $\boldsymbol{u}, \boldsymbol{v} \in V$. Prove that $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = 0$ if and only if $\|\boldsymbol{u}\| \leqslant \|\boldsymbol{u} + a\boldsymbol{v}\|$ for all $a \in \mathbb{F}$.*

PROOF. If $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$, then $\langle \boldsymbol{u}, a\boldsymbol{v} \rangle = 0$ and so

$$\|\boldsymbol{u} + a\boldsymbol{v}\|^2 = \langle \boldsymbol{u} + a\boldsymbol{v}, \boldsymbol{u} + a\boldsymbol{v} \rangle = \|\boldsymbol{u}\|^2 + \|a\boldsymbol{v}\|^2 \geqslant \|\boldsymbol{u}\|^2.$$

Now suppose that $\|\boldsymbol{u}\| \leqslant \|\boldsymbol{u} + a\boldsymbol{v}\|$ for any $a \in \mathbb{F}$. If $\boldsymbol{v} = \boldsymbol{0}$, then $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$ holds trivially. Thus we assume that $\boldsymbol{v} \neq \boldsymbol{0}$. We first have

$$\begin{aligned}
\|\boldsymbol{u} + a\boldsymbol{v}\|^2 &= \langle \boldsymbol{u} + a\boldsymbol{v}, \boldsymbol{u} + a\boldsymbol{v} \rangle \\
&= \langle \boldsymbol{u}, \boldsymbol{u} + a\boldsymbol{v} \rangle + \langle a\boldsymbol{v}, \boldsymbol{u} + a\boldsymbol{v} \rangle \\
&= \|\boldsymbol{u}\|^2 + \overline{a}\langle \boldsymbol{u}, \boldsymbol{v} \rangle + a\overline{\langle \boldsymbol{u}, \boldsymbol{v} \rangle} + \|a\boldsymbol{v}\|^2 \\
&= \|\boldsymbol{u}\|^2 + \|a\boldsymbol{v}\|^2 + \overline{a}\langle \boldsymbol{u}, \boldsymbol{v} \rangle + \overline{\overline{a}\langle \boldsymbol{u}, \boldsymbol{v} \rangle} \\
&= \|\boldsymbol{u}\|^2 + \|a\boldsymbol{v}\|^2 + 2\mathrm{Re}\left(\overline{a}\langle \boldsymbol{u}, \boldsymbol{v} \rangle\right).
\end{aligned}$$

Therefore, $\|\boldsymbol{u}\| \leqslant \|\boldsymbol{u} + a\boldsymbol{v}\|$ for all $a \in \mathbb{F}$ implies that for all $a \in \mathbb{F}$,

$$2\mathrm{Re}\left(\overline{a}\langle \boldsymbol{u}, \boldsymbol{v} \rangle\right) \geqslant -\|a\boldsymbol{v}\|^2 = -|a|^2 \|\boldsymbol{v}\|^2. \tag{6.2}$$

Take $a = -\alpha\langle \boldsymbol{u}, \boldsymbol{v} \rangle$, with $\alpha > 0$; then (6.2) becomes

$$2\left|\langle \boldsymbol{u}, \boldsymbol{v} \rangle\right|^2 \leqslant \alpha\left|\langle \boldsymbol{u}, \boldsymbol{v} \rangle\right|^2 \|\boldsymbol{v}\|^2. \tag{6.3}$$

Let $\alpha = 1/\|\boldsymbol{v}\|^2$. Then (6.3) becomes

$$2\left|\langle \boldsymbol{u}, \boldsymbol{v} \rangle\right|^2 \leqslant \left|\langle \boldsymbol{u}, \boldsymbol{v} \rangle\right|^2.$$

Hence, $\langle \boldsymbol{u}, \boldsymbol{v} \rangle = 0$. □

► EXERCISE 90 (6.3). *Prove that* $\left(\sum_{j=1}^n a_j b_j\right)^2 \leqslant \left(\sum_{j=1}^n j a_j^2\right)\left(\sum_{j=1}^n b_j^2/j\right)$ *for all* $a_j, b_j \in \mathbb{R}$.

PROOF. Since $a_j, b_j \in \mathbb{R}$, we can write any $\boldsymbol{a} = (a_1, \dots, a_n)$ and $\boldsymbol{b} = (b_1, \dots, b_n)$ as $\boldsymbol{a} = (a_1', a_2'/\sqrt{2}, \dots, a_n'/\sqrt{n})$ and $\boldsymbol{b} = (b_1', \sqrt{2}b_2', \dots, \sqrt{n}b_n')$ for some $\boldsymbol{a}' = (a_1', \dots, a_n')$ and $\boldsymbol{b}' = (b_1', \dots, b_n')$. Then

$$\left(\sum_{j=1}^n a_j b_j\right)^2 = \left(\sum_{j=1}^n a_j' b_j'\right)^2 = \langle \boldsymbol{a}', \boldsymbol{b}' \rangle^2,$$

$$\sum_{j=1}^n j a_j^2 = \sum_{j=1}^n j \frac{a_j'^2}{j} = \sum_{j=1}^n a_j'^2 = \|\boldsymbol{a}'\|^2,$$

and

$$\sum_{j=1}^n \frac{b_j^2}{j} = \sum_{j=1}^n \frac{j b_j'^2}{j} = \sum_{j=1}^n b_j'^2 = \|\boldsymbol{b}'\|^2.$$

Hence, by the Cauchy-Schwarz Inequality,

$$\left(\sum_{j=1}^{n} a_j b_j\right)^2 = \langle \boldsymbol{a}', \boldsymbol{b}' \rangle^2 \leqslant \|\boldsymbol{a}'\|^2 \|\boldsymbol{b}'\|^2 = \left(\sum_{j=1}^{n} j a_j^2\right)\left(\sum_{j=1}^{n} \frac{b_j^2}{j}\right). \qquad \square$$

► EXERCISE 91 (6.4). *Suppose* $\boldsymbol{u}, \boldsymbol{v} \in V$ *are such that* $\|\boldsymbol{u}\| = 3$, $\|\boldsymbol{u}+\boldsymbol{v}\| = 4$, *and* $\|\boldsymbol{u}-\boldsymbol{v}\| = 6$. *What number must* $\|\boldsymbol{v}\|$ *equal?*

SOLUTION. By the parallelogram equality, $\|\boldsymbol{u}+\boldsymbol{v}\|^2 + \|\boldsymbol{u}-\boldsymbol{v}\|^2 = 2(\|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2)$, so we have $\|\boldsymbol{v}\| = \sqrt{17}$. □

► EXERCISE 92 (6.5). *Prove or disprove: there is an inner product on* $\mathbb{R}^2$ *such that the associated norm is given by* $\|(x_1, x_2)\| = |x_1| + |x_2|$ *for all* $(x_1, x_2) \in \mathbb{R}^2$.

PROOF. There is no such inner product on $\mathbb{R}^2$. For example, let $\boldsymbol{u} = (1,0)$ and $\boldsymbol{v} = (0,1)$. Then $\|\boldsymbol{u}\| = \|\boldsymbol{v}\| = 1$ and $\|\boldsymbol{u}+\boldsymbol{v}\| = \|\boldsymbol{u}-\boldsymbol{v}\| = 2$. But then the Parallelogram Equality fails. □

► EXERCISE 93 (6.6). *Prove that if* V *is a real inner-product space, then* $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = (\|\boldsymbol{u}+\boldsymbol{v}\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2)/4$ *for all* $\boldsymbol{u}, \boldsymbol{v} \in V$.

PROOF. If V is a real inner-product space, then for any $\boldsymbol{u}, \boldsymbol{v} \in V$,

$$\begin{aligned}\frac{\|\boldsymbol{u}+\boldsymbol{v}\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2}{4} &= \frac{\langle \boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v}\rangle - \langle \boldsymbol{u}-\boldsymbol{v}, \boldsymbol{u}-\boldsymbol{v}\rangle}{4}\\ &= \frac{\left(\|\boldsymbol{u}\|^2 + 2\langle \boldsymbol{u}, \boldsymbol{v}\rangle + \|\boldsymbol{v}\|^2\right) - \left(\|\boldsymbol{u}\|^2 - 2\langle \boldsymbol{u}, \boldsymbol{v}\rangle + \|\boldsymbol{v}\|^2\right)}{4}\\ &= \langle \boldsymbol{u}, \boldsymbol{v}\rangle. \qquad \square\end{aligned}$$

► EXERCISE 94 (6.7). *Prove that if* V *is a complex inner-product space, then*

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \frac{\|\boldsymbol{u}+\boldsymbol{v}\|^2 - \|\boldsymbol{u}-\boldsymbol{v}\|^2 + \|\boldsymbol{u}+i\boldsymbol{v}\|^2 i - \|\boldsymbol{u}-i\boldsymbol{v}\|^2 i}{4}$$

for all $\boldsymbol{u}, \boldsymbol{v} \in V$.

PROOF. If V is a complex inner-product space, then for any $\boldsymbol{u}, \boldsymbol{v} \in V$ we have

$$\begin{aligned}\|\boldsymbol{u}+\boldsymbol{v}\|^2 = \langle \boldsymbol{u}+\boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v}\rangle &= \langle \boldsymbol{u}, \boldsymbol{u}+\boldsymbol{v}\rangle + \langle \boldsymbol{v}, \boldsymbol{u}+\boldsymbol{v}\rangle\\ &= \|\boldsymbol{u}\|^2 + \langle \boldsymbol{u}, \boldsymbol{v}\rangle + \langle \boldsymbol{v}, \boldsymbol{u}\rangle + \|\boldsymbol{v}\|^2,\end{aligned}$$

$$\begin{aligned}\|\boldsymbol{u}-\boldsymbol{v}\|^2 = \langle \boldsymbol{u}-\boldsymbol{v}, \boldsymbol{u}-\boldsymbol{v}\rangle &= \langle \boldsymbol{u}, \boldsymbol{u}-\boldsymbol{v}\rangle - \langle \boldsymbol{v}, \boldsymbol{u}-\boldsymbol{v}\rangle\\ &= \|\boldsymbol{u}\|^2 - \langle \boldsymbol{u}, \boldsymbol{v}\rangle - \langle \boldsymbol{v}, \boldsymbol{u}\rangle + \|\boldsymbol{v}\|^2,\end{aligned}$$

$$\begin{aligned}\|\boldsymbol{u}+i\boldsymbol{v}\|^2 i &= \langle \boldsymbol{u}+i\boldsymbol{v}, \boldsymbol{u}+i\boldsymbol{v}\rangle\, i = \left(\langle \boldsymbol{u}, \boldsymbol{u}+i\boldsymbol{v}\rangle + \langle i\boldsymbol{v}, \boldsymbol{u}+i\boldsymbol{v}\rangle\right) i\\ &= \left(\|\boldsymbol{u}\|^2 + \bar{i}\langle \boldsymbol{u}, \boldsymbol{v}\rangle + i\,\langle \boldsymbol{v}, \boldsymbol{u}\rangle + i\bar{i}\|\boldsymbol{v}\|^2\right) i\\ &= \|\boldsymbol{u}\|^2 i + \langle \boldsymbol{u}, \boldsymbol{v}\rangle - \langle \boldsymbol{v}, \boldsymbol{u}\rangle + \|\boldsymbol{v}\|^2 i,\end{aligned}$$

and

$$\begin{aligned}\|u - i\,v\|^2\, i &= \langle u - i\,v, u - i\,v\rangle\, i = \big(\langle u, u - i\,v\rangle - \langle i\,v, u - i\,v\rangle\big)\, i \\ &= \left(\|u\|^2 - \bar{i}\,\langle u, v\rangle - i\,\langle v, u\rangle + i\bar{i}\,\|v\|^2\right) i \\ &= \|u\|^2 i - \langle u, v\rangle + \langle v, u\rangle + \|v\|^2 i.\end{aligned}$$

Hence,

$$\begin{aligned}&\frac{\|u + v\|^2 - \|u - v\|^2 + \|u + i\,v\|^2\, i - \|u - i\,v\|^2\, i}{4} \\ &\qquad = \frac{2\langle u, v\rangle + 2\,\langle v, u\rangle + 2\langle u, v\rangle - 2\,\langle v, u\rangle}{4} \\ &\qquad = \langle u, v\rangle.\end{aligned}$$ □

► EXERCISE 95 (6.10). *On $\mathfrak{P}_2(\mathbb{R})$, consider the inner product given by*

$$\langle p, q\rangle = \int_0^1 p\,(x)\, q\,(x)\,\,\mathrm{d}x.$$

Apply the Gram-Schmidt procedure to the basis $(1, x, x^2)$ to produce an orthonormal basis of $\mathfrak{P}_2(\mathbb{R})$.

SOLUTION. It is clear that $e_1 = 1$ since $\|1\|^2 = \int_0^1 (1 \times 1)\,\,\mathrm{d}x = 1$. As for e_2, let

$$e_2 = \frac{x - \langle x, e_1\rangle\, e_1}{\left\|x - \langle x, e_1\rangle\, e_1\right\|}.$$

Since

$$\langle x, e_1\rangle = \int_0^1 x\,\mathrm{d}x = \frac{1}{2};$$

we have

$$e_2 = \frac{x - 1/2}{\|x - 1/2\|} = \frac{x - 1/2}{\sqrt{\int_0^1 \left(x - 1/2\right)^2\,\,\mathrm{d}x}} = \sqrt{3}\,(2x - 1)\,.$$

As for e_3,

$$e_3 = \frac{x^2 - \langle x^2, e_1\rangle e_1 - \langle x^2, e_2\rangle e_2}{\left\|x^2 - \langle x^2, e_1\rangle e_1 - \langle x^2, e_2\rangle e_2\right\|}.$$

Since

$$\begin{aligned}\left\langle x^2, e_1\right\rangle &= \int_0^1 x^2\,\mathrm{d}x = \frac{1}{3}, \\ \left\langle x^2, e_2\right\rangle &= \int_0^1 x^2\left[\sqrt{3}\,(2x - 1)\right]\,\mathrm{d}x = \frac{\sqrt{3}}{6},\end{aligned}$$

and

$$\left\|x^2 - \langle x^2, e_1\rangle e_1 - \langle x^2, e_2\rangle e_2\right\| = \sqrt{\int_0^1 \left(x^2 - x + \frac{1}{6}\right)^2\,\mathrm{d}x} = \frac{\sqrt{1/5}}{6},$$

we know that

$$e_3 = \frac{x^2 - x + 1/6}{\sqrt{1/5}\Big/6} = \sqrt{5}\left(6x^2 - 6x + 1\right). \qquad \square$$

► EXERCISE 96 (6.11). *What happens if the Gram-Schmidt procedure is applied to a list of vectors that is not linearly independent.*

SOLUTION. If $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ is not linearly independent, then

$$\boldsymbol{e}_j = \frac{\boldsymbol{v}_j - \langle \boldsymbol{v}_j, \boldsymbol{e}_j \rangle \boldsymbol{e}_1 - \cdots - \langle \boldsymbol{v}_j, \boldsymbol{e}_{j-1} \rangle \boldsymbol{e}_{j-1}}{\left\| \boldsymbol{v}_j - \langle \boldsymbol{v}_j, \boldsymbol{e}_j \rangle \boldsymbol{e}_1 - \cdots - \langle \boldsymbol{v}_j, \boldsymbol{e}_{j-1} \rangle \boldsymbol{e}_{j-1} \right\|}$$

may not be well defined since if $\boldsymbol{v}_j \in \operatorname{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_{j-1})$, then

$$\left\| \boldsymbol{v}_j - \langle \boldsymbol{v}_j, \boldsymbol{e}_j \rangle \boldsymbol{e}_1 - \cdots - \langle \boldsymbol{v}_j, \boldsymbol{e}_{j-1} \rangle \boldsymbol{e}_{j-1} \right\| = 0. \qquad \square$$

► EXERCISE 97 (6.12). *Suppose V is a real inner-product space and $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)$ is a linearly independent list of vectors in V. Prove that there exist exactly 2^m orthonormal lists $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m)$ of vectors in V such that*

$$\operatorname{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_j) = \operatorname{span}(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_j)$$

for all $j \in \{1, \ldots, m\}$.

PROOF. Given the linearly independent list $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m)$, we have a corresponding orthonormal list $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m)$ by the Gram-Schmidt procedure, such that $\operatorname{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_j) = \operatorname{span}(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_j)$ for all $j \in \{1, \ldots, m\}$.

Now, for every $i = 1, \ldots, m$, the list $(\boldsymbol{e}_1, \ldots, -\boldsymbol{e}_i, \ldots, \boldsymbol{e}_m)$ is also an orthonormal list; further,

$$\operatorname{span}(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_i) = \operatorname{span}(\boldsymbol{e}_1, \ldots, -\boldsymbol{e}_i).$$

The above shows that there are at least 2^m orthonormal lists satisfying the requirement.

On the other hand, if there is an orthonormal list $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_m)$ satisfying

$$\operatorname{span}(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_j) = \operatorname{span}(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_j)$$

for all $j \in \{1, \ldots, m\}$, then $\operatorname{span}(\boldsymbol{v}_1) = \operatorname{span}(\boldsymbol{f}_1)$ implies that

$$\boldsymbol{f}_1 = \pm \frac{\boldsymbol{v}_1}{\|\boldsymbol{v}_1\|} = \pm \boldsymbol{e}_1;$$

Similarly, $\operatorname{span}(\boldsymbol{e}_1, \boldsymbol{e}_2) = \operatorname{span}(\boldsymbol{v}_1, \boldsymbol{v}_2) = \operatorname{span}(\boldsymbol{e}_1, \boldsymbol{f}_2)$ implies that

$$\boldsymbol{f}_2 = a_1 \boldsymbol{e}_1 + a_2 \boldsymbol{e}_2, \quad \text{for some } a_1, a_2 \in \mathbb{R}.$$

Then the orthonormality implies that

$$\langle \boldsymbol{e}_1, a_1 \boldsymbol{e}_1 + a_2 \boldsymbol{e}_2 \rangle = 0 \Longrightarrow a_1 = 0,$$
$$\langle a_2 \boldsymbol{e}_2, a_2 \boldsymbol{e}_2 \rangle = 1 \Longrightarrow a_2 = \pm 1;$$

that is, $\boldsymbol{f}_2 = \pm \boldsymbol{e}_2$. By induction, we can show that $\boldsymbol{f}_i = \pm \boldsymbol{e}_i$ for all $i = 1, \ldots, m$, and this completes the proof. □

► EXERCISE 98 (6.13). *Suppose $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m)$ is an orthonormal list of vectors in V. Let $\boldsymbol{v} \in V$. Prove that $\|\boldsymbol{v}\|^2 = \left|\langle \boldsymbol{v}, \boldsymbol{e}_1 \rangle\right|^2 + \cdots + \left|\langle \boldsymbol{v}, \boldsymbol{e}_m \rangle\right|^2$ if and only if $\boldsymbol{v} \in \operatorname{span}(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m)$.*

PROOF. It follows from Corollary 6.25 that the list $(e_1, \ldots, e_m)$ can be extended to an orthonormal basis $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m, \boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ of V. Then by Theorem 6.17, every vector $\boldsymbol{v} \in V$ can be presented uniquely as $\boldsymbol{v} = \sum_{i=1}^{m} \langle \boldsymbol{v}, \boldsymbol{e}_i \rangle \boldsymbol{e}_i + \sum_{j=1}^{n} \langle \boldsymbol{v}, \boldsymbol{f}_j \rangle \boldsymbol{f}_j$, and so

$$\begin{aligned}
\|\boldsymbol{v}\|^2 &= \left\| \sum_{i=1}^{m} \langle \boldsymbol{v}, \boldsymbol{e}_i \rangle \boldsymbol{e}_i + \sum_{j=1}^{n} \langle \boldsymbol{v}, \boldsymbol{f}_j \rangle \boldsymbol{f}_j \right\|^2 \\
&= \left\langle \sum_{i=1}^{m} \langle \boldsymbol{v}, \boldsymbol{e}_i \rangle \boldsymbol{e}_i + \sum_{j=1}^{n} \langle \boldsymbol{v}, \boldsymbol{f}_j \rangle \boldsymbol{f}_j, \sum_{i=1}^{m} \langle \boldsymbol{v}, \boldsymbol{e}_i \rangle \boldsymbol{e}_i + \sum_{j=1}^{n} \langle \boldsymbol{v}, \boldsymbol{f}_j \rangle \boldsymbol{f}_j \right\rangle \\
&= \sum_{i=1}^{m} \left|\langle \boldsymbol{v}, \boldsymbol{e}_i \rangle\right|^2 + \sum_{j=1}^{n} \left|\langle \boldsymbol{v}, \boldsymbol{f}_j \rangle\right|^2.
\end{aligned}$$

Hence,

$$\begin{aligned}
\|\boldsymbol{v}\|^2 = \left|\langle \boldsymbol{v}, \boldsymbol{e}_1 \rangle\right|^2 + \cdots + \left|\langle \boldsymbol{v}, \boldsymbol{e}_m \rangle\right|^2 &\iff \langle \boldsymbol{v}, \boldsymbol{f}_j \rangle = 0, \quad \forall\, j = 1, \ldots, n \\
&\iff \boldsymbol{v} = \sum_{i=1}^{m} \langle \boldsymbol{v}, \boldsymbol{e}_i \rangle \boldsymbol{e}_j \\
&\iff \boldsymbol{v} \in \operatorname{span}(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m).
\end{aligned}$$

□

► EXERCISE 99 (6.14). *Find an orthonormal basis of $\mathfrak{P}_2(\mathbb{R})$ such that the differentiation operator on $\mathfrak{P}_2(\mathbb{R})$ has an upper-triangular matrix with respect to this basis.*

SOLUTION. Consider the orthonormal basis $(1, \sqrt{3}(2x-1), \sqrt{5}(6x^2 - 6x + 1)) = (\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ in Exercise 95. Let T be the differentiation operator on $\mathfrak{P}_2(\mathbb{R})$. We have

$$\mathrm{T}\boldsymbol{e}_1 = 0 \in \operatorname{span}(\boldsymbol{e}_1),$$

$$\mathrm{T}\boldsymbol{e}_2 = \left[\sqrt{3}(2x-1)\right]' = 2\sqrt{3} \in \operatorname{span}(\boldsymbol{e}_1, \boldsymbol{e}_2),$$

and

$$\mathrm{T}\boldsymbol{e}_3 = \left[\sqrt{5}(6x^2 - 6x + 1)\right]' = 12\sqrt{5}x - 6\sqrt{5} \in \operatorname{span}(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3).$$

It follows from Proposition 5.12 that T has an upper-triangular matrix. □

► EXERCISE 100 (6.15). *Suppose U is a subspace of V. Prove that $\dim U^{\perp} = \dim V - \dim U$.*

PROOF. We have $V = U \oplus U^{\perp}$; hence,

$$\begin{aligned}\dim V &= \dim U + \dim U^{\perp} - \dim U \cap U^{\perp} \\ &= \dim U + \dim U^{\perp};\end{aligned}$$

that is, $\dim U^{\perp} = \dim V - \dim U$. □

► EXERCISE 101 (6.16). *Suppose U is a subspace of V. Prove that $U^{\perp} = \{\mathbf{0}\}$ if and only if $U = V$.*

PROOF. If $U^{\perp} = \{\mathbf{0}\}$, then $V = U \oplus U^{\perp} = U \oplus \{\mathbf{0}\} = U$. To see the converse direction, let $U = V$. For any $\boldsymbol{w} \in U^{\perp}$, we have $\langle \boldsymbol{w}, \boldsymbol{w} \rangle = 0$ since $\boldsymbol{w} \in U^{\perp} \subseteq V = U$; then $\boldsymbol{w} = \mathbf{0}$, that is, $U^{\perp} = \{\mathbf{0}\}$. □

► EXERCISE 102 (6.17). *Prove that if $\mathbf{P} \in \mathcal{L}(V)$ is such that $\mathbf{P}^2 = \mathbf{P}$ and every vector in $\mathcal{N}_{\mathbf{P}}$ is orthogonal to every vector in $\mathcal{R}_{\mathbf{P}}$, then $\mathbf{P}$ is an orthogonal projection.*

PROOF. For every $\boldsymbol{w} \in \mathcal{R}_{\mathbf{P}}$, there exists $\boldsymbol{v}_{\boldsymbol{w}} \in V$ such that $\mathbf{P}\boldsymbol{v}_{\boldsymbol{w}} = \boldsymbol{w}$. Hence,

$$\mathbf{P}\boldsymbol{w} = \mathbf{P}(\mathbf{P}\boldsymbol{v}_{\boldsymbol{w}}) = \mathbf{P}^2\boldsymbol{v}_{\boldsymbol{w}} = \mathbf{P}\boldsymbol{v}_{\boldsymbol{w}} = \boldsymbol{w}.$$

By Exercise 84, $V = \mathcal{N}_{\mathbf{P}} \oplus \mathcal{R}_{\mathbf{P}}$ if $\mathbf{P}^2 = \mathbf{P}$. Then any $\boldsymbol{v} \in V$ can be uniquely written as $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$ with $\boldsymbol{u} \in \mathcal{N}_{\mathbf{P}}$ and $\boldsymbol{w} \in \mathcal{R}_{\mathbf{P}}$, and

$$\mathbf{P}\boldsymbol{v} = \mathbf{P}(\boldsymbol{u} + \boldsymbol{w}) = \mathbf{P}\boldsymbol{w} = \boldsymbol{w},$$

Hence, $\mathbf{P} = \mathbf{P}_{\mathcal{R}_{\mathbf{P}}}$ when $\mathcal{N}_{\mathbf{P}} \perp \mathcal{R}_{\mathbf{P}}$. □

► EXERCISE 103 (6.18). *Prove that if $\mathbf{P} \in \mathcal{L}(V)$ is such that $\mathbf{P}^2 = \mathbf{P}$ and $\|\mathbf{P}\boldsymbol{v}\| \leqslant \|\boldsymbol{v}\|$ for every $\boldsymbol{v} \in V$, then $\mathbf{P}$ is an orthogonal projection.*

PROOF. It follows from the previous exercise that if $\mathbf{P}^2 = \mathbf{P}$, then $\mathbf{P}\boldsymbol{v} = \boldsymbol{w}$ for every $\boldsymbol{v} \in V$, where $\boldsymbol{v}$ is uniquely written as $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$ with $\boldsymbol{u} \in \mathcal{N}_{\mathbf{P}}$ and $\boldsymbol{w} \in \mathcal{R}_{\mathbf{P}}$.

It now suffices to show that $\mathcal{N}_{\mathbf{P}} \perp \mathcal{R}_{\mathbf{P}}$. Take an arbitrary $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w} \in V$, where $\boldsymbol{u} \in \mathcal{N}_{\mathbf{P}}$ and $\boldsymbol{w} \in \mathcal{R}_{\mathbf{P}}$. Then $\|\mathbf{P}\boldsymbol{v}\| \leqslant \|\boldsymbol{v}\|$ implies that

$$\langle \mathbf{P}\boldsymbol{v}, \mathbf{P}\boldsymbol{v} \rangle = \langle \boldsymbol{w}, \boldsymbol{w} \rangle \leqslant \langle \boldsymbol{u} + \boldsymbol{w}, \boldsymbol{u} + \boldsymbol{w} \rangle \iff -\|\boldsymbol{u}\|^2 \leqslant 2\mathrm{Re}\left(\langle \boldsymbol{u}, \boldsymbol{w} \rangle\right).$$

The above inequality certainly fails for some $\boldsymbol{v}$ if $\langle \boldsymbol{u}, \boldsymbol{w} \rangle \neq 0$ (see Exercise 89). Therefore, $\mathcal{N}_{\mathbf{P}} \perp \mathcal{R}_{\mathbf{P}}$ and $\mathbf{P} = \mathbf{P}_{\mathcal{R}_{\mathbf{P}}}$. □

► EXERCISE 104 (6.19). *Suppose $\mathrm{T} \in \mathcal{L}(V)$ and U is a subspace of V. Prove that U is invariant under T if and only if $\mathbf{P}_U \mathrm{T} \mathbf{P}_U = \mathrm{T}\mathbf{P}_U$.*

PROOF. It follows from Theorem 6.29 that $V = U \oplus U^{\perp}$.

Only if: Suppose that U is invariant under T. For any $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$ with $\boldsymbol{u} \in U$ and $\boldsymbol{w} \in U^{\perp}$, we have

$$\left(\mathbf{P}_U \mathrm{T} \mathbf{P}_U\right)(\boldsymbol{v}) = \left(\mathbf{P}_U \mathrm{T}\right)(\boldsymbol{u}) = \mathbf{P}_U\left(\mathrm{T}\boldsymbol{u}\right) = \mathrm{T}\boldsymbol{u},$$

where the last equality holds since $\boldsymbol{u} \in U$ and U is invariant under T. We also have

$$(\mathrm{T}\mathbf{P}_U)(\boldsymbol{v}) = \mathrm{T}\boldsymbol{u}.$$

If: Now suppose that $\mathbf{P}_U \mathrm{T} \mathbf{P}_U = \mathrm{T}\mathbf{P}_U$. Take any $\boldsymbol{u} \in U$ and we have

$$\mathrm{T}\boldsymbol{u} = \mathrm{T}\left(\mathbf{P}_U(\boldsymbol{u})\right) = (\mathrm{T}\mathbf{P}_U)(\boldsymbol{u}) = (\mathbf{P}_U \mathrm{T}\mathbf{P}_U)(\boldsymbol{u}) = \mathbf{P}_U(\mathrm{T}\boldsymbol{u}) \in U$$

by the definition of $\mathbf{P}_U$. This proves that U is invariant under T. □

► EXERCISE 105 (6.20). *Suppose* $\mathrm{T} \in \mathcal{L}(V)$ *and* U *is a subspace of* V. *Prove that* U *and* $U^\perp$ *are both invariant under* T *if and only if* $\mathbf{P}_U \mathrm{T} = \mathrm{T}\mathbf{P}_U$.

PROOF. Suppose first that both U and $U^\perp$ are both invariant under T. Then for any $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$, where $\boldsymbol{u} \in U$ and $\boldsymbol{w} \in U^\perp$, we have

$$(\mathbf{P}_U \mathrm{T})(\boldsymbol{v}) = (\mathbf{P}_U \mathrm{T})(\boldsymbol{u} + \boldsymbol{w}) = \mathbf{P}_U(\mathrm{T}\boldsymbol{u} + \mathrm{T}\boldsymbol{w}) = \mathrm{T}\boldsymbol{u},$$

and $(\mathrm{T}\mathbf{P}_U)(\boldsymbol{v}) = \mathrm{T}\boldsymbol{u}$.

Now suppose $\mathbf{P}_U \mathrm{T} = \mathrm{T}\mathbf{P}_U$. For any $\boldsymbol{u} \in U$, we have $\mathrm{T}\boldsymbol{u} = (\mathrm{T}\mathbf{P}_U)(\boldsymbol{u}) = (\mathbf{P}_U \mathrm{T})(\boldsymbol{u}) = \mathbf{P}_U(\mathrm{T}\boldsymbol{u}) \in U$. Applying the previous argument to $U^\perp$ proves that $U^\perp$ is invariant. □

► EXERCISE 106 (6.21). *In* $\mathbb{R}^4$, *let* $U = \operatorname{span}((1,1,0,0),(1,1,1,2))$. *Find* $\boldsymbol{u} \in U$ *such that* $\|\boldsymbol{u} - (1,2,3,4)\|$ *is as small as possible.*

SOLUTION. We first need to find the orthonormal basis of U. Using the Gram-Schmidt procedure, we have

$$\boldsymbol{e}_1 = \frac{(1,1,0,0)}{\|(1,1,0,0,)\|} = \left(\sqrt{2}/2, \sqrt{2}/2, 0, 0\right),$$

and

$$\boldsymbol{e}_2 = \frac{(1,1,1,2) - \langle (1,1,1,2), \boldsymbol{e}_1 \rangle \boldsymbol{e}_1}{\left\|(1,1,1,2) - \langle (1,1,1,2), \boldsymbol{e}_1 \rangle \boldsymbol{e}_1\right\|} = \left(0, 0, \sqrt{5}/5, 2\sqrt{5}/5\right).$$

Then by 6.35,

$$\mathbf{P}_U(1,2,3,4) = \langle (1,2,3,4), e_1 \rangle e_1 + \langle (1,2,3,4), e_2 \rangle e_2 = \left(3/2, 3/2, 11/5, 22/5\right).$$

REMARK. We can use *Maple* to obtain the orthonormal basis easily:

```
>with(LinearAlgebra):
>v1:=<1,1,0,0>:
>v2:=<1,1,1,2>:
>GramSchmidt({v1,v2}, normalized)
```

□

► EXERCISE 107 (6.22). *Find* $p \in \mathcal{P}_3(\mathbb{R})$ *such that* $p(0) = 0$, $p'(0) = 0$, *and* $\int_0^1 |2 + 3x - p(x)|^2 \, \mathrm{d}x$ *is as small as possible.*

PROOF. $p(0) = p'(0) = 0$ implies that $p(x) = ax^2 + bx^3$, where $a, b \in \mathbb{R}$. We want to find $p \in U \equiv \operatorname{span}(x^2, x^3)$ such that distance from $q = 2_3 x$ to U is as small as possible. With the Gram-Schmidt procedure, the orthonomal basis is

$$e_1 = \frac{x^2}{\|x^2\|} = \frac{x^2}{\sqrt{\int_0^1 \left|x^2 \cdot x^2\right| \, \mathrm{d}x}} = \sqrt{5}x^2,$$

and

$$e_2 = \frac{x^3 - \left(\int_0^1 \left|\sqrt{5}x^5\right| \, \mathrm{d}x\right)\sqrt{5}x^2}{\left\|x^3 - \left(\int_0^1 \left|\sqrt{5}x^5\right| \, \mathrm{d}x\right)\sqrt{5}x^2\right\|} = \frac{x^3 - \frac{5}{6}x^2}{\sqrt{7}/42} = 6\sqrt{7}x^3 - 5\sqrt{7}x^2.$$

Hence,

$$\begin{aligned}\mathbf{P}_U(2+3x) &= \left[\int_0^1 (2+3x)\sqrt{5}x^2 \, \mathrm{d}x\right]\sqrt{5}x^2 \\ &\quad + \left[\int_0^1 (2+3x)\left(6\sqrt{7}x^3 - 5\sqrt{7}x^2\right) \mathrm{d}x\right]\left(6\sqrt{7}x^3 - 5\sqrt{7}x^2\right). \qquad \square\end{aligned}$$

► EXERCISE 108 (6.24). *Find a polynomial* $q \in \mathfrak{P}_2(\mathbb{R})$ *such that*

$$p\left(\frac{1}{2}\right) = \int_0^1 p(x)q(x) \, \mathrm{d}x$$

for every $p \in \mathfrak{P}_2(\mathbb{R})$.

SOLUTION. For every $p \in \mathfrak{P}_2(\mathbb{R})$, we define a function $\mathrm{T}\colon \mathfrak{P}_2(\mathbb{R}) \to \mathbb{R}$ by letting $\mathrm{T}p = p(1/2)$. It is clear that $\mathrm{T} \in \mathfrak{L}(\mathfrak{P}_2(\mathbb{R}), \mathbb{R})$.

It follows from Exercise 95 that $(e_1, e_2, e_3) = (1, \sqrt{3}(2x-1), \sqrt{5}(6x^2 - 6x + 1))$ is an orthonormal basis of $\mathfrak{P}_2(\mathbb{R})$. Then, by Theorem,

$$\begin{aligned}\mathrm{T}p &= \mathrm{T}\left(\langle p, e_1\rangle e_1 + \langle p, e_2\rangle e_2 + \langle p, e_3\rangle e_3\right) \\ &= \langle p, \mathrm{T}(e_1)e_1 + \mathrm{T}(e_2)e_2 + \mathrm{T}(e_3)e_3\rangle;\end{aligned}$$

hence,

$$\begin{aligned}q(x) &= e_1(1/2)e_1 + e_2(1/2)e_2 + e_3(1/2)e_3 \\ &= 1 + 0 - \frac{\sqrt{5}}{2}\left[\sqrt{5}(6x^2 - 6x + 1)\right] \\ &= -\frac{3}{2} + 15x - 15x^2. \qquad \square\end{aligned}$$

► EXERCISE 109 (6.25). *Find a polynomial* $q \in \mathfrak{P}_2(\mathbb{R})$ *such that*

$$\int_0^1 p(x)(\cos \pi x)\,\mathrm{d}x = \int_0^1 p(x)q(x)\,\mathrm{d}x$$

for every $p \in \mathfrak{P}_2(\mathbb{R})$.

SOLUTION. As in the previous exercise, we let $\mathrm{T}\colon p \mapsto \int_0^1 p(x)(\cos \pi x)\,\mathrm{d}x$ for every $p \in \mathfrak{P}_2(\mathbb{R})$. Then $\mathrm{T} \in \mathfrak{L}(\mathfrak{P}_2(\mathbb{R}), \mathbb{R})$. Let

$$q(x) = \mathrm{T}(e_1)e_1 + \mathrm{T}(e_2)e_2 + \mathrm{T}(e_3)e_3 = 12/\pi^2 - 24x/\pi^2. \qquad \square$$

► EXERCISE 110 (6.26). *Fix a vector* $\boldsymbol{v} \in V$ *and define* $\mathrm{T} \in \mathfrak{L}(V, \mathbb{F})$ *by* $\mathrm{T}\boldsymbol{u} = \langle \boldsymbol{u}, \boldsymbol{v} \rangle$. *For* $a \in \mathbb{F}$, *find a formula for* $\mathrm{T}^* a$.

PROOF. Take any $\boldsymbol{u} \in V$. We have $\langle \mathrm{T}\boldsymbol{u}, a \rangle = \big\langle \langle \boldsymbol{u}, \boldsymbol{v} \rangle, a \big\rangle = \langle \boldsymbol{u}, \boldsymbol{v} \rangle a = \langle \boldsymbol{u}, a\boldsymbol{v} \rangle$; thus, $\mathrm{T}^* a = a\boldsymbol{v}$. □

► EXERCISE 111 (6.27). *Suppose* n *is a positive integer. Define* $\mathrm{T} \in \mathfrak{L}(\mathbb{F}^n)$ *by* $\mathrm{T}\,(z_1, \ldots, z_n) = (0, z_1, \ldots, z_{n-1})$. *Find a formula for* $\mathrm{T}^*(z_1, \ldots, z_n)$.

SOLUTION. Take the standard basis of $\mathbb{F}^n$, which is also a orgonormal basis of $\mathbb{F}^n$. We then have

$$\begin{aligned}
\mathrm{T}\,(1, 0, 0, \ldots, 0) &= (0, 1, 0, 0, \ldots, 0)\,, \\
\mathrm{T}\,(0, 1, 0, \ldots, 0) &= (0, 0, 1, 0, \ldots, 0)\,, \\
&\cdots \\
\mathrm{T}\,(0, 0, \ldots, 0, 1) &= (0, 0, 0, 0, \ldots, 0)\,.
\end{aligned}$$

Therefore, $\mathcal{M}\,(\mathrm{T})$ is given by

$$\mathcal{M}\,(\mathrm{T}) = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix},$$

and so

$$\mathcal{M}(\mathrm{T}^*) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Then $\mathrm{T}^*(z_1, \ldots, z_n) = \mathcal{M}(\mathrm{T}^*)\mathcal{M}\,(z_1, \ldots, z_n) = (z_2, z_3, \ldots, z_{n-1}, z_n, 0)$. □

► EXERCISE 112 (6.28). *Suppose* $\mathrm{T} \in \mathfrak{L}(V)$ *and* $\lambda \in \mathbb{F}$. *Prove that* λ *is an eigenvalue of* T *if and only if* $\overline{\lambda}$ *is an eigenvalue of* T^*.

PROOF. If $\lambda \in \mathbb{F}$ is an eigenvalue of T, then there exists $\boldsymbol{v} \neq \mathbf{0}$ such that $\mathrm{T}\boldsymbol{v} = \lambda\boldsymbol{v}$. Take $\boldsymbol{w} \in V$ with $\boldsymbol{w} \neq \mathbf{0}$. Then

$$\langle \mathrm{T}\boldsymbol{v}, \boldsymbol{w} \rangle = \langle \lambda\boldsymbol{v}, \boldsymbol{w} \rangle = \langle \boldsymbol{v}, \overline{\lambda}\boldsymbol{w} \rangle = \langle \boldsymbol{v}, \mathrm{T}^*\boldsymbol{w} \rangle,$$

which implies that $\mathrm{T}^*\boldsymbol{w} = \overline{\lambda}\boldsymbol{w}$; that is, $\overline{\lambda}$ is an eigenvalue of T^*. With the same logic, we can show the inverse direction. □

► EXERCISE 113 (6.29). *Suppose* $\mathrm{T} \in \mathfrak{L}(V)$ *and* U *is a subspace of* V*. Prove that* U *is invariant under* T *if and only if* $U^{\perp}$ *is invariant under* T^**.*

PROOF. Take any $\boldsymbol{u} \in U$ and $\boldsymbol{w} \in U^{\perp}$. If U is invariant under T, then $\mathrm{T}\boldsymbol{u} \in U$ and so

$$0 = \langle \mathrm{T}\boldsymbol{u}, \boldsymbol{w} \rangle = \left\langle \boldsymbol{u}, \mathrm{T}^*\boldsymbol{w} \right\rangle;$$

that is, $\mathrm{T}^*\boldsymbol{w} \in U^{\perp}$. Applying T^* then we obtain the inverse direction. □

► EXERCISE 114 (6.30). *Suppose* $\mathrm{T} \in \mathfrak{L}(V, W)$*. Prove that*

a. T *is injective if and only if* T^* *is surjective;*

b. T *is surjective if and only if* T^* *is injective.*

PROOF. (a) If T is injective, then $\dim \mathcal{N}_{\mathrm{T}} = 0$. Then

$$\dim \mathcal{R}_{\mathrm{T}^*} = \dim \mathcal{R}_{\mathrm{T}} = \dim V - \dim \mathcal{N}_{\mathrm{T}} = \dim V,$$

i.e., $\mathrm{T} \in \mathfrak{L}(W, V)$ is surjective. If T^* is surjective, then $\dim \mathcal{R}_{\mathrm{T}^*} = \dim V$ and so

$$\dim \mathcal{N}_{\mathrm{T}} = \dim V - \dim \mathcal{R}_{\mathrm{T}} = \dim V - \dim \mathcal{R}_{\mathrm{T}^*} = 0,$$

that is, $\mathrm{T} \in \mathfrak{L}(V, W)$ is injective.

(b) Using the fact that $(\mathrm{T}^*)^* = \mathrm{T}$ and the result in part (a) we get (b) immediately. □

► EXERCISE 115 (6.31). *Prove that* $\dim \mathcal{N}_{\mathrm{T}^*} = \dim \mathcal{N}_{\mathrm{T}} + \dim W - \dim V$ *and* $\dim \mathcal{R}_{\mathrm{T}^*} = \dim \mathcal{R}_{\mathrm{T}}$ *for every* $\mathrm{T} \in \mathfrak{L}(V, W)$*.*

PROOF. It follows from Proposition 6.46 that $\mathcal{N}_{\mathrm{T}^*} = (\mathcal{R}_{\mathrm{T}})^{\perp}$. Since $\mathcal{R}_{\mathrm{T}}$ is a subspace of W, and $W = \mathcal{R}_{\mathrm{T}} \oplus (\mathcal{R}_{\mathrm{T}})^{\perp}$, we thus have

$$\begin{aligned} \dim V &= \dim \mathcal{N}_{\mathrm{T}} + \dim \mathcal{R}_{\mathrm{T}} \\ &= \dim \mathcal{N}_{\mathrm{T}} + \dim W - \dim \mathcal{R}_{\mathrm{T}}^{\perp} \\ &= \dim \mathcal{N}_{\mathrm{T}} + \dim W - \dim \mathcal{N}_{\mathrm{T}^*}, \end{aligned} \tag{6.4}$$

which proves the first claim. As for the second equality, we first have

$$\begin{aligned} \dim \mathcal{R}_{\mathrm{T}} &= \dim V - \dim \mathcal{N}_{\mathrm{T}}, \\ \dim \mathcal{R}_{\mathrm{T}^*} &= \dim W - \dim \mathcal{N}_{\mathrm{T}^*}. \end{aligned}$$

Thus, $\dim \mathcal{R}_{\mathrm{T}} - \dim \mathcal{R}_{\mathrm{T}^*} = 0$ by (6.4), that is, $\dim \mathcal{R}_{\mathrm{T}} = \dim \mathcal{R}_{\mathrm{T}^*}$. □

► EXERCISE 116 (6.32). *Suppose* $\mathbf{A}$ *is an* $m \times n$ *matrix of real numbers. Prove that the dimension of the span of the columns of* $\mathbf{A}$ *(in* $\mathbb{R}^m$*) equals the dimension of the span of the rows of* $\mathbf{A}$ *(in* $\mathbb{R}^n$*).*

PROOF. Without loss of generality, we can assume that $\mathrm{T} \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ is the linear map induced by $\mathbf{A}$, where $\mathbf{A}$ corresponds to an orthonormal basis of $\mathbb{R}^n$ and an orthonormal basis of $\mathbb{R}^m$; that is, $\mathrm{T}\boldsymbol{x} = \mathbf{A}\boldsymbol{x}$ for all $\boldsymbol{x} \in \mathbb{R}^n$. By Proposition 6.47, we know that for any $\boldsymbol{y} \in \mathbb{R}^m$,

$$\mathrm{T}^* \boldsymbol{y} = \mathbf{A}' \boldsymbol{y},$$

where $\mathbf{A}'$ is the (conjugate) transpose of $\mathbf{A}$. Let

$$\mathbf{A} = \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{b}_1 \\ \vdots \\ \boldsymbol{b}_m \end{pmatrix}.$$

Then

$$\mathbf{A}' = \begin{pmatrix} \boldsymbol{a}_1' \\ \vdots \\ \boldsymbol{a}_n' \end{pmatrix} = \begin{pmatrix} \boldsymbol{b}_1' & \cdots & \boldsymbol{b}_m' \end{pmatrix}.$$

It is easy to see that

$$\operatorname{span}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n) = \mathcal{R}_{\mathrm{T}}, \quad \text{and} \quad \operatorname{span}(\boldsymbol{a}_1', \ldots, \boldsymbol{a}_n') = \mathcal{R}_{\mathrm{T}^*}.$$

It follows from Exercise 115 that $\dim \mathcal{R}_{\mathrm{T}} = \dim \mathcal{R}_{\mathrm{T}^*}$. □

7

OPERATORS ON INNER-PRODUCT SPACES

"As You Should Verify"

Remark (p.131). If T is normal, then $\mathrm{T}-\lambda\mathrm{Id}$ is normal, too.

Proof. Note that $(\mathrm{T}-\lambda\mathrm{Id})^* = \mathrm{T}^* - \overline{\lambda}\mathrm{Id}$. For any $\boldsymbol{v} \in V$,

$$\begin{aligned}(\mathrm{T}-\lambda\mathrm{Id})(\mathrm{T}^*-\overline{\lambda}\mathrm{Id})\boldsymbol{v} &= (\mathrm{T}-\lambda\mathrm{Id})(\mathrm{T}^*\boldsymbol{v}-\overline{\lambda}\boldsymbol{v}) \\ &= \mathrm{T}(\mathrm{T}^*\boldsymbol{v}-\overline{\lambda}\boldsymbol{v}) - \lambda\cdot(\mathrm{T}^*\boldsymbol{v}-\overline{\lambda}\boldsymbol{v}) \\ &= \mathrm{T}\mathrm{T}^*\boldsymbol{v} - \overline{\lambda}\mathrm{T}\boldsymbol{v} - \lambda\mathrm{T}^*\boldsymbol{v} + |\lambda|^2\,\boldsymbol{v},\end{aligned}$$

and

$$\begin{aligned}(\mathrm{T}^*-\overline{\lambda}\mathrm{Id})(\mathrm{T}-\lambda\mathrm{Id})\boldsymbol{v} &= (\mathrm{T}^*-\overline{\lambda}\mathrm{Id})\,(\mathrm{T}\boldsymbol{v}-\lambda\boldsymbol{v}) \\ &= \mathrm{T}^*\,(\mathrm{T}\boldsymbol{v}-\lambda\boldsymbol{v}) - \overline{\lambda}\cdot(\mathrm{T}\boldsymbol{v}-\lambda\boldsymbol{v}) \\ &= \mathrm{T}^*\mathrm{T}\boldsymbol{v} - \lambda\mathrm{T}^*\boldsymbol{v} - \overline{\lambda}\mathrm{T}\boldsymbol{v} + |\lambda|^2\,\boldsymbol{v}.\end{aligned}$$

Hence, $(\mathrm{T}-\lambda\mathrm{Id})(\mathrm{T}^*-\overline{\lambda}\mathrm{Id}) = (\mathrm{T}^*-\overline{\lambda}\mathrm{Id})(\mathrm{T}-\lambda\mathrm{Id})$ since $\mathrm{T}\mathrm{T}^* = \mathrm{T}^*\mathrm{T}$. □

Exercises

► Exercise 117 (7.1). *Make $\mathfrak{P}_2(\mathbb{R})$ into an inner-product space by defining $\langle p,q\rangle = \int_0^1 p(x)\,q(x)\,\mathrm{d}x$. Define $\mathrm{T} \in \mathfrak{L}(\mathfrak{P}_2(\mathbb{R}))$ by $\mathrm{T}\left(a_0 + a_1x + a_2x^2\right) = a_1x$.*

a. *Show that* T *is not self-adjoint.*

b. *The matrix of* T *with respect to the basis $\left(1, x, x^2\right)$ is*

$$\begin{pmatrix}0 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}.$$

This matrix equals its conjugate transpose, even though T *is not self-adjoint. Explain why this is not a contradiction.*

PROOF. (a) Suppose T is self-adjoint, that is, $T = T^*$. Take any $p, q \in \mathfrak{P}_2(\mathbb{R})$ with $p(x) = a_0 + a_1 x + a_2 x^2$ and $q(x) = b_0 + b_1 x + b_2 x^2$. Then $\langle Tp, q\rangle = \langle p, T^* q\rangle = \langle p, Tq\rangle$ implies that

$$\int_0^1 (a_1 x)\left(b_0 + b_1 x + b_2 x^2\right) \mathrm{d}x = \int_0^1 \left(a_0 + a_1 x + a_2 x^2\right)(b_1 x)\, \mathrm{d}x,$$

that is,

$$\frac{a_1 b_0}{2} + \frac{a_1 b_1}{3} + \frac{a_1 b_2}{4} = \frac{a_0 b_1}{2} + \frac{a_1 b_1}{3} + \frac{a_2 b_1}{4}. \tag{7.1}$$

Let $a_1 = 0$, then (7.1) becomes $0 = a_0 b_1/2$, which fails to hold for any $a_0 b_1 \neq 0$. Therefore, $T \neq T^*$.

(b) $\left(1, x, x^2\right)$ is not an orthonormal basis. See Proposition 6.47. □

► EXERCISE 118 (7.2). *Prove or give a counterexample: the product of any two self-adjoint operators on a finite-dimensional inner-product space is self-adjoint.*

PROOF. The claim is incorrect. Let $S, T \in \mathfrak{L}(V)$ be two self-adjoint operators. Then $(ST)^* = T^* S^* = TS$. It is not necessarily that $ST = TS$ since multiplication is not commutable.

For example, let $S, T \in \mathfrak{L}(\mathbb{R}^2)$ be defined by the following matrices (with respect to the stand basis of $\mathbb{R}^2$):

$$\mathcal{M}(S) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \mathcal{M}(T) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

Then both S and T are self-adjoint, but ST is not since $\mathcal{M}(S)\mathcal{M}(T) \neq \mathcal{M}(T)\mathcal{M}(S)$. □

► EXERCISE 119 (7.3). a. *Show that if V is a real inner-product space, then the set of self-adjoint operators on V is a subspace of $\mathfrak{L}(V)$.*

b. *Show that if V is a complex inner-product space, then the set of self-adjoint operators on V is not a subspace of $\mathfrak{L}(V)$.*

PROOF. (a) Let $\mathfrak{L}^{sa}(V)$ be the set of self-adjoint operators. Obviously, $0 = 0^*$ since for any $\boldsymbol{v}, \boldsymbol{w}$ we have $0 = \langle 0\boldsymbol{v}, \boldsymbol{w}\rangle = \langle \boldsymbol{v}, 0\boldsymbol{w}\rangle = \langle \boldsymbol{v}, 0^*\boldsymbol{w}\rangle$. To see $\mathfrak{L}^{sa}(V)$ is closed under addition, let $S, T \in \mathfrak{L}^{sa}(V)$. Then $(S + T)^* = S^* + T^* = S + T$ implies that $S + T \in \mathfrak{L}^{sa}(V)$. Finally, for any $a \in \mathbb{F}$ and $T \in \mathfrak{L}^{sa}(V)$, we have $(aT)^* = aT^* = aT \in \mathfrak{L}^{sa}(V)$.

(b) If V is a complex inner-product, then $(aT)^* = \overline{a}T^* = \overline{a}T$, so $\mathfrak{L}^{sa}(V)$ is not a subspace of $\mathfrak{L}(V)$. □

► EXERCISE 120 (7.4). *Suppose $\mathbf{P} \in \mathfrak{L}(V)$ is such that $\mathbf{P}^2 = \mathbf{P}$. Prove that $\mathbf{P}$ is an orthogonal projection if and only if $\mathbf{P}$ is self-adjoint.*

PROOF. If $\mathbf{P}^2 = \mathbf{P}$, then $V = \mathcal{N}_{\mathbf{P}} \oplus \mathcal{R}_{\mathbf{P}}$ (by Exercise 84), and $\mathbf{P}\boldsymbol{w} = \boldsymbol{w}$ for every $\boldsymbol{w} \in \mathcal{R}_{\mathbf{P}}$ (by Exercise 102).

Suppose first that $\mathbf{P} = \mathbf{P}^*$. Take arbitrary $\boldsymbol{u} \in \mathcal{N}_{\mathbf{P}}$ and $\boldsymbol{w} \in \mathcal{R}_{\mathbf{P}}$. Then

$$\langle \boldsymbol{u}, \boldsymbol{w} \rangle = \langle \boldsymbol{u}, \mathbf{P}\boldsymbol{w} \rangle = \left\langle \boldsymbol{u}, \mathbf{P}^*\boldsymbol{w} \right\rangle = \langle \mathbf{P}\boldsymbol{u}, \boldsymbol{w} \rangle = \langle \mathbf{0}, \boldsymbol{w} \rangle = 0.$$

Hence, $\mathcal{N}_{\mathbf{P}} \perp \mathcal{R}_{\mathbf{P}}$ and so $\mathbf{P} = \mathbf{P}_{\mathcal{R}_{\mathbf{P}}}$.

Now suppose that $\mathbf{P}$ is an orthogonal projection. Then there exists a subspace U of V such that $V = U \oplus U^{\perp}$ and $\mathbf{P}\boldsymbol{v} = \boldsymbol{u}$ if $\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{w}$ with $\boldsymbol{u} \in U$ and $\boldsymbol{w} \in U^{\perp}$. Take arbitrary $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$ with $\boldsymbol{v}_1 = \boldsymbol{u}_1 + \boldsymbol{w}_1$ and $\boldsymbol{v}_2 = \boldsymbol{u}_2 + \boldsymbol{w}_2$. Then $\langle \mathbf{P}\boldsymbol{v}_1, \boldsymbol{v}_2 \rangle = \langle \boldsymbol{u}_1, \boldsymbol{u}_2 + \boldsymbol{w}_2 \rangle = \langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle$. Similarly, $\langle \boldsymbol{v}_1, \mathbf{P}\boldsymbol{v}_2 \rangle = \langle \boldsymbol{u}_1 + \boldsymbol{w}_1, \boldsymbol{u}_2 \rangle = \langle \boldsymbol{u}_1, \boldsymbol{u}_2 \rangle$. Thus, $\mathbf{P} = \mathbf{P}^*$. □

► EXERCISE 121 (7.5). *Show that if* $\dim V \geqslant 2$, *then the set of normal operators on* V *is not a subspace of* $\mathcal{L}(V)$.

PROOF. Let $\mathcal{L}^{\mathrm{n}}(V)$ denote the set of normal operators on V and $\dim V \geqslant 2$. Let $\mathrm{S}, \mathrm{T} \in \mathcal{L}^{\mathrm{n}}(V)$. It is easy to see that

$$\begin{aligned}(\mathrm{S}+\mathrm{T})(\mathrm{S}+\mathrm{T})^* &= (\mathrm{S}+\mathrm{T})\left(\mathrm{S}^*+\mathrm{T}^*\right) \\ &\neq \left(\mathrm{S}^*+\mathrm{T}^*\right)(\mathrm{S}+\mathrm{T})\end{aligned}$$

generally since matrix multiplication is not commutable. □

► EXERCISE 122 (7.6). *Prove that if* $\mathrm{T} \in \mathcal{L}(V)$ *is normal, then* $\mathcal{R}_{\mathrm{T}} = \mathcal{R}_{\mathrm{T}^*}$.

PROOF. $\mathrm{T} \in \mathcal{L}(V)$ is normal if and only if $\|\mathrm{T}\boldsymbol{v}\| = \|\mathrm{T}^*\boldsymbol{v}\|$ for all $\boldsymbol{v} \in V$ (by Proposition 7.6). Then $\boldsymbol{v} \in \mathcal{N}_{\mathrm{T}} \iff \|\mathrm{T}\boldsymbol{v}\| = 0 \iff \|\mathrm{T}^*\boldsymbol{v}\| = 0 \iff \boldsymbol{v} \in \mathcal{N}_{\mathrm{T}^*}$, i.e., $\mathcal{N}_{\mathrm{T}} = \mathcal{N}_{\mathrm{T}^*}$. It follows from Proposition 6.46 that

$$\mathcal{R}_{\mathrm{T}^*} = \mathcal{N}_{\mathrm{T}}^{\perp} = \mathcal{N}_{\mathrm{T}^*}^{\perp} = \mathcal{R}_{\mathrm{T}}.$$ □

► EXERCISE 123 (7.7). *Prove that if* $\mathrm{T} \in \mathcal{L}(V)$ *is normal, then* $\mathcal{N}_{\mathrm{T}^k} = \mathcal{N}_{\mathrm{T}}$ *and* $\mathcal{R}_{\mathrm{T}^k} = \mathcal{R}_{\mathrm{T}}$ *for every positive integer* k.

PROOF. It is evident that $\mathcal{N}_{\mathrm{T}} \subseteq \mathcal{N}_{\mathrm{T}^k}$. So we take any $\boldsymbol{v} \in \mathcal{N}_{\mathrm{T}^k}$ with $\boldsymbol{v} \neq \mathbf{0}$ (if $\mathcal{N}_{\mathrm{T}^k} = \{\mathbf{0}\}$, there is nothing to prove). Then

$$\begin{aligned}\left\langle \mathrm{T}^*\mathrm{T}^{k-1}\boldsymbol{v}, \mathrm{T}^*\mathrm{T}^{k-1}\boldsymbol{v} \right\rangle = \left\langle \mathrm{T}\mathrm{T}^*\mathrm{T}^{k-1}\boldsymbol{v}, \mathrm{T}^{k-1}\boldsymbol{v} \right\rangle &= \left\langle \mathrm{T}^*\mathrm{T}\mathrm{T}^{k-1}\boldsymbol{v}, \mathrm{T}^{k-1}\boldsymbol{v} \right\rangle \\ &= \left\langle \mathrm{T}^*\mathrm{T}^{k}\boldsymbol{v}, \mathrm{T}^{k-1}\boldsymbol{v} \right\rangle \\ &= 0,\end{aligned}$$

and so $\left(\mathrm{T}^*\mathrm{T}^{k-1}\right)\boldsymbol{v} = 0$. Now

$$\left\langle \mathrm{T}^{k-1}\boldsymbol{v}, \mathrm{T}^{k-1}\boldsymbol{v} \right\rangle = \left\langle \mathrm{T}^{k-2}\boldsymbol{v}, \mathrm{T}^*\mathrm{T}^{k-1}\boldsymbol{v} \right\rangle = 0$$

implies that $\mathrm{T}^{k-1}\boldsymbol{v} = 0$, that is, $\boldsymbol{v} \in \mathcal{N}_{\mathrm{T}^{k-1}}$. With the same logic, we can show that $\boldsymbol{v} \in \mathcal{N}_{\mathrm{T}^{k-2}}, \ldots, \boldsymbol{v} \in \mathcal{N}_{\mathrm{T}}$. □

► EXERCISE 124 (7.8). *Prove that there does not exist a self-adjoint operator* $T \in \mathcal{L}(\mathbb{R}^3)$ *such that* $T(1,2,3) = (0,0,0)$ *and* $T(2,5,7) = (2,5,7)$.

PROOF. Suppose there exists such a operator $T \in \mathcal{L}(\mathbb{R}^3)$. Then

$$\langle T(1,2,3), (2,5,7)\rangle = \langle (0,0,0), (2,5,7)\rangle = 0,$$

but

$$\langle (1,2,3), T(2,5,7)\rangle = \langle (1,2,3), (2,5,7)\rangle \neq 0. \qquad \square$$

► EXERCISE 125 (7.9). *Prove that a normal operator on a complex inner-product space is self-adjoint if and only if all its eigenvalues are real.*

PROOF. It follows from Proposition 7.1 that every eigenvalue of a self-adjoint operator is real, so the "only if" part is clear.

To see the "if" part, let $T \in \mathcal{L}(V)$ be a normal operator, and all its eigenvalues be real. Then by the Complex Spectral Theorem, V has an orthonormal basis consisting of eigenvectors of T. Hence, $\mathcal{M}(T)$ is diagonal with respect this basis, and so the conjugate transpose of $\mathcal{M}(T)$ equals to $\mathcal{M}(T)$ since all eigenvalues are real. $\square$

► EXERCISE 126 (7.10). *Suppose* V *is a complex inner-product space and* $T \in \mathcal{L}(V)$ *is a normal operator such that* $T^9 = T^8$. *Prove that* T *is self-adjoint and* $T^2 = T$.

PROOF. Let $T \in \mathcal{L}(V)$ be normal and $v \in V$. Then by Exercise 123,

$$T^8(Tv - v) = 0 \Longrightarrow Tv - v \in \mathcal{N}_{T^8} = \mathcal{N}_T \Longrightarrow T(Tv - v) = 0 \Longrightarrow T^2 = T.$$

By the Complex Spectral Theorem, there exists an orthonormal basis of V such that $\mathcal{M}(T)$ is diagonal, and the entries on the diagonal line consists of the eigenvalues $(\lambda_1, \ldots, \lambda_n)$ of T. Now $T^2 = T$ implies that $\mathcal{M}(T)\mathcal{M}(T) = \mathcal{M}(T)$; that is,

$$\lambda_i^2 = \lambda_i, \quad i = 1, \ldots, n.$$

Then each $\lambda_i \in \mathbb{R}$. It follows from Exercise 125 that T is self-adjoint. $\square$

► EXERCISE 127 (7.11). *Suppose* V *is a complex inner-product space. Prove that every normal operator on* V *has a square root.*

PROOF. By the Complex Spectral Theorem, there exists an orthonormal basis of V such that $\mathcal{M}(T)$ is diagonal, and the entries on the diagonal line consists of the eigenvalues $(\lambda_1, \ldots, \lambda_n)$ of T. Let $S \in \mathcal{L}(V)$ be an operator whose matrix is

$$\mathcal{M}(\mathrm{S}) = \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_n} \end{pmatrix}.$$

Then $\mathrm{S}^2 = \mathrm{T}$; that is, S is a square root of T. □

► EXERCISE 128 (7.12). *Give an example of a real inner-product space V and $\mathrm{T} \in \mathfrak{L}(V)$ and real numbers α, β with $\alpha^2 < 4\beta$ such that $\mathrm{T}^2 + \alpha\mathrm{T} + \beta\mathrm{Id}$ is not invertible.*

PROOF. We use a normal, but not self-adjoint operator on V (See Lemma 7.15). Let

$$\mathcal{M}(\mathrm{T}) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Then

$$\mathcal{M}(\mathrm{T}^2) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

If we let $\alpha = 0$ and $\beta = 1$, then

$$\left(\mathrm{T}^2 + \mathrm{Id}\right)(x, y) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

for all $(x, y) \in \mathbb{R}^2$. Thus, $\mathrm{T}^2 + \mathrm{Id}$ is not injective, and so is not invertible. □

► EXERCISE 129 (7.13). *Prove or give a counterexample: every self-adjoint operator on V has a cube root.*

PROOF. By the Spectral Theorem, for any self-adjoint operator on V there is a orthonormal basis $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ such that

$$\mathcal{M}(\mathrm{T}) = \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix},$$

where there may some i with $\lambda_i = 0$. Then it is clear that there exists a matrix $\mathcal{M}(\mathrm{S})$ with

$$\mathcal{M}(\mathrm{S}) = \begin{pmatrix} \sqrt[3]{\lambda_1} & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \sqrt[3]{\lambda_n} \end{pmatrix}$$

such that $[\mathcal{M}(\mathrm{S})]^3 = \mathcal{M}(\mathrm{T})$. Let S be the operator with the matrix $\mathcal{M}(\mathrm{S})$ and so S is the cube root of T. □

► EXERCISE 130 (7.14). *Suppose* $\mathrm{T} \in \mathfrak{L}(V)$ *is self-adjoint,* $\lambda \in \mathbb{F}$*, and* $\varepsilon > 0$*. Prove that if there exists* $\boldsymbol{v} \in V$ *such that* $\|\boldsymbol{v}\| = 1$ *and* $\|\mathrm{T}\boldsymbol{v} - \lambda\boldsymbol{v}\| < \varepsilon$*, then* T *has an eigenvalue* λ' *such that* $|\lambda - \lambda'| < \varepsilon$*.*

PROOF. By the Spectral Theorem, there exists an orthonormal basis $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ consisting of eigenvectors of T. Write $\boldsymbol{v} = \sum_{i=1}^n a_i \boldsymbol{e}_i$, where $a_i \in \mathbb{F}$. Since $\|\boldsymbol{v}\| = 1$, we have

$$1 = \left\| \sum_{i=1}^n a_i \boldsymbol{e}_i \right\|^2 = |a_1|^2 + \cdots + |a_n|^2 .$$

Suppose that $|\lambda - \lambda_i| \geqslant \varepsilon$ for all eigenvalues $\lambda_i \in \mathbb{F}$. Then

$$\begin{aligned} \|\mathrm{T}\boldsymbol{v} - \lambda\boldsymbol{v}\|^2 = \left\| \mathrm{T}\left(\sum_{i=1}^n a_i \boldsymbol{e}_i \right) - \lambda \sum_{i=1}^n a_i \boldsymbol{e}_i \right\|^2 &= \left\| \sum_{i=1}^n a_i \lambda_i \boldsymbol{e}_i - \sum_{i=1}^n a_i \lambda \boldsymbol{e}_i \right\|^2 \\ &= \left\| \sum_{i=1}^n a_i \left(\lambda_i - \lambda\right) \boldsymbol{e}_i \right\|^2 \\ &= \sum_{i=1}^n |a_i|^2 \cdot |\lambda_i - \lambda|^2 \\ &\geqslant \sum_{i=1}^n |a_i|^2 \cdot \varepsilon^2 \\ &= \varepsilon^2, \end{aligned}$$

that is, $\|\mathrm{T}\boldsymbol{v} - \lambda\boldsymbol{v}\| \geqslant \varepsilon$. A contradiction. Thus, there exists some eigenvalue λ' so that $|\lambda - \lambda'| < \varepsilon$. □

► EXERCISE 131 (7.15). *Suppose* U *is a finite-dimensional real vector space and* $\mathrm{T} \in \mathfrak{L}(U)$*. Prove that* U *has a basis consisting of eigenvectors of* T *if and only if there is an inner product on* U *that makes* T *into a self-adjoint operator.*

PROOF. Suppose first that U has a basis consisting of eigenvectors $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ of T. Let the corresponding eigenvalues be $(\lambda_1, \dots, \lambda_n)$. Then

$$\mathcal{M}(\mathrm{T}) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

Define $\langle \cdot, \cdot \rangle : U \times U \to \mathbb{R}$ by letting

$$\langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle = \delta_{ij} .$$

Then, for arbitrary $\boldsymbol{u}, \boldsymbol{w} \in U$,

$$\begin{aligned}\langle \mathrm{T}\boldsymbol{u}, \boldsymbol{w}\rangle = \left\langle \sum_{i=1}^{n} a_i \mathrm{T}\boldsymbol{e}_i, \sum_{i=1}^{n} b_i \boldsymbol{e}_i \right\rangle &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i b_j \left\langle \mathrm{T}\boldsymbol{e}_i, \boldsymbol{e}_j \right\rangle \\ &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i \lambda_i b_j \left\langle \boldsymbol{e}_i, \boldsymbol{e}_j \right\rangle \\ &= \sum_{i=1}^{n} a_i \lambda_i b_i.\end{aligned}$$

Similarly, $\langle \boldsymbol{u}, \mathrm{T}\boldsymbol{w}\rangle = \sum_{i=1}^{n} a_i \lambda_i b_i$. Hence $\mathrm{T} = \mathrm{T}^*$.

The other direction follows from the Real Spectral Theorem directly. □

► Exercise 132 (7.16). *Give an example of an operator* T *on an inner-product space such that* T *has an invariant subspace whose orthogonal complement is not invariant under* T.

Solution. Let $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_m)$ be an orthonormal basis of U. Extend to an orthonormal basis $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_m, \boldsymbol{f}_1, \dots, \boldsymbol{f}_n)$ of V. Let U is invariant under T, but $U^\perp$ is not invariant under T. Then $\mathcal{M}(\mathrm{T})$ takes the following form

$$\mathcal{M}(\mathrm{T}) = \begin{array}{c} \\ \boldsymbol{e}_1 \\ \vdots \\ \boldsymbol{e}_m \\ \boldsymbol{f}_1 \\ \vdots \\ \boldsymbol{f}_n \end{array} \begin{array}{c} \begin{array}{cccccc} \boldsymbol{e}_1 & \cdots & \boldsymbol{e}_m & \boldsymbol{f}_1 & \cdots & \boldsymbol{f}_n \end{array} \\ \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & \mathbf{C} \end{pmatrix} \end{array}.$$

Since $(\boldsymbol{f}_1, \dots, \boldsymbol{f}_n)$ is a orthonormal basis of $U^\perp$, we know that $U^\perp$ is not invariant if $\mathbf{C} \neq \mathbf{0}$.

For example, let $V = \mathbb{R}^2$, U be the x-axis, and $U^\perp$ be the y-axis. Let $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ be the standard basis of $\mathbb{R}^2$. Let

$$\mathcal{M}(\mathrm{T}) = \begin{array}{c} \\ \boldsymbol{e}_1 \\ \boldsymbol{e}_2 \end{array} \begin{array}{c} \begin{array}{cc} \boldsymbol{e}_1 & \boldsymbol{e}_2 \end{array} \\ \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \end{array}$$

Notice that T is not normal:

$$\mathcal{M}(\mathrm{T})\mathcal{M}(\mathrm{T}^*) = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}, \quad \text{but} \quad \mathcal{M}(\mathrm{T}^*)\mathcal{M}(\mathrm{T}) = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$ □

► Exercise 133 (7.17). *Prove that the sum of any two positive operators on* V *is positive.*

Proof. Let $\mathrm{S}, \mathrm{T} \in \mathcal{L}(V)$ be positive. Then

$$(\mathrm{S} + \mathrm{T})^* = \mathrm{S}^* + \mathrm{T}^* = \mathrm{S} + \mathrm{T};$$

that is, $\mathrm{S}+\mathrm{T}$ is self-adjoint. Also, for an arbitrary $\boldsymbol{v} \in V$,

$$\langle(\mathrm{S}+\mathrm{T})\boldsymbol{v}, \boldsymbol{v}\rangle = \langle\mathrm{S}\boldsymbol{v}, \boldsymbol{v}\rangle + \langle\mathrm{T}\boldsymbol{v}, \boldsymbol{v}\rangle \geqslant 0.$$

Hence, $\mathrm{S}+\mathrm{T}$ is positive. □

► EXERCISE 134 (7.18). *Prove that if* $\mathrm{T} \in \mathfrak{L}(V)$ *is positive, then so is* T^k *for every positive integer* k*.*

PROOF. It is evident that T^k is self-adjoint. Pick an arbitrary $\boldsymbol{v} \in V$. If $k=2$, then $\left\langle\mathrm{T}^2\boldsymbol{v}, \boldsymbol{v}\right\rangle = \langle\mathrm{T}\boldsymbol{v}, \mathrm{T}\boldsymbol{v}\rangle = \|\mathrm{T}\boldsymbol{v}\| \geqslant 0$. Now suppose that $\left\langle\mathrm{T}^\ell\boldsymbol{v}, \boldsymbol{v}\right\rangle \geqslant 0$ for all integer $\ell < k$. Then

$$\left\langle\mathrm{T}^k\boldsymbol{v}, \boldsymbol{v}\right\rangle = \left\langle\mathrm{T}^{k-1}\boldsymbol{v}, \mathrm{T}\boldsymbol{v}\right\rangle = \left\langle\mathrm{T}^{k-2}(\mathrm{T}\boldsymbol{v}), \mathrm{T}\boldsymbol{v}\right\rangle \geqslant 0$$

by the induction hypothesis. □

► EXERCISE 135 (7.19). *Suppose that* T *is a positive operator on* V*. Prove that* T *is invertible if and only if* $\langle\mathrm{T}\boldsymbol{v}, \boldsymbol{v}\rangle > 0$ *for every* $\boldsymbol{v} \in V \smallsetminus \{\mathbf{0}\}$*.*

PROOF. First assume that $\langle\mathrm{T}\boldsymbol{v}, \boldsymbol{v}\rangle > 0$ for every $\boldsymbol{v} \in V \smallsetminus \{\mathbf{0}\}$. Then $\mathrm{T}\boldsymbol{v} \neq \mathbf{0}$; that is, T is injective, which means that T is invertible.

Now suppose that T is invertible. Since T is self-adjoint, there exists an orthonormal basis $(\boldsymbol{v}_1, \dots, \boldsymbol{v}_n)$ consisting of eigenvectors of T by the Real Spectral Theorem. Let $(\lambda_1, \dots, \lambda_n)$ be the corresponding eigenvalues. Since T injective, we know that $\mathrm{T}\boldsymbol{v}_i \neq \mathbf{0}$ for all $i = 1, \dots, n$; hence, $\lambda_i \neq 0$ for all $i = 1, \dots, n$.

For every $\boldsymbol{v} \in V \smallsetminus \{\mathbf{0}\}$, there exists a list $(a_1, \dots, a_n)$, not all zero, such that $\boldsymbol{v} = \sum_{i=1}^n a_i\boldsymbol{v}_i$. Then

$$\langle\mathrm{T}\boldsymbol{v}, \boldsymbol{v}\rangle = \left\langle\sum_{i=1}^n a_i\mathrm{T}\boldsymbol{v}_i, \sum_{i=1}^n a_i\boldsymbol{v}_i\right\rangle = \left\langle\sum_{i=1}^n a_i\lambda_i\boldsymbol{v}_i, \sum_{i=1}^n a_i\boldsymbol{v}_i\right\rangle = \sum_{i=1}^n \lambda_i|a_i|^2 > 0.$$ □

► EXERCISE 136 (7.20). *Prove or disprove: the identity operator on* $\mathbb{F}^2$ *has infinitely many self-adjoint square roots.*

PROOF. Let

$$\mathcal{M}(\mathrm{S}) = \begin{pmatrix}\cos\theta & \sin\theta \\ \sin\theta & -\cos\theta\end{pmatrix}.$$

Then $\mathrm{S}^2 = \mathrm{Id}$. Hence, there are infinitely many self-adjoint square roots. □

Part II
Linear Algebra and Its Application (Lax, 2007)

8

FUNDAMENTALS

► Exercise 137 (1.1). *Show that the zero of vector addition is unique.*

Proof. Suppose that $\mathbf{0}$ and $\mathbf{0}'$ are both additive identities for some vector. Then $\mathbf{0}' = \mathbf{0}' + \mathbf{0} = \mathbf{0}$. □

► Exercise 138 (1.2). *Show that the vector with all components zero serves as the zero element of classical vector addition.*

Proof. Let $\mathbf{0} = (0, \ldots, 0)$. Then $\boldsymbol{x} + \mathbf{0} = (x_1, \ldots, x_n) + (0, \ldots, 0) = (x_1, \ldots, x_n) = \boldsymbol{x}$. □

Example 1 (Examples of Linear Spaces).

(i) Set of all row vectors: $(a_1, \ldots, a_n)$, $a_j \in K$; addition, multiplication defined componentwise. This space is denoted as K^n.

(ii) Set of all real-valued functions $f(x)$ defined on the real line, $K = \mathbb{R}$.

(iii) Set of all functions with values in K, defined on an arbitrary set S.

(iv) Set of all polynomials of degree less than n with coefficients in K.

► Exercise 139 (1.3). *Show that (i) and (iv) are isomorphic.*

Proof. Let $\mathfrak{P}_{n-1}(K)$ denote the set of all polynomials of degree less than n with coefficients in K, that is,

$$\mathfrak{P}_{n-1}(K) = \left\{ a_1 + a_2 x + \cdots + a_n x^{n-1} \,\middle|\, a_1, \ldots, a_n \in K \right\}.$$

Then, $(a_1, \ldots, a_n) \mapsto a_1 + a_2 x + \cdots + a_n x^{n-1}$ is an isomorphism. □

► Exercise 140 (1.4). *Show that if S has n elements, (i) and (iii) are isomorphic.*

Proof. Let $|S| = n$. Then any function $f \in K^S$ can be written as

$$\big(f(s_1), \ldots, f(s_n) \big) = (a_1, \ldots, a_n),$$

where $s_1, \ldots, s_n \in S$. □

► EXERCISE 141 (1.5). *Show that when $K = \mathbb{R}$, (iv) is isomorphic with (iii) when S consists of n distinct points of $\mathbb{R}$.*

PROOF. We need to show that $\mathbb{R}^S$ is isomorphic to $\mathfrak{P}_{n-1}(\mathbb{R})$. We can write each $f \in \mathbb{R}^S$ as $(a_1, \ldots, a_n)$, and consider the map $(a_1, \ldots, a_n) \mapsto a_1 + a_2 x + \cdots + a_n x^{n-1}$. □

► EXERCISE 142 (1.6). *Prove that $Y + Z$ is a linear subspace of X if Y and Z are.*

PROOF. If $\boldsymbol{y}_1 + \boldsymbol{z}_1, \boldsymbol{y}_2 + \boldsymbol{z}_2 \in Y + Z$, then $(\boldsymbol{y}_1 + \boldsymbol{z}_1) + (\boldsymbol{y}_2 + \boldsymbol{z}_2) = (\boldsymbol{y}_1 + \boldsymbol{y}_2) + (\boldsymbol{z}_1 + \boldsymbol{z}_2) \in Y + Z$; if $\boldsymbol{y} + \boldsymbol{z} \in Y + Z$ and $k \in K$, then $k(\boldsymbol{y} + \boldsymbol{z}) = k\boldsymbol{y} + k\boldsymbol{z} \in Y + Z$. □

► EXERCISE 143 (1.7). *Prove that if Y and Z are linear subspaces of X, so is $Y \cap Z$.*

PROOF. If $\boldsymbol{x}, \boldsymbol{y} \in Y \cap Z$, then $\boldsymbol{x} + \boldsymbol{y} \in Y$ and $\boldsymbol{x} + \boldsymbol{y} \in Z$, which imply that $\boldsymbol{x} + \boldsymbol{y} \in Y \cap Z$; if $\boldsymbol{x} \in Y \cap Z$, then $\boldsymbol{x} \in Y$ and $\boldsymbol{x} \in Z$; since both Y and X are subspaces of X, we have $k\boldsymbol{x} \in Y$ and $k\boldsymbol{y} \in Z$ for all $k \in K$, that is $k\boldsymbol{x} \in Y \cap Z$. □

► EXERCISE 144 (1.8). *Show that the set $\{\mathbf{0}\}$ consisting of the zero element of a linear space X is a subspace of X. It is called the trivial subspace.*

PROOF. Trial. □

► EXERCISE 145 (1.9). *Show that the set of all linear combinations of $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j$ is a subspace of X, and that is the smallest subspace of X containing $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j$. This is called the* subspace spanned *by $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j$.*

PROOF. Let $\text{span}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j) \equiv \left\{\boldsymbol{x} : \boldsymbol{x} = \sum_{i=1}^{j} k_i \boldsymbol{x}_i\right\}$. Let $\boldsymbol{x} = \sum_{i=1}^{j} k_i \boldsymbol{x}_i$ and $\boldsymbol{x}' = \sum_{i=1}^{j} k_i' \boldsymbol{x}_i$. Then

$$\boldsymbol{x} + \boldsymbol{x}' = \sum_{i=1}^{j} \left(k_i + k_i'\right) \boldsymbol{x}_i,$$

and

$$k\boldsymbol{x} = \sum_{i=1}^{j} (k k_i)\, \boldsymbol{x}_i.$$

Hence, the set of all linear combinations of $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j$ is a subspace of X.

Since $\boldsymbol{x}_i = 1 \cdot \boldsymbol{x}_i + \sum_{\ell \neq i} 0 \cdot \boldsymbol{x}_\ell$, each $\boldsymbol{x}_i$ is a linear combination of $\left(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j\right)$. Thus, $\text{span}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j)$ contains each $\boldsymbol{x}_i$. Conversely, because subspaces are closed under scalar multiplication and addition, every subspace of V containing each $\boldsymbol{x}_i$ must contain $\text{span}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j)$. □

► EXERCISE 146 (1.10). *Show that if the vectors $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j$ are linearly independent, then none of the $\boldsymbol{x}_i$ is the zero vector.*

PROOF. Suppose that there is a vector $\boldsymbol{x}_i = \mathbf{0}$. Then

$$k_i \cdot \mathbf{0} + \sum_{\ell \neq i} 0 \cdot \boldsymbol{x}_\ell = 0, \quad \forall\, k \neq 0,$$

that is, the list $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_j)$ is linearly dependent. □

► EXERCISE 147 (1.11). *Prove that if X is finite dimensional and the direct sum of $Y_1, \ldots, Y_m$, then* $\dim X = \sum_{j=1}^m \dim Y_j$.

PROOF. Let $\left(\boldsymbol{y}_1^1, \ldots, \boldsymbol{y}_{n_1}^1\right)$ be a basis of Y_1, …, $\left(\boldsymbol{y}_1^m, \ldots, \boldsymbol{y}_{n_m}^m\right)$ be a basis of Y_m. We show that the list $B = \left(\boldsymbol{y}_1^1, \ldots, \boldsymbol{y}_{n_1}^1, \ldots, \boldsymbol{y}_1^m, \ldots, \boldsymbol{y}_{n_m}^m\right)$ is a basis of $X = Y_1 \oplus \cdots \oplus Y_m$. To see $X = \operatorname{span}(B)$, note that for any $\boldsymbol{x} \in X$, there exists a unique list $(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m)$ with $\boldsymbol{y}_i \in Y_i$ such that $\boldsymbol{x} = \sum_{i=1}^m \boldsymbol{y}_i$. But each $\boldsymbol{y}_i$ can be uniquely represented as $\boldsymbol{y}_i = \sum_{j=1}^{n_i} = a_j^i \boldsymbol{y}_j^i$; thus,

$$\boldsymbol{x} = \sum_{i=1}^m \sum_{j=1}^{n_i} a_j^i \boldsymbol{y}_j^i.$$

To see that the list B is linearly independent, suppose that there exists a list of scalars $\boldsymbol{b} = \left(b_1^1, \ldots, b_{n_1}^1, \ldots, b_1^m, \ldots, b_{m_n}^m\right)$, such that

$$b_1^1 \boldsymbol{y}_1^1 + \cdots + b_{n_1}^1 \boldsymbol{y}_{n_1}^1 + \cdots + b_1^m \boldsymbol{y}_1^m + \cdots + b_{m_n}^m \boldsymbol{y}_{m_n}^m = \mathbf{0}_X.$$

But $\mathbf{0}_X = 0\boldsymbol{y}_1^1 + \cdots + 0\boldsymbol{y}_{n_1}^1 + \cdots + 0\boldsymbol{y}_1^m + \cdots + 0\boldsymbol{y}_{m_n}^m$ and $X = Y_1 \oplus \cdots \oplus Y_m$ implies that all the scalars are zero, that is, B is linearly independent. Therefore, $\dim X = \sum_{j=1}^m \dim Y_j$. □

► EXERCISE 148 (1.12). *Show that every finite-dimensional space X over K is isomorphic to K^n, $n = \dim X$. Show that this isomorphism is not unique when n is > 1.*

PROOF. Let $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ be a basis of X, and $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ be a basis of K^n. Define a linear map $\mathrm{T} \in \mathfrak{L}(X, K^n)$ by letting $\mathrm{T}\boldsymbol{x}_i = \boldsymbol{e}_i$. Then for any $\boldsymbol{x} = \sum_{i=1}^n a_i \boldsymbol{x}_i \in X$, we have

$$\mathrm{T}\boldsymbol{x} = \mathrm{T}\left(\sum_{i=1}^n a_i \boldsymbol{x}_i\right) = \sum_{i=1}^n a_i \mathrm{T}\boldsymbol{x}_i = \sum_{i=1}^n a_i \boldsymbol{e}_i.$$

We first show that T is surjective. For any $\boldsymbol{k} \in K^n$, there exists $(k_1, \ldots, k_n)$ such that $\boldsymbol{k} = \sum_{i=1}^n k_i \boldsymbol{e}_i$, and so there exists $\boldsymbol{x}_{\boldsymbol{k}} = \sum_{i=1}^n k_i \boldsymbol{x}_i \in X$ such that $\mathrm{T}\boldsymbol{x}_{\boldsymbol{k}} = \sum_{i=1}^n k_i \boldsymbol{e}_i = \boldsymbol{k}$. To see T is injective, let

$$\mathrm{T}\left(\sum_{i=1}^n a_i \boldsymbol{x}_i\right) = \mathrm{T}\left(\sum_{i=1}^n b_i \boldsymbol{x}_i\right),$$

that is,

$$\sum_{i=1}^{n} a_i \boldsymbol{e}_i = \sum_{i=1}^{n} b_i \boldsymbol{e}_i \iff \sum_{i=1}^{n} (a_i - b_i)\, \boldsymbol{e}_i = \boldsymbol{0} \iff a_i = b_i \quad \forall\, i = 1, \dots, n$$

since $(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ is linearly independent. Thus, $\sum_{i=1}^{n} a_i \boldsymbol{x}_i = \sum_{i=1}^{n} b_i \boldsymbol{x}_i$.

The isomorphism is not unique when $n > 1$ since there are many, many basis. □

► Exercise 149 (1.13). *Congruence mod Y is an equivalence relation. Show further that if $\boldsymbol{x}_1 \equiv \boldsymbol{x}_2$, then $k\boldsymbol{x}_1 \equiv k\boldsymbol{x}_2$ for every scalar k.*

Proof. (i) If $\boldsymbol{x}_1 \equiv \boldsymbol{x}_2$, then $\boldsymbol{x}_1 - \boldsymbol{x}_2 \in Y$, which means that $\boldsymbol{x}_2 - \boldsymbol{x}_1 = -(\boldsymbol{x}_1 - \boldsymbol{x}_2) \in Y$ since Y is a subspace; (ii) $\boldsymbol{x} - \boldsymbol{x} = \boldsymbol{0} \in Y$; (iii) if $\boldsymbol{x}_1 - \boldsymbol{x}_2 \in Y$ and $\boldsymbol{x}_2 - \boldsymbol{x}_3 \in Y$, then $\boldsymbol{x}_1 - \boldsymbol{x}_3 = (\boldsymbol{x}_1 - \boldsymbol{x}_2) + (\boldsymbol{x}_2 - \boldsymbol{x}_3) \in Y$, that is, $\boldsymbol{x}_1 \equiv \boldsymbol{x}_3$.

If $\boldsymbol{x}_1 \equiv \boldsymbol{x}_2 \mod Y$, then $\boldsymbol{x}_1 - \boldsymbol{x}_2 \in Y$ and so $k(\boldsymbol{x}_1 - \boldsymbol{x}_2) \in Y$ since Y is a subspace of X. But then $k\boldsymbol{x}_1 - k\boldsymbol{x}_2 \in Y$, i.e., $k\boldsymbol{x}_1 \equiv k\boldsymbol{x}_2 \mod Y$. □

► Exercise 150 (1.14). *Show that two congruence classes are either identical or disjoint.*

Proof. Let $\boldsymbol{x}_3 \in [\boldsymbol{x}_1] \cap [\boldsymbol{x}_2]$. Then $\boldsymbol{x}_1 \equiv \boldsymbol{x}_3$ and $\boldsymbol{x}_3 \equiv \boldsymbol{x}_2$. By transitivity of $\equiv$ we have $\boldsymbol{x}_1 \equiv \boldsymbol{x}_2$, that is, $[\boldsymbol{x}_1] = [\boldsymbol{x}_2]$. □

► Exercise 151 (1.15). *Show that the above definition of addition and multiplication by scalars is independent of the choice of representatives in the congruence class.*[1]

Proof. By definition, $[\boldsymbol{x}] + [\boldsymbol{z}] = [\boldsymbol{x} + \boldsymbol{z}] = (\boldsymbol{x} + \boldsymbol{z}) + Y$, and $k[\boldsymbol{x}] = [k\boldsymbol{x}] = k\boldsymbol{x} + Y$. Note that $[\boldsymbol{x}'] = [\boldsymbol{x}]$ if $\boldsymbol{x}' \in [\boldsymbol{x}]$. □

► Exercise 152 (1.16). *Denote by X the linear space of all polynomials $p(t)$ of degree $< n$, and denote by Y the set of polynomials that are zero at $t_1, \dots, t_j$, $j < n$.*

a. *Show that Y is a subspace of X.*

b. *Determine* $\dim Y$.

c. *Determine* $\dim X/Y$.

Proof.

a. Any $p \in \mathfrak{P}_{n-1}(K)$ with roots $t_1, \dots, t_j$ can be written in the form

$$q(t) \prod_{i=1}^{j} \left(t - t_j\right),$$

where $q(t) \in \mathfrak{P}_{n-1-j}(K)$. These clearly form a vector space.

[1] We have $[\boldsymbol{x}] = \boldsymbol{x} + Y$. **Proof:** If $\boldsymbol{z} \in [\boldsymbol{x}]$, then there exists $\boldsymbol{y} \in Y$ such that $\boldsymbol{z} - \boldsymbol{x} = \boldsymbol{y}$; then $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y} \in \boldsymbol{x} + Y$. Conversely, if $\boldsymbol{z} \in \boldsymbol{x} + Y$, then $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$ for some $\boldsymbol{y} \in Y$; hence, $\boldsymbol{z} - \boldsymbol{x} = \boldsymbol{y} \in Y$, i.e., $\boldsymbol{z} \in [\boldsymbol{x}]$.

b. $\dim Y = n - j$.

c. $\dim X/Y = \dim X - \dim Y = n - \big(n - j\big) = j$.

□

► EXERCISE 153 (1.17). *A subspace Y of a finite-dimensional linear space X whose dimension is the same as the dimension of X is all of X.*

PROOF. Suppose that $Y \subsetneq X$, then there exists $\boldsymbol{x} \in X \smallsetminus Y$ and $\boldsymbol{x} \neq \mathbf{0}_X$ since $\mathbf{0}_X \in Y$. Let $[\boldsymbol{x}] = \boldsymbol{x} + Y$. Thus, $[\boldsymbol{x}] \in X/Y$ and $[\boldsymbol{x}] \neq Y = \mathbf{0}_{X/Y}$ and so $\dim X/Y \geqslant 1$, which implies that $\dim Y = \dim X - \dim X/Y < \dim X$ by Theorem 1.6. A contradiction. □

► EXERCISE 154 (1.18). *Show that* $\dim X_1 \oplus X_2 = \dim X_1 + \dim X_2$.

PROOF. $X_1 \oplus X_2$ implies that $X_1 \cap X_2 = \{\mathbf{0}\}$, that is, $\dim X_1 \cap X_2 = 0$. Therefore, $\dim X_1 \oplus X_2 = \dim X_1 + \dim X_2$. See Exercise 147. □

► EXERCISE 155 (1.19). *X is a linear space, Y a subspace. Show that $Y \oplus X/Y$ is isomorphic to X.*

PROOF. According to Exercise 148, we only need to show that $\dim Y \oplus X/Y = \dim X$. This holds since

$$\dim Y \oplus X/Y = \dim Y + \dim X/Y = \dim X.$$

□

► EXERCISE 156 (1.20). *Which of the following sets of vectors $\boldsymbol{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ are a subspace of $\mathbb{R}^n$? Explain your answer.*

a. *All $\boldsymbol{x}$ such that $x_1 \geqslant 0$.*

b. *All $\boldsymbol{x}$ such that $x_1 + x_2 = 0$.*

c. *All $\boldsymbol{x}$ such that $x_1 + x_2 + 1 = 0$.*

d. *All $\boldsymbol{x}$ such that $x_1 = 0$.*

e. *All $\boldsymbol{x}$ such that x_1 is an integer.*

PROOF.

a. No. $-\boldsymbol{x}$ is not in that set if $x_1 > 0$.

b. Yes.

c. No. $k\boldsymbol{x}$ is not in that set if $k \neq 1$.

d. Yes.

e. No. $k\boldsymbol{x}$ is not in that set if $k \notin \mathbb{Z}$.

□

► EXERCISE 157 (1.21). *Let U, V, and W be subspaces of some finite-dimensional vector space X. Is the statement*

$$\begin{aligned} \dim U + V + W = {} & \dim U + \dim V + \dim W - \dim U \cap V \\ & - \dim U \cap W - \dim V \cap W + \dim U \cap V \cap W, \end{aligned}$$

true or false? If true, prove it. If false, provide a counterexample.

PROOF. It is false. See Exercise 30. □

9

DUALITY

REMARK (Theorem 3). The bilinear function $\langle \boldsymbol{x}, \ell \rangle$ gives a natural identification of X with X''.

PROOF. For $\langle \boldsymbol{x}, \ell \rangle$, fix $\boldsymbol{x} = \boldsymbol{x}_0$, then we observe that the function of the vectors in X', whose value at ℓ is $\langle \boldsymbol{x}_0, \ell \rangle = \ell(\boldsymbol{x}_0)$, is a scalar-valued function that happens to be linear [**Proof:** Let $z_0 \in X''$ be so defined. For any $\ell, \ell' \in X'$, we have $z_0(\ell + \ell') = \langle \boldsymbol{x}_0, \ell + \ell' \rangle = (\ell + \ell')(\boldsymbol{x}_0) = \ell(\boldsymbol{x}_0) + \ell'(\boldsymbol{x}_0) = z_0(\ell) + z_0(\ell')$. For any $k \in K$ and $\ell \in X'$, we have $z_0(k\ell) = \langle \boldsymbol{x}_0, k\ell \rangle = (k\ell)(\boldsymbol{x}_0) = k\ell(\boldsymbol{x}_0) = k \cdot z_0(\ell)$.] Thus,

> $\langle \boldsymbol{x}_0, \ell \rangle$ defines a linear functional on X', and consequently, and element of X''.

By this method we have exhibited *some* linear functionals on X'; have we exhibited them all? For the finite-dimensional case the following theorem furnishes the affirmative answer.

> If X is a finite-dimensional vector space, then corresponding to every linear functional z_0 on X' there is a vector $\boldsymbol{x}_0 \in X$ such that $z_0(\ell) = \langle \boldsymbol{x}_0, \ell \rangle = \ell(\boldsymbol{x}_0)$ for every $\ell \in X'$; the correspondence $z_0 \leftrightarrow \boldsymbol{x}_0$ between X'' and X is an isomorphism.

Proof: To every $\boldsymbol{x}_0 \in X$, we make correspond a vector $z_{\boldsymbol{x}_0} \in X''$ defined by $z_{\boldsymbol{x}_0}(\ell) = \ell(\boldsymbol{x}_0)$ for every $\ell \in X'$. We first show that the transformation $\boldsymbol{x}_0 \mapsto z_{\boldsymbol{x}_0}$ is linear. For any $\boldsymbol{x}_0, \boldsymbol{x}_1 \in X$, we have $\boldsymbol{x}_0 + \boldsymbol{x}_1 \mapsto z_{\boldsymbol{x}_0 + \boldsymbol{x}_1}$; by definition, $z_{\boldsymbol{x}_0+\boldsymbol{x}_1}(\ell) = \ell(\boldsymbol{x}_0 + \boldsymbol{x}_1) = \ell(\boldsymbol{x}_0) + \ell(\boldsymbol{x}_1) = z\boldsymbol{x}_0(\ell) + z_{\boldsymbol{x}_1}(\ell)$ for any $\ell \in X'$. For any $k \in K$ and $\boldsymbol{x}_0 \in X$, we have $k\boldsymbol{x}_0 \mapsto z_{k\boldsymbol{x}_0}$ and so $z_{k\boldsymbol{x}_0}(\ell) = \ell(k\boldsymbol{x}_0) = k \cdot \ell(\boldsymbol{x}_0) = k \cdot z_{\boldsymbol{x}_0}$ for any $\ell \in X'$.

We shall show that this transformation is injective. Take any $z_{\boldsymbol{x}_1}, z_{\boldsymbol{x}_2} \in X''$ with $z_{\boldsymbol{x}_1} = z_{\boldsymbol{x}_2}$. To say that $z_{\boldsymbol{x}_1} = z_{\boldsymbol{x}_2}$ means that $\langle \boldsymbol{x}_1, \ell \rangle = \langle \boldsymbol{x}_2, \ell \rangle$ for every $\ell \in X'$. But then $\boldsymbol{x}_1 = \boldsymbol{x}_2$ by Exercise 158 (iii).

Therefore, the set $Z \equiv \{z_{\boldsymbol{x}} : \boldsymbol{x} \in X\}$ is a subspace of X'' since Z is the range under a linear map, and Z is isomorphic to X'', and so $\dim Z = \dim X$. Since $\dim X = \dim X' = \dim X''$, we have $\dim Z = \dim X''$. It follows that $X'' = Z$ by Exercise 153. □

REMARK (p. 16). $Y^{\perp}$ is isomorphic to $(X/Y)'$.

PROOF. Given $\ell \in Y^{\perp}$, we make correspond a linear functional $L_\ell \in \left(X/Y\right)'$ defined by

$$L_\ell[\boldsymbol{x}] = \ell(\boldsymbol{x}).$$

We show first that $L_\ell[\boldsymbol{x}] = \ell(\boldsymbol{x})$ is well defined.[1] Let $[\boldsymbol{x}_1] = [\boldsymbol{x}]$; then there exists $\boldsymbol{y}_1 \in Y$ such that $\boldsymbol{x}_1 = \boldsymbol{x} + \boldsymbol{y}_1$, as depicted in Figure 9.1. Thus,

$$\ell(\boldsymbol{x}_1) = \ell(\boldsymbol{x} + \boldsymbol{y}_1) = \ell(\boldsymbol{x}) + \ell(\boldsymbol{y}_1) = \ell(\boldsymbol{y}),$$

that is, $L_\ell[\boldsymbol{x}_1] = L_\ell[\boldsymbol{x}]$ if $[\boldsymbol{x}_1] = [\boldsymbol{x}]$. We then show L_ℓ such defined is linear. For any $[\boldsymbol{x}], [\boldsymbol{y}] \in X/Y$, we have $L_\ell([\boldsymbol{x}] + [\boldsymbol{y}]) \overset{\langle 1 \rangle}{=} L_\ell[\boldsymbol{x} + \boldsymbol{y}] = \ell(\boldsymbol{x} + \boldsymbol{y}) = \ell(\boldsymbol{x}) + \ell(\boldsymbol{y}) = L_\ell[\boldsymbol{x}] + L_\ell[\boldsymbol{y}]$, where $\langle 1 \rangle$ holds since $[\boldsymbol{x}] + [\boldsymbol{y}] = (\boldsymbol{x} + Y) + (\boldsymbol{y} + Y) = (\boldsymbol{x} + \boldsymbol{y}) + Y = [\boldsymbol{x} + \boldsymbol{y}]$. To see L_ℓ is homogenous, take any $k \in K$ and $[\boldsymbol{x}] \in X/Y$. Then $L_\ell(k \cdot [\boldsymbol{x}]) = L_\ell[k\boldsymbol{x}] = \ell(k\boldsymbol{x}) = k\ell(\boldsymbol{x}) = k \cdot L_\ell[\boldsymbol{x}]$.

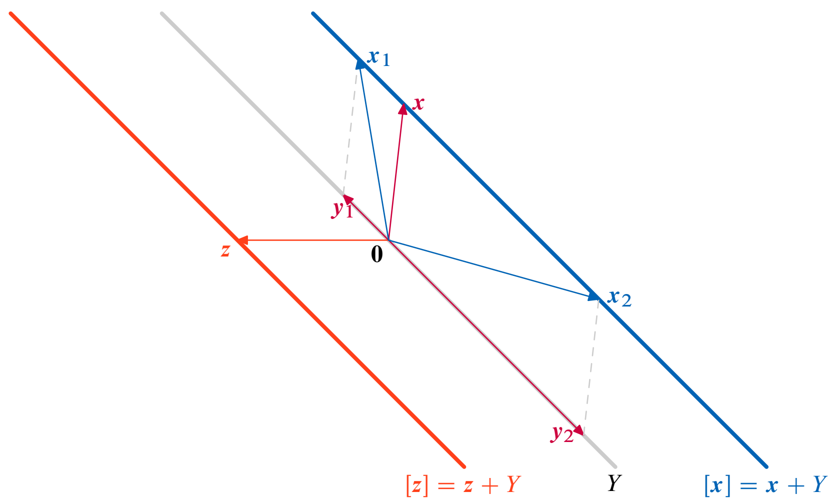


FIGURE 9.1. $Y^{\perp} \cong (X/Y)'$

Conversely, given any $L \in \left(X/Y\right)'$, define a linear functional ℓ_L on X as

$$\ell_L(\boldsymbol{x}) = L[\boldsymbol{x}].$$

It follows from the above definition that $\ell_L \in Y^{\perp}$: for any $\boldsymbol{y} \in Y$, we have $[\boldsymbol{y}] = \boldsymbol{y} + Y = Y$ and so $\ell_L(\boldsymbol{y}) = L[Y] = 0$. This also proves that the correspondence between $Y^{\perp}$ and $\left(X/Y\right)'$ is surjective. We finally to show that the mapping $\ell \mapsto L_\ell$ is injective. Take two $\ell_1, \ell_2 \in Y^{\perp}$ such that $L_{\ell_1} = L_{\ell_2}$, where $L_{\ell_1}, L_{\ell_2} \in \left(X/Y\right)'$. To say $L_{\ell_1} = L_{\ell_2}$ means that $L_{\ell_1}[\boldsymbol{x}] = L_{\ell_2}[\boldsymbol{x}]$ for all $[\boldsymbol{x}] \in X/Y$, but which means that $\ell_1(\boldsymbol{x}) = \ell_2(\boldsymbol{x})$ for all $\boldsymbol{x} \in X$, i.e., $\ell_1 = \ell_2$. Thus, $\ell \mapsto L_\ell$ is injective. □

[1] To see the relation between congruence classes and affine sets (linear manifolds), refer Rockafellar (1970).

▶ EXERCISE 158 (2.1). *Given a nonzero vector $\boldsymbol{x}_1 \in X$, show that there is a linear function ℓ such that $\ell(\boldsymbol{x}_1) \neq 0$.*

PROOF. See Halmos (1974, Sec. 15).

(

i) If X is an n-dimensional vector space, if $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ is a basis of X, and if $(a_1, \dots, a_n)$ is any list of n scalars, then there is one and only one linear functional ℓ on X such that $\langle \boldsymbol{x}_i, \ell \rangle = a_i$ for $i = 1, \dots, n$.

Proof: Every $\boldsymbol{x} \in X$ can be represented uniquely as $\boldsymbol{x} = \sum_{i=1}^{n} k_i \boldsymbol{x}_i$, where $k_i \in K$. If ℓ is any linear functional, then

$$\langle \boldsymbol{x}, \ell \rangle = \left\langle \sum_{i=1}^{n} k_i \boldsymbol{x}_i, \ell \right\rangle = k_1 \langle \boldsymbol{x}_1, \ell \rangle + \cdots + k_n \langle \boldsymbol{x}_i, \ell \rangle .$$

From this relation the uniqueness of ℓ is clear: if $\langle \boldsymbol{x}_i, \ell \rangle = a_i$, then the value of $\langle \boldsymbol{x}, \ell \rangle$ is determined, for every $\boldsymbol{x}$, by $\langle \boldsymbol{x}, \ell \rangle = \sum_{i=1}^{n} k_i a_i$. The argument can also be turned around; if we define ℓ by

$$\langle \boldsymbol{x}, \ell \rangle = k_1 a_1 + \cdots k_n a_n,$$

then ℓ is indeed a linear functional, and $\langle \boldsymbol{x}_i, \ell \rangle = a_i$.

(

ii) If X is an n-dimensional vector space and if $B = (\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ is a basis of X, then there is a uniquely determined basis B' in X', $B' = (\ell_1, \dots, \ell_n)$, with the property that $\langle \boldsymbol{x}_i, \ell_j \rangle = \delta_{ij}$. Consequently the dual space of an n-dimensional space is n-dimensional.

Proof: It follows from (i) that, for each $j = 1, \dots, n$, a unique $\ell_i \in X'$ can be found so that $\langle \boldsymbol{x}_i, \ell_j \rangle = \delta_{ij}$; we have only to prove that the list $B' = (\ell_1, \dots, \ell_n)$ is a basis in X'. In the first place, B' is linearly independent, for if we had $a_1 \ell_1 + \cdots + a_n \ell_n = 0$, in other words, if

$$\langle \boldsymbol{x}, a_1 \ell_1 + \cdots + a_n \ell_n \rangle = a_1 \langle \boldsymbol{x}, \ell_1 \rangle + \cdots + a_n \langle \boldsymbol{x}, \ell_n \rangle = 0$$

for all $\boldsymbol{x} \in X$, then we should have, for $\boldsymbol{x} = \boldsymbol{x}_i$,

$$0 = \sum_{j=1}^{n} a_j \langle \boldsymbol{x}_i, \ell_j \rangle = \sum_{j=1}^{n} a_j \delta_{ij} = a_i .$$

In the second place, $X' = \operatorname{span}(\ell_1, \dots, \ell_n)$. To prove this, write $\langle \boldsymbol{x}_i, \ell \rangle = a_i$; then, for $\boldsymbol{x} = \sum_{i=1}^{n} k_i \boldsymbol{x}_i$, we have

$$\langle \boldsymbol{x}, \ell \rangle = \left\langle \sum_{i=1}^{n} k_i \boldsymbol{x}_i, \ell \right\rangle = \sum_{i=1}^{n} k_i \langle \boldsymbol{x}_i, \ell \rangle = \sum_{i=1}^{n} k_i a_i.$$

On the other hand,

$$\langle \boldsymbol{x}, \ell_j \rangle = \sum_{i=1}^{n} k_i \langle \boldsymbol{x}_i, \ell_j \rangle = k_j,$$

so that, substituting in the preceding equation, we get

$$\langle \boldsymbol{x}, \ell \rangle = \sum_{i=1}^{n} k_i a_i = \sum_{i=1}^{n} a_i \cdot \langle \boldsymbol{x}, \ell_i \rangle = \left\langle \boldsymbol{x}, \sum_{i=1}^{n} a_i \ell_i \right\rangle.$$

Consequently $\ell = \sum_{i=1}^{n} a_i \ell_i$, and the proof of (ii) is complete.

(

iii) For any non-zero vector $\boldsymbol{x} \in X$ there corresponds a $\ell \in X'$ such that $\langle \boldsymbol{x}, \ell \rangle \neq 0$.

Proof: Let $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ be a basis of X, and let $(\ell_1, \ldots, \ell_n)$ be the dual basis in X'. If $\boldsymbol{x} = \sum_{i=1}^{n} k_i \boldsymbol{x}_i$, then $\langle \boldsymbol{x}, \ell_j \rangle = k_j$. Hence if $\langle \boldsymbol{x}, \ell \rangle = 0$ for all ℓ, in particular, if $\langle \boldsymbol{x}, \ell_j \rangle = 0$ for $j = 1, \ldots, n$, then $k_j = 0$ and so $\boldsymbol{x} = \boldsymbol{0}_X$. □

► Exercise 159 (2.2). *Verify that $Y^{\perp}$ is a subspace of X'.*

Proof. (i) Obviously that $0 \in Y^{\perp}$ since $\langle \boldsymbol{x}, 0 \rangle = 0$ for any $\boldsymbol{x} \in X$, including $\boldsymbol{y} \in Y \subseteq X$. (ii) Let $\ell, m \in Y^{\perp}$. Then $\langle \boldsymbol{y}, \ell \rangle = 0 = \langle \boldsymbol{y}, m \rangle$ for all $\boldsymbol{y} \in Y$ and so $\langle \boldsymbol{y}, \ell + m \rangle = \langle \boldsymbol{y}, \ell \rangle + \langle \boldsymbol{y}, m \rangle = 0$, i.e., $\ell + m \in Y^{\perp}$. (iii) If $\ell \in Y^{\perp}$, then $k \langle \boldsymbol{y}, \ell \rangle = 0$ for any $\boldsymbol{y} \in Y$, and so $k\ell \in Y^{\perp}$. Thus $Y^{\perp}$ is a subspace of X'. □

► Exercise 160 (2.3). *Denote by Y the smallest subspace containing S. Then $S^{\perp} = Y^{\perp}$.*

Proof. It is clear that $Y^{\perp} \subseteq S^{\perp}$. If $S = \varnothing$, then $Y = \{\boldsymbol{0}\}$ and the conclusion is obvious. Similarly, the proof is trivial if $S = \{\boldsymbol{0}\}$. So we suppose that $S \neq \varnothing$ and $S \neq \{\boldsymbol{0}\}$. Take any $\boldsymbol{y}_1 \in S$ with $\boldsymbol{y}_1 \neq \boldsymbol{0}$. If $S \subseteq \operatorname{span}(\boldsymbol{y}_1)$, let $Y = \operatorname{span}(\boldsymbol{y}_1)$; if there is $\boldsymbol{y}_2 \in S \setminus \operatorname{span}(\boldsymbol{y}_1)$, let $Y = \operatorname{span}(\boldsymbol{y}_1, \boldsymbol{y}_2)$; …. Since the embedding vector space is finite-dimensional, the process will be ended with a list $(\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$ with $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n \in S$, and this list is a basis of Y. Then for any $\ell \in S^{\perp}$ and any $\boldsymbol{y} \in Y$, we have

$$\langle \boldsymbol{y}, \ell \rangle = \left\langle \sum_{i=1}^{n} k_i \boldsymbol{y}_i, \ell \right\rangle = \sum_{i=1}^{n} k_i \langle \boldsymbol{y}_i, \ell \rangle = 0$$

since $\ell(\boldsymbol{y}_i) = 0$. Thus, $\ell \in Y^{\perp}$. □

► Exercise 161 (2.4). *In Theorem 7 take the interval I to be $[-1, 1]$, and take $n = 3$. Choose the three points to be $t_1 = -a$, $t_2 = 0$, and $t_3 = a$.*

a. *Determine the weights* m_1, m_2, m_3 *so that* $\int_I p(t)\,\mathrm{d}t = m_1 p(t_1) + m_2 p(t_2) + m_3 p(t_3)$ *holds for all polynomials* $p \in \mathfrak{P}_2(K)$.

b. *Show that for* $a > \sqrt{1/3}$*, all three weights are positive.*

c. *Show that for* $a = \sqrt{3/5}$*, (9) holds for all* $p \in \mathfrak{P}_5(K)$.

PROOF.

a. If $p(t) = t$, then $\int_{-1}^{1} t\,\mathrm{d}t = 0$ and so $0 = m_1(-a) + m_3 a$, i.e., $m_1 = m_3$. Then (9) can be rewritten as

$$\int_{-1}^{1} p(t)\,\mathrm{d}t = m_1\left[p(-a) + p(a)\right] + m_2 p(0)\,. \tag{9.1}$$

Take $p(t) = 1$ now. Then $2 = \int_{-1}^{1} \mathrm{d}t = 2m_1 + m_2$, i.e., $m_2 = 2(1 - m_1)$. So we rewrite (9.1) as

$$\int_{-1}^{1} p(t)\,\mathrm{d}t = m_1\left[p(-a) + p(a)\right] + 2(1 - m_1)\,p(0)\,. \tag{9.2}$$

Now let $p(t) = t^2$ and hence $p(0) = 0$. We then have $\frac{2}{3} = \int_{-1}^{1} t^2\,\mathrm{d}t = m_1 2a^2$ implies that

$$m_1 = m_3 = \frac{1}{3a^2}, \quad \text{and} \quad m_2 = 2 - \frac{2}{3a^2}.$$

□

10

LINEAR MAPPINGS

► EXERCISE 162 (3.1). *The image of a subspace of X under a linear map* T *is a subspace of U. The inverse image of a subspace of U, that is the set of all vectors in X mapped by* T *into the subspace, is a subspace of X.*

PROOF. Let Y be a subspace of X; then $\mathbf{0}_X \in Y$ and so $\mathbf{0}_U = \mathrm{T}\mathbf{0}_X \in \mathrm{T}[Y]$. To see $\mathrm{T}[Y]$ is closed under addition, take any $\mathrm{T}\boldsymbol{x}, \mathrm{T}\boldsymbol{y} \in \mathrm{T}[Y]$. Then $\boldsymbol{x} + \boldsymbol{y} \in Y$ and $\mathrm{T}\boldsymbol{x} + \mathrm{T}\boldsymbol{y} = \mathrm{T}(\boldsymbol{x} + \boldsymbol{y}) \in \mathrm{T}[Y]$; to see $\mathrm{T}[Y]$ is closed under scalar multiplication, take any $k \in K$ and $\mathrm{T}\boldsymbol{x} \in \mathrm{T}[Y]$; then $k \ln \in Y$ and $k\mathrm{T}\boldsymbol{x} = \mathrm{T}(k\boldsymbol{x}) \in \mathrm{T}[Y]$. Thus $\mathrm{T}[Y]$ is a subspace of U.

We then show that $\mathrm{T}^{-1}[V]$ is a subspace of X is V is a subspace of U. (i) $\mathbf{0}_X \in \mathrm{T}^{-1}[V]$ since $\mathbf{0}_U \in V$ and $\mathrm{T}\mathbf{0}_X = \mathbf{0}_U$. (ii) For any $\boldsymbol{x}, \boldsymbol{y} \in \mathrm{T}^{-1}[V]$, we have $\mathrm{T}\boldsymbol{x}, \mathrm{T}\boldsymbol{y} \in V$ and so $\mathrm{T}(\boldsymbol{x} + \boldsymbol{y}) = \mathrm{T}\boldsymbol{x} + \mathrm{T}\boldsymbol{y} \in V$, i.e., $\boldsymbol{x} + \boldsymbol{y} \in \mathrm{T}^{-1}[V]$. (iii) For any $k \in K$ and $\boldsymbol{x} \in \mathrm{T}^{-1}[V]$, we have $\mathrm{T}\boldsymbol{x} \in V$ and $k\mathrm{T}\boldsymbol{x} \in V$; then $\mathrm{T}(k\boldsymbol{x})\, k\mathrm{T}\boldsymbol{x} \in V$, that is, $k\boldsymbol{x} \in k\mathrm{T}\boldsymbol{x}$. □

► EXERCISE 163 (3.3).

a. *The composite of linear mappings is also a linear mapping.*

b. *Composition is distributive with respect to the addition of linear maps, that is,* $(\mathrm{R} + \mathrm{S}) \circ \mathrm{T} = \mathrm{R} \circ \mathrm{T} + \mathrm{S} \circ \mathrm{T}$ *and* $\mathrm{S} \circ (\mathrm{T} + \mathrm{P}) = \mathrm{S} \circ \mathrm{T} + \mathrm{S} \circ \mathrm{P}$*, where* R *and* S *map* $U \to V$ *and* P *and* T *map* $X \to U$.

PROOF.

a. Let $\mathrm{S}, \mathrm{T} \in \mathfrak{L}(X, U)$ and consider $\mathrm{S} \circ \mathrm{T}$. To see $\mathrm{S} \circ \mathrm{T}$ is additive, take any $\boldsymbol{x}, \boldsymbol{y} \in X$; then $(\mathrm{S} \circ \mathrm{T})(\boldsymbol{x} + \boldsymbol{y}) = \mathrm{S}\big[\mathrm{T}(\boldsymbol{x} + \boldsymbol{y})\big] = \mathrm{S}[\mathrm{T}\boldsymbol{x} + \mathrm{T}\boldsymbol{y}] = (\mathrm{ST})\,\boldsymbol{x} + (\mathrm{ST})\,\boldsymbol{y} = (\mathrm{S} \circ \mathrm{T})\,\boldsymbol{x} + (\mathrm{S} \circ \mathrm{T})\,\boldsymbol{y}$. To see $\mathrm{S} \circ \mathrm{T}$ is homogenous, take any $k \in K$ and $\boldsymbol{x} \in X$. Then $(\mathrm{S} \circ \mathrm{T})(k\boldsymbol{x}) = \mathrm{S}\big[\mathrm{T}(k\boldsymbol{x})\big] = \mathrm{S}(k\mathrm{T}\boldsymbol{x}) = k\mathrm{ST}\boldsymbol{x} = k(\mathrm{S} \circ \mathrm{T})\,\boldsymbol{x}$.

b. Let

$$V \underset{\mathrm{S}}{\overset{\mathrm{R}}{\leftarrow}} U \underset{\mathrm{T}}{\overset{\mathrm{P}}{\leftarrow}} X.$$

For any $\boldsymbol{x} \in X$, we have $\big[(\mathrm{R} + \mathrm{S}) \circ \mathrm{T}\big](\boldsymbol{x}) = (\mathrm{R} + \mathrm{S})(\mathrm{T}\boldsymbol{x}) = \mathrm{R}(\mathrm{T}\boldsymbol{x}) + \mathrm{S}(\mathrm{T}\boldsymbol{x}) = (\mathrm{R} \circ \mathrm{T})\,\boldsymbol{x} + (\mathrm{S} \circ \mathrm{T})\,\boldsymbol{x}$. The other claim is proved similarly.

□

► EXERCISE 164 (3.7). *Show that whenever meaningful,*

$$(\mathrm{ST})' = \mathrm{T}'\mathrm{S}', \quad (\mathrm{T}+\mathrm{R})' = \mathrm{T}' + \mathrm{R}', \quad \textit{and} \quad \left(\mathrm{T}^{-1}\right)' = \left(\mathrm{T}'\right)^{-1}.$$

PROOF. For a generic linear mapping $\mathrm{T} \in \mathfrak{L}(X, U)$, we have the following diagram:

$$X \xrightarrow{\mathrm{T}} U,$$
$$X' \xleftarrow{\mathrm{T}'} U'.$$

For the first equality[1], let $\mathbb{F} \xleftarrow{\ell} V \xleftarrow{\mathrm{S}} U \xleftarrow{\mathrm{T}} X$, i.e., $\mathbb{F} \xleftarrow{\ell} V \xleftarrow{\mathrm{ST}} X$, and so $(\mathrm{ST})' : V' \to X'$. We have

$$\left\langle \left(\mathrm{T}'\mathrm{S}'\right)\ell, \boldsymbol{x} \right\rangle = \left\langle \mathrm{T}'\left(\mathrm{S}'\ell\right), \boldsymbol{x} \right\rangle = \left\langle \mathrm{S}'\ell, \mathrm{T}\boldsymbol{x} \right\rangle = \langle \ell, (\mathrm{ST})\,\boldsymbol{x} \rangle = \left\langle (\mathrm{ST})'\,\ell, \boldsymbol{x} \right\rangle,$$

and this establish the first equality. As for the second equality,

$$\begin{aligned}
\left\langle \left(\mathrm{T}' + \mathrm{S}'\right)\ell, \boldsymbol{x} \right\rangle &= \left\langle \mathrm{T}'\ell + \mathrm{S}'\ell, \boldsymbol{x} \right\rangle \\
&= \left\langle \mathrm{T}'\ell, \boldsymbol{x} \right\rangle + \left\langle \mathrm{S}'\ell, \boldsymbol{x} \right\rangle \\
&= \langle \ell, \mathrm{T}\boldsymbol{x} \rangle + \langle \ell, \mathrm{S}\boldsymbol{x} \rangle \\
&= \langle \ell, (\mathrm{T}+\mathrm{S})\,\boldsymbol{x} \rangle \\
&= \left\langle (\mathrm{T}+\mathrm{S})'\,\ell, \boldsymbol{x} \right\rangle.
\end{aligned}$$

Finally, let $\mathbb{F} \xleftarrow{\ell} U \xleftarrow{\mathrm{T}} X$, then $\mathrm{T}' : U' \to X'$ and $\left(\mathrm{T}'\right)^{-1} : \mathfrak{R}_{\mathrm{T}'} \to U'$. Take any $m \in \mathfrak{R}_{\mathrm{T}'}$; then there exists $\ell \in U'$ such that $\left(\mathrm{T}'\right)^{-1}(m) = \ell$, or equivalently, $\mathrm{T}'\ell = m$. Now consider $\left(\mathrm{T}^{-1}\right)'(m)$. Then

$$\left(\mathrm{T}^{-1}\right)'(m) = \left(\mathrm{T}^{-1}\right)'\left(\mathrm{T}'\ell\right) = \left(\mathrm{T}\mathrm{T}^{-1}\right)'\ell = \mathrm{Id}'\ell = \ell$$

since $\mathrm{Id}' = \mathrm{Id}$. □

► EXERCISE 165 (3.8). *Show that if X'' is identified with X and U'' with U, then* $\mathrm{T}'' = \mathrm{T}$.

PROOF. We have

$$\langle \mathrm{T}\ell, \boldsymbol{x} \rangle = \left\langle \ell, \mathrm{T}'\boldsymbol{x} \right\rangle = \left\langle \mathrm{T}''\ell, \boldsymbol{x} \right\rangle.$$

□

► EXERCISE 166 (3.9). *Show that if* $\mathrm{A} \in \mathfrak{L}(X)$ *is a left inverse of* $\mathrm{B} \in \mathfrak{L}(X)$*, that is,* $\mathrm{A}\,\mathrm{B} = \mathrm{Id}$*, then it is also a right inverse:* $\mathrm{B}\,\mathrm{A} = \mathrm{Id}$.

[1] Notation Warning: We occasionally use $\mathbb{F}$, instead of K, to denote the field. From now on we also be back to Lax's notation of linear mapping, that is, $\ell(\boldsymbol{x}) = \langle \ell, \boldsymbol{x} \rangle$.

PROOF. $AB = \mathrm{Id} \Longrightarrow ABA = A \Longrightarrow A(BA) = A \Longrightarrow BA = \mathrm{Id}.$ □

► EXERCISE 167 (3.10). *Show that if* M *is invertible, and similar to* K*, then* K *also is invertible, and* K^{-1} *is similar to* M^{-1}.

PROOF. M similar to K means that $K = SMS^{-1}$; then

$$K^{-1} = \left[(SM)S^{-1}\right]^{-1} = S(SM)^{-1} = SM^{-1}S^{-1},$$

and

$$M = S^{-1}KS \Longrightarrow M^{-1} = S^{-1}K^{-1}S.$$

□

► EXERCISE 168 (3.11). *If either* A *or* B *in* $\mathcal{L}(X)$ *is invertible, then* AB *and* BA *are similar.*

PROOF. Suppose B is invertible. Then

$$B(AB)B^{-1} = BA,$$

i.e., AB similar to BA. □

► EXERCISE 169 (3.14). *Suppose* T *is a linear map of rank 1 of a finite dimensional vector space into itself.*

a. *Show there exists a unique number* c *such that* $T^2 = cT$.

b. *Show that if* $c \neq 1$ *then* $\mathrm{Id} - T$ *has an inverse.*

PROOF. Let $T \in \mathcal{L}(X)$. By definition rank (T) = 1 means that $\dim \mathcal{R}_T = 1$. Let $\dim X = n$. Then

$$\dim X = \dim \mathcal{N}_T + \dim \mathcal{R}_T$$

implies that

$$\dim \mathcal{N}_T = n - 1.$$

Let $(\boldsymbol{v})$, where $\boldsymbol{v} \neq \mathbf{0}_X$, be a basis of $\mathcal{R}_T$, and extend it to a basis $(\boldsymbol{v}, \boldsymbol{u}_1, \dots, \boldsymbol{u}_{n-1})$ of X. Since $\boldsymbol{u}_i \notin \operatorname{span}(\boldsymbol{v})$ for all $\boldsymbol{u}_i$, we have $T\boldsymbol{u}_i = \mathbf{0}_X$; since $\mathcal{R}_T = \operatorname{span}(\boldsymbol{v})$, there exists $c \in K$ such that $T\boldsymbol{v} = a\boldsymbol{v}$. For any $\boldsymbol{x} \in X$, there exists a list of scalars $(b, k_1, \dots, k_{n-1})$ such that $\boldsymbol{x} = b\boldsymbol{v} + \sum_{i=1}^{n-1} k_i \boldsymbol{u}_i$. Then

$$T\boldsymbol{x} = bT\boldsymbol{v} = b(c\boldsymbol{v}) = cb\boldsymbol{v},$$

and

$$T^2\boldsymbol{x} = T(T\boldsymbol{x}) = T(cb\boldsymbol{v}) = cbT\boldsymbol{v} = c^2b\boldsymbol{v} = c(cb\boldsymbol{v}) = cT\boldsymbol{x}.$$

Since the above display holds for any $\boldsymbol{x} \in X$, we have $T^2 = cT$. □

► EXERCISE 170 (3.15[2][3]). *Suppose* T *and* S *are linear maps of a finite dimensional vector space into itself. Show that* $\operatorname{rank}(\mathrm{ST}) \leqslant \operatorname{rank}(\mathrm{S})$. *Show that* $\dim \mathcal{N}_{\mathrm{ST}} \leqslant \dim \mathcal{N}_{\mathrm{S}} + \dim \mathcal{N}_{\mathrm{T}}$.

PROOF. Let $X \overset{\mathrm{S}}{\leftarrow} X \overset{\mathrm{T}}{\leftarrow} X$. By definition, $\operatorname{rank}(\mathrm{ST}) \leqslant \operatorname{rank}(\mathrm{S})$ if and only if $\dim \mathcal{R}_{\mathrm{ST}} \leqslant \dim \mathcal{R}_{\mathrm{S}}$. But this is obvious. As for the second claim, we have

$$\mathcal{R}_{\mathrm{ST}} = (\mathrm{ST})[X] = \mathrm{S}\big[\mathrm{T}[X]\big] = \mathrm{S}[\mathcal{R}_{\mathrm{T}}],$$

so that

$$\operatorname{rank}(\mathrm{ST}) = \dim \mathcal{R}_{\mathrm{ST}} = \dim \mathrm{S}[\mathcal{R}_{\mathrm{T}}].$$

If M is a subspace of dimension m, say, and if N is any complement of M so that $X = M + N$, then[4]

$$\mathcal{R}_{\mathrm{S}} = \mathrm{S}[X] = \mathrm{S}[M] + \mathrm{S}[N].$$

It follows that

$$\operatorname{rank}(\mathrm{S}) = \dim \mathcal{R}_{\mathrm{S}} \leqslant \dim \mathrm{S}[M] + \dim \mathrm{S}[N] \leqslant \dim \mathrm{S}[M] + \dim N,$$

and hence that

$$\dim X - \mathcal{N}_{\mathrm{S}} \leqslant \dim \mathrm{S}[M] + \dim X - m.$$

If in particular

$$M = \mathcal{R}_{\mathrm{T}} = \mathrm{T}[X],$$

then the last inequality implies that

$$\operatorname{rank}(\mathrm{T}) - \mathcal{N}_{\mathrm{S}} \leqslant \operatorname{rank}(\mathrm{ST}),$$

or, equivalently, that

$$\dim X - \dim \mathcal{N}_{\mathrm{S}} - \dim \mathcal{N}_{\mathrm{T}} \leqslant \dim X - \dim \mathcal{N}_{\mathrm{ST}},$$

that is,

$$\dim \mathcal{N}_{\mathrm{ST}} \leqslant \dim \mathcal{N}_{\mathrm{S}} + \dim \mathcal{N}_{\mathrm{T}}.$$

□

[2] See Exercise 48.

[3] See Halmos (1995, Problem 95, p. 270).

[4] Theorem 1.5 (b).

11

MATRICES

12

DETERMINANT AND TRACE

► EXERCISE 171 (5.1). *Prove the properties of signature:*[1]

$$\operatorname{sign}(\pi) = \pm 1. \tag{5-a}$$

$$\operatorname{sign}(\pi \circ \sigma) = \operatorname{sign}(\pi)\operatorname{sign}(\sigma). \tag{5-b}$$

PROOF. The discriminant of $(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n)$ is $P(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) = \prod_{i<j}\left(\boldsymbol{x}_i - \boldsymbol{x}_j\right)$. Thus

$$(P\pi)(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) = \prod_{i<j}\left(\boldsymbol{x}_{\pi_i} - \boldsymbol{x}_{\pi_j}\right).$$

A typical factor in $P\pi$ is $\boldsymbol{x}_{\pi_i} - \boldsymbol{x}_{\pi_j}$. Now if $\pi_i < \pi_j$, this is also a factor of P, while if $\pi_i > \pi_j$, then $-\left(\boldsymbol{x}_{\pi_i} - \boldsymbol{x}_{\pi_j}\right)$ is a factor of P. Consequently, $P\pi = +P$ if the number of inversions of the natural order in π is even and $P\pi = -P$ if it is odd. Then (5-a) holds since

$$\operatorname{sign}(\pi) = \frac{P\pi}{P} = \pm 1.$$

We not prove (5-b). Let $P = \prod_{i<j}\left(\boldsymbol{x}_i - \boldsymbol{x}_j\right)$. Then, since $P\pi = \operatorname{sign}(\pi) P$, we have

$$\begin{aligned}(P\pi\sigma)(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) &= \left[(P\pi)\sigma\right](\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \\ &= \operatorname{sign}(\sigma)(P\pi)(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n) \\ &= \operatorname{sign}(\sigma)\operatorname{sign}(\pi) P(\boldsymbol{x}_1, \dots, \boldsymbol{x}_n).\end{aligned}$$

But $P\pi\sigma = \sigma\pi\sigma P$. Hence, $\operatorname{sign}(\pi\sigma) = \operatorname{sign}(\pi)\operatorname{sign}(\sigma)$. □

► EXERCISE 172 (5.2). *Prove that transposition has the following properties:*

a. *The signature of a transposition t is minus one:*

$$\operatorname{sign}(t) = -1. \tag{5-c}$$

[1] See Robinson (2003, Sec. 3.1) for a detailed discussion of permutation, signature function, and so on.

b. *Every permutation π can be written as a composition of transpositions:*

$$\pi = t_k \circ \cdots \circ t_1. \tag{5-d}$$

PROOF. (5-c) is clear. For (5-d), see Robinson (2003, 3.1.3 & 3.1.4, p. 34-35). □

References

[1] Axler, Sheldon (1997) *Linear Algebra Done Right*, Undergraduate Texts in Mathematics, New York: Springer-Verlag, 2nd edition. [i, v, 1, 38]

[2] Halmos, Paul R. (1974) *Finite-Dimensional Vector Spaces*, Undergraduate Texts in Mathematics, New York: Springer-Verlag. [69]

[3] ——— (1995) *Linear Algebra Problem Book*, **16** of The Dolciani Mathematical Expositions, Washington, DC: The Mathematical Association of America. [19, 76]

[4] Lax, Peter D. (2007) *Linear Algebra and Its Applications*, Pure and Applied Mathematics: A Wiley-Interscience Series of Texts, Monographs and Tracts, New Jersey: Wiley-Interscience, 2nd edition. [i, v, 59]

[5] Robinson, Derek J. S. (2003) *An Introduction to Abstract Algebra*, Berlin: Walter de Gruyter. [79, 80]

[6] Rockafellar, R. Tyrrell (1970) *Convex Analysis*, Princeton Mathematical Series, New Jersey: Princeton University Press, 2nd edition. [68]

[7] Roman, Steven (2008) *Advanced Linear Algebra*, **135** of Graduate Texts in Mathematics, New York: Springer Science+Business Media, LLC, 3rd edition. [i, 5]

General Topology

A Solution Manual for Willard (2004)

Jianfei Shen

School of Economics, The University of New South Wales

Sydney, Australia October 15, 2011

Preface

Sydney, *Jianfei Shen*
October 15, 2011

Acknowledgements

Contents

Acronyms

$\mathbb{R}$	the set of real numbers
$\mathbb{I}$	$[0, 1]$
$\mathbb{P}$	$\mathbb{R} \smallsetminus \mathbb{Q}$

1

SET THEORY AND METRIC SPACES

1.1 Set Theory

1A. Russell's Paradox

► Exercise 1. *The phenomenon to be presented here was first exhibited by Russell in 1901, and consequently is known as* Russell's Paradox.

Suppose we allow as sets things A for which $A \in A$. Let $\mathcal{P}$ be the set of all sets. Then $\mathcal{P}$ can be divided into two nonempty subsets, $\mathcal{P}_1 = \{A \in \mathcal{P} : A \notin A\}$ and $\mathcal{P}_2 = \{A \in \mathcal{P} : A \in A\}$. Show that this results in the contradiction: $\mathcal{P}_1 \in \mathcal{P}_1 \iff \mathcal{P}_1 \notin \mathcal{P}_1$. Does our (naive) restriction on sets given in 1.1 eliminate the contradiction?

Proof. If $\mathcal{P}_1 \in \mathcal{P}_1$, then $\mathcal{P}_1 \in \mathcal{P}_2$, i.e., $\mathcal{P}_1 \notin \mathcal{P}_1$. But if $\mathcal{P}_1 \notin \mathcal{P}_1$, then $\mathcal{P}_1 \in \mathcal{P}_1$. A contradiction. □

1B. De Morgan's laws and the distributive laws

► Exercise 2. a. $A \smallsetminus \left(\bigcap_{\lambda \in \Lambda} B_\lambda\right) = \bigcup_{\lambda \in \Lambda} (A \smallsetminus B_\lambda)$.

b. $B \cup \left(\bigcap_{\lambda \in \Lambda} B_\lambda\right) = \bigcap_{\lambda \in \Lambda} (B \cup B_\lambda)$.

c. *If A_{nm} is a subset of A for $n = 1, 2, \ldots$ and $m = 1, 2, \ldots$, is it necessarily true that*

$$\bigcup_{n=1}^{\infty}\left[\bigcap_{m=1}^{\infty} A_{nm}\right] = \bigcap_{m=1}^{\infty}\left[\bigcup_{n=1}^{\infty} A_{nm}\right]?$$

Proof. **(a)** If $x \in A \smallsetminus \left(\bigcap_{\lambda \in \Lambda} B_\lambda\right)$, then $x \in A$ and $x \notin \bigcap_{\lambda \in \Lambda} B_\lambda$; thus, $x \in A$ and $x \notin B_\lambda$ for some λ, so $x \in (A \smallsetminus B_\lambda)$ for some λ; hence $x \in \bigcup_{\lambda \in \Lambda} (A \smallsetminus B_\lambda)$. On the other hand, if $x \in \bigcup_{\lambda \in \Lambda} (A \smallsetminus B_\lambda)$, then $x \in A \smallsetminus B_\lambda$ for some $\lambda \in \Lambda$, i.e., $x \in A$ and $x \notin B_\lambda$ for some $\lambda \in \Lambda$. Thus, $x \in A$ and $x \notin \bigcap_{\lambda \in \Lambda} B_\lambda$; that is, $x \in A \smallsetminus \left(\bigcap_{\lambda \in \Lambda} B_\lambda\right)$.

(b) If $x \in B \cup \left(\bigcap_{\lambda\in\Lambda} B_\lambda\right)$, then $x \in B_\lambda$ for all λ, then $x \in (B \cup B_\lambda)$ for all λ, i.e., $x \in \bigcap_{\lambda\in\Lambda}(B \cup B_\lambda)$. On the other hand, if $x \in \bigcap_{\lambda\in\Lambda}(B \cup B_\lambda)$, then $x \in (B \cup B_\lambda)$ for all λ, i.e., $x \in B$ or $x \in B_\lambda$ for all λ; that is, $x \in B \cup \left(\bigcap_{\lambda\in\Lambda} B_\lambda\right)$.

(c) They are one and the same set. □

1C. Ordered pairs

► EXERCISE 3. *Show that, if (x_1, x_2) is defined to be $\{\{x_1\}, \{x_1, x_2\}\}$, then $(x_1, x_2) = (y_1, y_2)$ iff $x_1 = y_1$ and $x_2 = y_2$.*

PROOF. If $x_1 = y_1$ and $x_2 = y_2$, then, clearly, $(x_1, x_2) = \{\{x_1\}, \{x_1, x_2\}\} = \{\{y_1\}, \{y_1, y_2\}\} = (y_1, y_2)$. Now assume that $\{\{x_1\}, \{x_1, x_2\}\} = \{\{y_1\}, \{y_1, y_2\}\}$. If $x_1 \neq x_2$, then $\{x_1\} = \{y_1\}$ and $\{x_1, x_2\} = \{y_1, y_2\}$. So, first, $x_1 = y_1$ and then $\{x_1, x_2\} = \{y_1, y_2\}$ implies that $x_2 = y_2$. If $x_1 = x_2$, then $\{\{x_1\}, \{x_1, x_1\}\} = \{\{x_1\}\}$. So $\{y_1\} = \{y_1, y_2\} = \{x_1\}$, and we get $y_1 = y_2 = x_1$, so $x_1 = y_1$ and $x_2 = y_2$ holds in this case, too. □

1D. Cartesian products

► EXERCISE 4. *Provide an inductive definition of "the ordered n-tuple $(x_1, \dots, x_n)$ of elements $x_1, \dots, x_n$ of a set" so that $(x_1, \dots, x_n)$ and $(y_1, \dots, y_n)$ are equal iff their coordinates are equal in order, i.e., iff $x_1 = y_1, \dots, x_n = y_n$.*

PROOF. Define $(x_1, \dots, x_n) = \{(1, x_1), \dots, (n, x_n)\}$ as a finite sequence. □

► EXERCISE 5. *Given sets $X_1, \dots, X_n$ define the Cartesian product $X_1 \times \cdots \times X_n$*

a. *by using the definition of ordered n-tuple you gave in Exercise 4,*

b. *inductively from the definition of the Cartesian product of two sets,*

and show that the two approaches are the same.

PROOF. **(a)** $X_1 \times \cdots \times X_n = \{f \in (\bigcup_{i=1}^n X_i)^n : f(i) \in X_i\}$.

(b) From the definition of the Cartesian product of two sets, $X_1 \times \cdots \times X_n = \{(x_1, \dots, x_n) : x_i \in X_i\}$, where $(x_1, \dots, x_n) = ((x_1, \dots, x_{n-1}), x_n)$.

These two definitions are equal essentially since there is a bijection between them. □

► EXERCISE 6. *Given sets $X_1, \dots, X_n$ let $X = X_1 \times \cdots \times X_n$ and let X^* be the set of all functions f from $\{1, \dots, n\}$ into $\bigcup_{k=1}^n X_k$ having the property that $f(k) \in X_k$ for each $k = 1, \dots, n$. Show that X^* is the "same" set as X.*

PROOF. Each function f can be written as $\{(1, x_1), \dots, (n, x_n)\}$. So define $F : X^* \to X$ as $F(f) = (x_1, \dots, x_n)$. □

► EXERCISE 7. *Use what you learned in Exercise 6 to define the Cartesian product $X_1 \times X_2 \times \cdots$ of denumerably many sets as a collection of certain functions with domain $\mathbb{N}$.*

PROOF. $X_1 \times X_2 \times \cdots$ consists of functions $f : \mathbb{N} \to \bigcup_{n=1}^{\infty} X_n$ such that $f(n) \in X_n$ for all $n \in \mathbb{N}$. □

1.2 METRIC SPACES

2A. Metrics on $\mathbb{R}^n$

► EXERCISE 8. *Verify that each of the following is a metric on $\mathbb{R}^n$:*

a. $\rho(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$.

b. $\rho_1(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{n} |x_i - y_i|$.

c. $\rho_2(\boldsymbol{x}, \boldsymbol{y}) = \max\{|x_1 - y_1|, \ldots, |x_n - y_n|\}$.

PROOF. Clearly, it suffices to verify the triangle inequalities for all of the three functions. Pick arbitrary $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^n$.

(a) By Minkowski's Inequality, we have

$$\begin{aligned}\rho(\boldsymbol{x}, \boldsymbol{z}) = \sqrt{\sum_{i=1}^{n} (x_i - z_i)^2} &= \sqrt{\sum_{i=1}^{n} [(x_i - y_i) + (y_i - z_i)]^2} \\ &\leqslant \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2} + \sqrt{\sum_{i=1}^{n} (y_i - z_i)^2} \\ &= \rho(\boldsymbol{x}, \boldsymbol{y}) + \rho(\boldsymbol{y}, \boldsymbol{z}).\end{aligned}$$

(b) We have

$$\rho_1(\boldsymbol{x}, \boldsymbol{z}) = \sum_{i=1}^{n} |x_i - z_i| = \sum_{i=1}^{n} (|x_i - y_i| + |y_i - z_i|) = \rho_1(\boldsymbol{x}, \boldsymbol{y}) + \rho_1(\boldsymbol{y}, \boldsymbol{z}).$$

(c) We have

$$\begin{aligned}\rho_2(\boldsymbol{x}, \boldsymbol{z}) &= \max\{|x_1 - z_1|, \ldots, |x_n - z_n|\} \\ &\leqslant \max\{|x_1 - y_1| + |y_1 - z_1|, \ldots, |x_n - y_n| + |y_n - z_n|\} \\ &\leqslant \max\{|x_1 - y_1|, \ldots, |x_n - y_n|\} + \max\{|y_1 - z_1|, \ldots, |y_n - z_n|\} \\ &= \rho_2(\boldsymbol{x}, \boldsymbol{y}) + \rho_2(\boldsymbol{y}, \boldsymbol{z}).\end{aligned}$$

□

2B. Metrics on $\mathscr{C}(\mathbb{I})$

► EXERCISE 9. *Let $\mathscr{C}(\mathbb{I})$ denote the set of all continuous real-valued functions on the unit interval $\mathbb{I}$ and let x_0 be a fixed point of $\mathbb{I}$.*

a. $\rho(f,g) = \sup_{x\in\mathbb{I}} |f(x) - g(x)|$ *is a metric on* $\mathscr{C}(\mathbb{I})$.

b. $\sigma(f,g) = \int_0^1 |f(x) - g(x)|\,\mathrm{d}x$ *is a metric on* $\mathscr{C}(\mathbb{I})$.

c. $\eta(f,g) = |f(x_0) - g(x_0)|$ *is a pseudometric on* $\mathscr{C}(\mathbb{I})$.

PROOF. Let $f,g,h \in \mathscr{C}(\mathbb{I})$. It is clear that ρ, σ, and η are positive, symmetric; it is also clear that ρ and σ satisfy M-b.

(a) We have

$$\begin{aligned}\rho(f,h) = \sup_{x\in\mathbb{I}} |f(x) - h(x)| &\leqslant \sup_{x\in\mathbb{I}}(|f(x) - g(x)| + |g(x) - h(x)|)\\ &\leqslant \sup_{x\in\mathbb{I}} |f(x) - g(x)| + \sup_{x\in\mathbb{I}} |g(x) - h(x)|\\ &= \rho(f,g) + \rho(g,h).\end{aligned}$$

(b) We have

$$\begin{aligned}\sigma(f,h) = \int_0^1 |f(x) - h(x)| &\leqslant \int_0^1 |f(x) - g(x)| + \int_0^1 |g(x) - h(x)|\\ &= \sigma(f,g) + \sigma(g,h).\end{aligned}$$

(c) For arbitrary $f,g \in \mathscr{C}(\mathbb{I})$ with $f(x_0) = g(x_0)$ we have $\eta(f,g) = 0$, so $\eta(f,g) = 0$ does not imply that $f = g$. Further, $\eta(f,h) = |f(x_0) - h(x_0)| \leqslant |f(x_0) - g(x_0)| + |g(x_0) - h(x_0)| = \eta(f,g) + \eta(g,h)$. □

2C. Pseudometrics

► EXERCISE 10. *Let (M,ρ) be a pseudometric space. Define a relation $\sim$ on M by $x \sim y$ iff $\rho(x,y) = 0$. Then $\sim$ is an equivalence relation.*

PROOF. (i) $x \sim x$ since $\rho(x,x) = 0$ for all $x \in M$. (ii) $x \sim y$ iff $\rho(x,y) = 0$ iff $\rho(y,x) = 0$ iff $y \sim x$. (iii) Suppose $x \sim y$ and $y \sim z$. Then $\rho(x,z) \leqslant \rho(x,y) + \rho(y,z) = 0$; that is, $\rho(x,z) = 0$. So $x \sim z$. □

► EXERCISE 11. *If M^* is he set of equivalence classes in M under the equivalence relation $\sim$ and if ρ^* is defined on M^* by $\rho^*([x],[y]) = \rho(x,y)$, then ρ^* is a well-defined metric on M^*.*

PROOF. ρ^* is well-defined since it does not dependent on the representative of $[x]$: let $x' \in [x]$ and $y' \in [y]$. Then

$$\rho(x',y') \leqslant \rho(x',x) + \rho(x,y) + \rho(y,y') = \rho(x,y).$$

Symmetrically, $\rho(x, y) \leqslant \rho(x', y')$. To verify ρ^* is a metric on M^*, it suffices to show that ρ^* satisfies the triangle inequality. Let $[x], [y], [z] \in M^*$. Then

$$\rho^*([x], [z]) = \rho(x, z) \leqslant \rho(x, y) + \rho(y, z) = \rho^*([x], [y]) + \rho^*([y], [z]). \qquad \square$$

► EXERCISE 12. *If $h\colon M \to M^*$ is the mapping $h(x) = [x]$, then a set A in M is closed (open) iff $h(A)$ is closed (open) in M^*.*

PROOF. Let A be open in M and $h(x) = [x] \in h(A)$ for some $x \in A$. Since A is open, there exist an ε-disk $U_\rho(x, \varepsilon)$ contained in A. For each $y \in U_\rho(x, \varepsilon)$, we have $h(y) = [y] \in h(A)$, and $\rho^*([x], [y]) = \rho(x, y) \leqslant \varepsilon$. Hence, for each $[x] \in h(A)$, there exists an ε-disk $U_{\rho^*}([x], \varepsilon) = h(U_\rho(x, \varepsilon))$ contained in $h(A)$; that is, $h(A)$ is open in M^*. Since h is surjective, it is now easy to see that $h(A)$ is closed in M^* whenever A is closed in M. □

► EXERCISE 13. *If f is any real-valued function on a set M, then the distance function $\rho_f(x, y) = |f(x) - f(y)|$ is a pseudometric on M.*

PROOF. Easy. □

► EXERCISE 14. *If (M, ρ) is any pseudometric space, then a function $f : M \to \mathbb{R}$ is continuous iff each set open in (M, ρ_f) is open in (M, ρ).*

PROOF. Suppose that f is continuous and G is open in (M, ρ_f). For each $x \in G$, there is an $\varepsilon > 0$ such that if $|f(y) - f(x)| < \varepsilon$ then $y \in G$. The continuity of f at x implies that there exists $\delta > 0$ such that if $\rho(y, x) < \delta$ then $|f(y) - f(x)| < \varepsilon$, and so $y \in U$. We thus proved that for each $x \in U$ there exists a δ-disk $U_\rho(x, \rho)$ contained in G; that is, G is open in (M, ρ).

Conversely, suppose that each set is open in (M, ρ) whenever it is open in (M, ρ_f). For each $x \in (M, \rho_f)$, there is an ε-disk $U_{\rho_f}(x, \varepsilon)$ contained in M since M is open under ρ_f; then $U_{\rho_f}(x, \varepsilon)$ is open in (M, ρ) since $U_{\rho_f}(x, \varepsilon)$ is open in (M, ρ_f). Hence, there is an δ-disk $U_\rho(x, \delta)$ such that $U_\rho(x, \delta) \subset U_{\rho_f}(x, \varepsilon)$; that is, if $\rho(y, x) < \delta$, then $|f(y) - f(x)| < \varepsilon$. So f is continuous on M. □

2D. Disks Are Open

► EXERCISE 15. *For any subset A of a metric space M and any $\varepsilon > 0$, the set $U(A, \varepsilon)$ is open.*

PROOF. Let $A \subset M$ and $\varepsilon > 0$. Take an arbitrary point $x \in U(A, \varepsilon)$; take an arbitrary point $y \in A$ such that $\rho(x, y) < \varepsilon$. Observe that every ε-disk $U(y, \varepsilon)$ is contained in $U(A, \varepsilon)$. Since $x \in U(y, \varepsilon)$ and $U(y, \varepsilon)$ is open, there exists an δ-disk $U(x, \delta)$ contained in $U(y, \varepsilon)$. Therefore, $U(A, \varepsilon)$ is open. □

2E. Bounded Metrics

► EXERCISE 16. *If ρ is any metric on M, the distance function $\rho^*(x,y) = \min\{\rho(x,y,),1\}$ is a metric also and is bounded.*

PROOF. To see ρ^* is a metric, it suffices to show the triangle inequality. Let $x,y,z \in M$. Then

$$\begin{aligned}\rho^*(x,z) = \min\{\rho(x,z),1\} &\leqslant \min\{\rho(x,y)+\rho(y,z),1\}\\ &\leqslant \min\{\rho(x,y),1\}+\min\{\rho(y,z),1\}\\ &= \rho^*(x,y)+\rho^*(y,z).\end{aligned}$$

It is clear that ρ^* is bounded above by 1. □

► EXERCISE 17. *A function f is continuous on (M,ρ) iff it is continuous on (M,ρ^*).*

PROOF. It suffices to show that ρ and ρ^* are equivalent. If G is open in (M,ρ), then for each $x \in G$ there is an ε-disk $U_\rho(x,\varepsilon) \subset G$. Since $U_{\rho^*}(x,\varepsilon) \subset U_\rho(x,\varepsilon)$, we know G is open in (M,ρ^*). Similarly, we can show that G is open in (M,ρ^*) whenever it is open in (M,ρ). □

2F. The Hausdorff Metric

Let ρ be a bounded metric on M; that is, for some constant A, $\rho(x,y) \leqslant A$ for all x and y in M.

► EXERCISE 18. *Show that the elevation of ρ to the power set $\mathcal{P}(M)$ as defined in 2.4 is not necessarily a pseudometric on $\mathcal{P}(M)$.*

PROOF. Let $M := \{(x_1,x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leqslant 1\}$, and let ρ be the usual metric. Then ρ is a bounded metric on M. We show that the function $\rho^*\colon (E,F) \mapsto \inf_{x\in E, y\in F} \rho(x,y)$, for all $E,F \in \mathcal{P}(M)$, is not a pseudometric on $\mathcal{P}(M)$ by showing that the triangle inequality fails. Let $E,F,G \in \mathcal{P}(M)$, where $E = U_\rho((-1/4,0),1/4)$, $G = U_\rho((1/4,0),1/4)$, and F meets both E and G. Then $\rho^*(E,G) > 0$, but $\rho^*(E,F) = \rho^*(F,G) = 0$. □

► EXERCISE 19. *Let $\mathcal{F}(M)$ be all nonempty closed subsets of M and for $A,B \in \mathcal{F}(M)$ define*

$$\begin{aligned}d_A(B) &= \sup\{\rho(A,x) : x \in B\}\\ d(A,B) &= \max\{d_A(B), d_B(A)\}.\end{aligned}$$

Then d is a metric on $\mathcal{F}(M)$ with the property that $d(\{x\},\{y\}) = \rho(x,y)$. It is called the Hausdorff metric *on $\mathcal{F}(M)$.*

PROOF. Clearly, d is nonnegative and symmetric. If $d(A,B)=0$, then $d_A(B)=d_B(A)=0$, i.e., $\sup_{y\in B}\rho(A,y)=\sup_{x\in A}\rho(B,x)=0$. But then $\rho(A,y)=0$ for all $y\in B$ and $\rho(B,x)=0$ for all $x\in A$. Since A is closed, we have $y\in A$ for all $y\in B$; that is, $B\subset A$. Similarly, $A\subset B$. Hence, $A=B$.

We next show the triangle inequality of d. Let $A,B,C\in\mathcal{F}(M)$. For an arbitrary point $a\in A$, take a point $b\in C$ such that $\rho(a,b)=\rho(B,a)$ (since B is closed, such a point exists). Then

$$\rho(a,b)\leqslant\sup_{x\in A}\rho(B,x)=d_B(A)\leqslant d(A,B).$$

For this $b\in B$, we take a point $c\in C$ such that $\rho(b,c)\leqslant d(B,C)$. Therefore,

$$\rho(a,c)\leqslant\rho(a,b)+\rho(b,c)\leqslant d(A,B)+d(B,C).$$

We thus proved that for every $a\in A$, there exists $c\in C$ (depends on a), such that $\rho(a,c)\leqslant d(A,B)+d(B,C)$. In particular, we have

$$\rho(a,C)=\inf_{z\in C}\rho(a,z)\leqslant d(A,B)+d(B,C).$$

Since the above inequality holds for all $a\in A$, we obtain

$$d_C(A)=\sup_{x\in A}\rho(a,C)\leqslant d(A,B)+d(B,C). \tag{1.1}$$

Similarly, for each $c\in C$ there exists $b\in B$ with $\rho(c,b)\leqslant d(B,C)$; for this b, there exists $a\in A$ with $\rho(a,b)\leqslant d(A,B)$. Hence $\rho(a,c)\leqslant d(A,B)+d(B,C)$ for all $c\in C$. The same argument shows that

$$d_A(C)\leqslant d(A,B)+d(B,C). \tag{1.2}$$

Combining (1.1) and (1.2) we get the desired result.

Finally, notice that $d_{\{x\}}(\{y\})=d_{\{y\}}(\{x\})=\rho(x,y)$; hence, $d(\{x\},\{y\})=\rho(x,y)$. □

► EXERCISE 20. *Prove that closed sets A and B are "close" in the Hausdorff metric iff they are "uniformly close"; that is, $d(A,B)<\varepsilon$ iff $A\subset U_\rho(B,\varepsilon)$ and $B\subset U_\rho(A,\varepsilon)$.*

PROOF. If $d(A,B)<\varepsilon$, then $\sup_{y\in B}\rho(A,y)=\rho_A(B)<\varepsilon$; that is, $\rho(A,y)<\varepsilon$ for all $y\in B$, so $B\subset U_\rho(A,\varepsilon)$. Similarly, $A\subset U_\rho(B,\varepsilon)$.

Conversely, if $A\subset U_\rho(B,\varepsilon)$, then $\rho(B,x)<\varepsilon$ for all $x\in A$. Since A is closed, we have $d_B(A)<\varepsilon$; similarly, $B\subset U_\rho(A,\varepsilon)$ implies that $d_A(B)<\varepsilon$. Hence, $d(A,B)<\varepsilon$. □

2G. Isometry

Metric spaces (M, ρ) and (N, σ) are *isometric* iff there is a one-one function f from M onto N such that $\rho(x, y) = \sigma(f(x), f(y))$ for all x and y in M; f is called an *isometry*.

► EXERCISE 21. *If f is an isometry from M to N, then both f and f^{-1} are continuous functions.*

PROOF. By definition, f is (uniformly) continuous on M: for every $\varepsilon > 0$, let $\delta = \varepsilon$; then $\rho(x, y) < \delta$ implies that $\sigma(f(x), f(y)) = \rho(x, y) < \varepsilon$.

On the other hand, for every $\varepsilon > 0$ and $y \in N$, pick the unique $f^{-1}(y) \in M$ (since f is bijective). For each $z \in N$ with $\sigma(y, z) < \varepsilon$, we must have $\rho(f^{-1}(y), f^{-1}(z)) = \sigma(f(f^{-1}(y)), f(f^{-1}(z))) = \sigma(y, z) < \varepsilon$; that is, f^{-1} is continuous. □

► EXERCISE 22. *$\mathbb{R}$ is not isometric to $\mathbb{R}^2$ (each with its usual metric).*

PROOF. Consider $S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$. Notice that there are only two points around $f^{-1}(0, 0)$ with distance 1. □

► EXERCISE 23. *$\mathbb{I}$ is isometric to any other closed interval in $\mathbb{R}$ of the same length.*

PROOF. Consider the function $f : \mathbb{I} \to [a, a + 1]$ defined by $f(x) = a + x$ for all $x \in \mathbb{I}$. □

2

TOPOLOGICAL SPACES

2.1 Fundamental Concepts

3A. Examples of Topologies

▶ Exercise 24. *If $\mathscr{F}$ is the collection of all closed, bounded subset of $\mathbb{R}$ (in its usual topology), together with $\mathbb{R}$ itself, then $\mathscr{F}$ is the family of closed sets for a topology on $\mathbb{R}$ strictly weaker than the usual topology.*

Proof. It is easy to see that $\mathscr{F}$ is a topology. Further, for instance, $(-\infty, 0]$ is a closed set of $\mathbb{R}$, but it is not in $\mathscr{F}$. □

▶ Exercise 25. *If $A \subset X$, show that the family of all subsets of X which contain A, together with the empty set $\varnothing$, is a topology on X. Describe the closure and interior operations. What topology results when $A = \varnothing$? when $A = X$?*

Proof. Let

$$\mathscr{E} = \{E \subset X : A \subset E\} \cup \{\varnothing\}.$$

Now suppose that $E_\lambda \in \mathscr{E}$ for each $\lambda \in \Lambda$. Then $A \subset \bigcup_\lambda E_\lambda \subset X$ and so $\bigcup E_\lambda \in \mathscr{E}$. The other postulates are easy to check.

For any set $B \subset X$, if $A \subset B$, then $B \in \mathscr{E}$ and so $B^\circ = B$; if not, then $B^\circ = \varnothing$.

If $A = \varnothing$, then $\mathscr{E}$ is the discrete topology; if $A = X$, then $\mathscr{E} = \{\varnothing, X\}$. □

3D. Regularly Open and Regularly Closed Sets

An open subset G in a topological space is *regular open* iff G is the interior of its closure. A closed subset is *regularly closed* iff it is the closure of its interior.

▶ Exercise 26. *The complement of a regularly open set is regularly closed and vice versa.*

Proof. Suppose G is regular open; that is, $G = (\overline{G})^\circ$. Then

$$X \setminus G = X \setminus (\overline{G})^\circ = \overline{X \setminus \overline{G}} = \overline{(X \setminus G)^\circ}.$$

Hence, $X \setminus G$ is regularly closed. If F is regular closed, i.e., $F = \overline{F^\circ}$, then

$$X \setminus F = X \setminus \overline{F^\circ} = (X \setminus F^\circ)^\circ = (\overline{X \setminus F})^\circ;$$

that is, $X \setminus F$ is regularly open. □

► Exercise 27. *There are open sets in* $\mathbb{R}$ *which are not regularly open.*

Proof. Consider $\mathbb{Q}$. We have $(\overline{\mathbb{Q}})^\circ = \mathbb{R}^\circ = \mathbb{R} \neq \mathbb{Q}$. So $\mathbb{Q}$ is not regularly open. □

► Exercise 28. *If A is any subset of a topological space, then* $\operatorname{int}(\operatorname{cl}(A))$ *is regularly open.*

Proof. Let A be a subset of a topological space X. We then have

$$\operatorname{int}(\operatorname{cl}(A)) \subset \operatorname{cl}(\operatorname{int}(\operatorname{cl}(A))) \implies \operatorname{int}(\operatorname{cl}(A)) = \operatorname{int}(\operatorname{int}(\operatorname{cl}(A))) \subset \operatorname{int}(\operatorname{cl}(\operatorname{int}(\operatorname{cl}(A)))),$$

and

$$\begin{aligned} \operatorname{int}(\operatorname{cl}(A)) \subset \operatorname{cl}(A) &\implies \operatorname{cl}(\operatorname{int}(\operatorname{cl}(A))) \subset \operatorname{cl}(\operatorname{cl}(A)) = \operatorname{cl}(A) \\ &\implies \operatorname{int}(\operatorname{cl}(\operatorname{int}(\operatorname{cl}(A)))) \subset \operatorname{int}(\operatorname{cl}(A)). \end{aligned}$$

Therefore, $\operatorname{int}(\operatorname{cl}(A)) = \operatorname{int}(\operatorname{cl}(\operatorname{int}(\operatorname{cl}(A))))$; that is, $\operatorname{int}(\operatorname{cl}(A))$ is regularly open. □

► Exercise 29. *The intersection, but not necessarily the union, of two regularly open sets is regularly open.*

Proof. Let A and B be two regularly open sets in a topological space X. Then

$$(\overline{A \cap B})^\circ \subset (\overline{A} \cap \overline{B})^\circ = (\overline{A})^\circ \cap (\overline{B})^\circ = A \cap B,$$

and

$$\begin{aligned} (\overline{A} \cap \overline{B})^\circ = (\overline{A})^\circ \cap (\overline{B})^\circ = A \cap B &\subset \overline{A \cap B} \\ \implies A \cap B = (\overline{A} \cap \overline{B})^\circ &= \left[(\overline{A} \cap \overline{B})^\circ\right]^\circ \subset (\overline{A \cap B})^\circ. \end{aligned}$$

Hence, $A \cap B = (\overline{A \cap B})^\circ$.

To see that the union of two regularly open sets is not necessarily regularly open, consider $A = (0,1)$ and $B = (1,2)$ in $\mathbb{R}$ with its usual topology. Then

$$(\overline{A \cup B})^\circ = [0,2]^\circ = (0,2) \neq A \cup B.$$ □

3E. Metrizable Spaces

Let X be a metrizable space whose topology is generated by a metric ρ.

► EXERCISE 30. *The metric 2ρ defined by $2\rho(x,y) = 2 \cdot \rho(x,y)$ generates the same topology on X.*

PROOF. Let $\mathcal{O}_\rho$ be the collection of open sets in (X,ρ), and let $\mathcal{O}_{2\rho}$ be the collection of open sets in $(X, 2\rho)$. If $O \in \mathcal{O}_\rho$, then for every $x \in O$, there exists an open ball $\mathbb{B}_\rho(x,\varepsilon) \subseteq O$; but then $\mathbb{B}_{2\rho}(x,\varepsilon/2) \subset O$. Hence, $O \in \mathcal{O}_{2\rho}$. Similarly, we can show that $\mathcal{O}_{2\rho} \subset \mathcal{O}_\rho$. In fact, ρ and 2ρ are equivalent metrics. □

► EXERCISE 31. *The closure of a set $E \subset X$ is given by $\overline{E} = \{y \in X : \rho(E,y) = 0\}$.*

PROOF. Denote $\widetilde{E} := \{y \in X : \rho(E,y) = 0\}$. We first show that $\widetilde{E}$ is closed (see Definition 2.5, p. 17). Take an arbitrary $x \in X$ such that for every $n \in \mathbb{N}$, there exists $y_n \in \widetilde{E}$ with $\rho(x,y_n) < 1/2n$. For each $y_n \in \widetilde{E}$, take $z_n \in E$ with $\rho(y_n,z_n) < 1/2n$. Then

$$\rho(x,z_n) \leqslant \rho(x,y_n) + \rho(y_n,z_n) < 1/n, \quad \text{for all } n \in \mathbb{N}.$$

Thus, $\rho(x,E) = 0$, i.e., $x \in \widetilde{E}$. Therefore, $\widetilde{E}$ is closed. It is clear that $E \subseteq \widetilde{E}$, and so $\overline{E} \subset \widetilde{E}$.

We next show that $\widetilde{E} \subseteq \overline{E}$. Take an arbitrary $x \in \widetilde{E}$ and a closed set K containing E. If $x \in X \smallsetminus K$, then $\rho(x,K) > 0$ (see Exercise 35). But then $\rho(x,E) > 0$ since $E \subset K$ and so

$$\inf_{y \in E} \rho(x,y) \geqslant \inf_{z \in K} \rho(x,z).$$

Hence, $\widetilde{E} \subset \overline{E}$. □

► EXERCISE 32. *The closed disk $U(x,\overline{\varepsilon}) = \{y : \rho(x,y) \leqslant \varepsilon\}$ is closed in X, but may not be the closure of the open disk $U(x,\varepsilon)$.*

PROOF. Fix $x \in X$. We show that the function $\rho(x,\cdot)\colon X \to \mathbb{R}$ is (uniformly) continuous. For any $y,z \in X$, the triangle inequality yields

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

Hence, for every $\varepsilon > 0$, take $\delta = \varepsilon$, and $\rho(x,\cdot)$ satisfies the ε-δ criterion. Therefore, $U(x,\overline{\varepsilon})$ is closed since $U(x,\overline{\varepsilon}) = \rho^{-1}(x,[0,\varepsilon])$ and $[0,\varepsilon]$ is closed in $\mathbb{R}$.

To see it is not necessary that $U(x,\overline{\varepsilon}) = \overline{U(x,\varepsilon)}$, consider $\varepsilon = 1$ and the usual metric on

$$\left\{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\right\} \cup \left\{(x,0) \in \mathbb{R}^2 : 0 \leqslant x \leqslant 1\right\};$$

see Figure 2.1. Observe that $(0,0) \notin U(x,1)$, but $(0,0) \in U(x,\overline{1})$. It follows from Exercise 31 that $(0,0) \notin \overline{U(x,1)}$. □

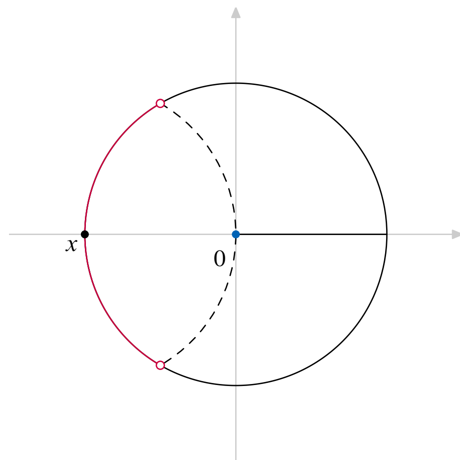


FIGURE 2.1. $U(x, \bar{1}) \neq \overline{U(x, 1)}$.

3H. G_δ and F_σ Sets

► EXERCISE 33. *The complement of a G_δ is an F_σ, and vice versa.*

PROOF. If A is a G_δ set, then there exists a sequence of open sets $\{U_n\}$ such that $A = \bigcap_{n=1}^{\infty} U_n$. Then $A^c = \bigcup_{n=1}^{\infty} U_n^c$ is F_σ. Vice versa. ☐

► EXERCISE 34. *An F_σ can be written as the union of an* increasing *sequence $F_1 \subset F_2 \subset \cdots$ of closed sets.*

PROOF. Let $B = \bigcup_{n=1}^{\infty} E_n$, where E_n is closed for all $n \in \mathbb{N}$. Define $F_1 = E_1$ and $F_n = \bigcup_{i=1}^{n} E_i$ for $n \geqslant 2$. Then each F_n is closed, $F_1 \subset F_2 \subset \cdots$, and $\bigcup_{n=1}^{\infty} F_n = \bigcup_{n=1}^{\infty} = B$. ☐

► EXERCISE 35. *A closed set in a metric space is a G_δ.*

PROOF. For an arbitrary set $A \subset X$ and a point $x \in X$, define

$$\rho(x, A) = \inf_{y \in A} \{\rho(x, y)\}.$$

We first show that $\rho(\cdot, A) \colon X \to \mathbb{R}$ is (uniformly) continuous by showing

$$|\rho(x, A) - \rho(y, A)| \leqslant \rho(x, y), \quad \text{for all } x, y \in X. \tag{2.1}$$

For an arbitrary $z \in A$, we have

$$\rho(x, A) \leqslant \rho(x, z) \leqslant \rho(x, y) + \rho(y, z).$$

Take the infimum over $z \in A$ and we get

$$\rho(x, A) \leqslant \rho(x, y) + \rho(y, A). \tag{2.2}$$

Symmetrically, we have

$$\rho(y, A) \leqslant \rho(x, y) + \rho(x, A). \tag{2.3}$$

Hence, (2.1) follows from (2.2) and (2.3). We next show that if A is closed, then $\rho(x, A) = 0$ iff $x \in A$. The "if" part is trivial, so we do the "only if" part. If $\rho(x, A) = 0$, then for every $n \in \mathbb{N}$, there exists $y_n \in A$ such that $\rho(x, y_n) < 1/n$; that is, $y_n \to x$. Since $\{y_n\} \subset A$ and A is closed, we must have $x \in A$.

Therefore,

$$A = \bigcap_{n=1}^{\infty} \{x \in X : \rho(x, A) < 1/n\}.$$

The continuity of $\rho(\cdot, A)$ implies that $\{x \in X : \rho(x, A) < 1/n\}$ is open for all n. Thus, A is a G_δ set. □

► EXERCISE 36. *The rationals are an F_σ in $\mathbb{R}$.*

PROOF. $\mathbb{Q}$ is countable, and every singleton set in $\mathbb{R}$ is closed; hence, $\mathbb{Q}$ is an F_σ. □

3I. Borel Sets

2.2 NEIGHBORHOODS

4A. The Sorgenfrey Line

► EXERCISE 37. *Verify that the set $[x, z)$, for $z > x$, do form a nhood base at x for a topology on the real line.*

PROOF. We need only check that for each $x \in \mathbb{R}$, the family $\mathcal{B}_x := \{[x, z) : z > x\}$ satisfies V-a, V-b, and V-c in Theorem 4.5. V-a is trivial. If $[x, z_1) \in \mathcal{B}_x$ and $[x, z_2) \in \mathcal{B}_x$, then $[x, z_1) \cap [x, z_2) = [x, z_1 \wedge z_2) \in \mathcal{B}_x$ and is in $[x, z_1) \cap [x, z_2)$. For V-c, let $[x, z) \in \mathcal{B}_x$. Let $z' \in (x, z]$. Then $\left[x, z'\right) \in \mathcal{B}_x$, and if $y \in \left[x, z'\right)$, the right-open interval $\left[y, z'\right) \in \mathcal{B}_y$ and $\left[y, z'\right) \subset [x, z)$.

Then, define open sets using V-d: $G \subset \mathbb{R}$ is open if and only if G contains a set $[x, z)$ of each of its points x. □

► EXERCISE 38. *Which intervals on the real line are open sets in the Sorgenfrey topology?*

SOLUTION.

- Sets of the form $(-\infty, x)$, $[x, z)$, or $[x, \infty)$ are both open and closed.
- Sets of the form (x, z) or $(x, +\infty)$ are open in $\mathbb{R}$, since

$$(x, z) = \bigcup \{[y, z) : x < y < z\}.$$ □

► EXERCISE 39. *Describe the closure of each of the following subset of the Sorgenfrey line: the rationals* $\mathbb{Q}$*, the set* $\{1/n : n = 1, 2, \dots, \}$*, the set* $\{-1/n : n = 1, 2, \dots\}$*, the integers* $\mathbb{Z}$*.*

SOLUTION. Recall that, by Theorem 4.7, for each $E \subset \mathbb{R}$, we have

$$\bar{E} = \left\{x \in \mathbb{R} : \text{each basic nhood of } x \text{ meets } E\right\}.$$

Then $\overline{\mathbb{Q}} = \mathbb{R}$ since for any $x \in \mathbb{R}$, we have $[x, z) \cap \mathbb{Q} \neq \varnothing$ for $z > x$. Similarly, $\overline{\{1/n : n = 1, 2, \dots\}} = \{1/n : n = 1, 2, \dots\}$, and $\overline{\mathbb{Z}} = \mathbb{Z}$. □

4B. The Moore Plane

► EXERCISE 40. *Verify that this gives a topology on* Γ*.*

PROOF. Verify (V-a)—(V-c). It is easy. □

4E. Topologies from nhoods

► EXERCISE 41. *Show that if each point* x *in a set* X *has assigned a collection* $\mathcal{U}_x$ *of subsets of* X *satisfying N-a through N-d of 4.2, then the collection*

$$\tau = \left\{G \subset X : \textit{for each } x \textit{ in } G, x \in U \subset G \textit{ for some } U \in \mathcal{U}_x\right\}$$

is a topology for X*, in which the nhood system at each* x *is just* $\mathcal{U}_x$*.*

PROOF. We need to check G1—G3 in Definition 3.1. Since G1 and G3 are evident, we focus on G2. Let $E_1, E_2 \in \tau$. Take any $x \in E_1 \cap E_2$. Then there exist some $U_1, U_2 \in \mathcal{U}_x$ such that $x \in U_1 \subset E_1$ and $x \in U_2 \subset E_2$. By N-b, we know that $U_1 \cap U_2 \in \mathcal{U}_x$. Hence,

$$x \in U_1 \cap U_2 \subset E_1 \cap E_2,$$

and so $E_1 \cap E_2 \in \tau$. The induction principle then means that τ is closed under finite intersections. □

4F. Spaces of Functions

► EXERCISE 42. *For each* $f \in \mathbb{R}^{\mathbb{I}}$*, each finite subset* F *of* $\mathbb{I}$ *and each positive* δ*, let*

$$U(f, F, \delta) = \left\{g \in \mathbb{R}^{\mathbb{I}} : |g(x) - f(x)| < \delta, \textit{ for each } x \in F\right\}.$$

Show that the sets $U(f, F, \delta)$ *form a nhood base at* f*, making* $\mathbb{R}^{\mathbb{I}}$ *a topological space.*

PROOF. Denote

$$\mathcal{B}_f = \left\{U(f, F, \delta) : F \subset \mathbb{I}, |F| < \infty, \delta > 0\right\}.$$

(*V*-a) For each $U(f, F, \delta) \in \mathcal{B}_f$, we have $|f(x) - f(x)| = 0 < \delta$ for all $x \in F$; hence, $f \in U(f, F, \delta)$.

(*V*-b) Let $U(f, F_1, \delta_1), U(f, F_2, \delta_2) \in \mathcal{B}_f$. Define $U(f, F_3, \delta_3)$ by letting

$$F_3 = F_1 \cup F_2, \quad \text{and} \quad \delta_3 = \min\{\delta_1, \delta_2\}.$$

Clearly, $U(f, F_3, \delta_3) \in \mathcal{B}_f$. If $g \in U(f, F_3, \delta_3)$, then

$$|g(x) - f(x)| < \min\{\delta_1, \delta_2\}, \quad \text{for all } x \in F_1 \cup F_2.$$

Hence, $|g(x) - f(x)| < \delta_1$ for all $x \in F_1$ and $|g(x) - f(x)| < \delta_2$ for all $x \in F_2$; that is, $g \in U(f, F_1, \delta_1) \cap U(f, F_2, \delta_2)$. Hence, there exists $U(f, F_3, \delta_3) \in \mathcal{B}_f$ such that $U(f, F_3, \delta_3) \subset U(f, F_1, \delta_1) \cap U(f, F_2, \delta_2)$.

(*V*-c) Pick $U(f, F, \delta) \in \mathcal{B}_f$. We must show that there exists some $U(f, F_0, \delta_0) \in \mathcal{B}_f$ such that if $g \in U(f, F_0, \delta_0)$, then there is some $U(g, F', \delta') \in \mathcal{B}_g$ with $U(g, F', \delta') \subset U(f, F, \delta)$.

Let $F_0 = F$, and $\delta_0 = \delta/2$. Then $U(f, F, \delta/2) \in \mathcal{B}_f$. For every $g \in U(f, F, \delta/2)$, we have

$$|g(x) - f(x)| < \delta/2, \quad \text{for all } x \in F.$$

Let $U(g, F', \delta') = U(g, F, \delta/2)$. If $h \in U(g, F, \delta/2)$, then

$$|h(x) - f(x)| < \delta/2, \quad \text{for all } x \in F.$$

Triangle inequality implies that

$$|h(x) - f(x)| \leqslant |h(x) - g(x)| + |g(x) - f(x)| < \delta/2 + \delta/2 = \delta, \quad \text{for all } x \in F;$$

that is, $h \in U(f, F, \delta)$. Hence, $U(g, F, \delta/2) \subset U(f, F, \delta)$.

Now, $G \subset \mathbb{R}^{\mathbb{I}}$ is open iff G is contains a $U(f, F, \delta)$ of each $f \in G$. This defines a topology on $\mathbb{R}^{\mathbb{I}}$. □

► EXERCISE 43. *For each $f \in \mathbb{R}^{\mathbb{I}}$, the closure of the one-point set $\{f\}$ is just $\{f\}$.*

PROOF. For every $g \in \mathbb{R}^{\mathbb{I}} \smallsetminus \{f\}$, pick $x \in \mathbb{I}$ with $g(x) \neq f(x)$. Define $U(g, F, \delta)$ with $F = \{x\}$ and $\delta < |g(x) - f(x)|$. Then $f \notin U(g, \{x\}, \delta)$; that is, $U(g, \{x\}, \delta) \in \mathbb{R}^{\mathbb{I}} \smallsetminus \{f\}$. Hence, $\mathbb{R}^{\mathbb{I}} \smallsetminus \{f\}$ is open, and so $\{f\}$ is closed. This proves that $\overline{\{f\}} = \{f\}$. □

► EXERCISE 44. *For $f \in \mathbb{R}^{\mathbb{I}}$ and $\varepsilon > 0$, let*

$$V(f, \varepsilon) = \left\{g \in \mathbb{R}^{\mathbb{I}} : |g(x) - f(x)| < \varepsilon, \text{ for each } x \in \mathbb{I}\right\}.$$

Verify that the sets $V(f, \varepsilon)$ form a nhood base at f, making $\mathbb{R}^{\mathbb{I}}$ a topological space.

PROOF. Denote $\mathcal{V}_f = \{V(f,\varepsilon) : \varepsilon > 0\}$. We verify the following properties.

(V-a) If $V(f,\varepsilon) \in \mathcal{V}_f$, then $|f(x) - f(x)| = 0 < \varepsilon$; that is, $f \in V(f,\varepsilon)$.

(V-b) Let $V(f,\varepsilon_1), V(f,\varepsilon_2) \in \mathcal{V}_f$. Let $\varepsilon_3 = \min\{\varepsilon_1,\varepsilon_2\}$. If $g \in V(f,\varepsilon_3)$, then

$$|g(x) - f(x)| < \varepsilon_3 = \min\{\varepsilon_1,\varepsilon_2\}, \quad \text{for all } x \in \mathbb{I}.$$

Hence, $V(f,\varepsilon_3) \subset V(f,\varepsilon_1) \cap V(f,\varepsilon_2)$.

(V-c) For an arbitrary $V(f,\varepsilon) \in \mathcal{V}_f$, pick $V(f,\varepsilon/2) \in \mathcal{V}_f$. For each $g \in V(f,\varepsilon/2)$, pick $V(g,\varepsilon/2) \in \mathcal{V}_g$. If $h \in V(g,\varepsilon/2)$, then $|h(x) - g(x)| < \varepsilon/2$ for all $x \in \mathbb{I}$. Hence

$$|h(x) - f(x)| \leqslant |h(x) - g(x)| + |g(x) - f(x)| < \varepsilon;$$

that is, $V(g,\varepsilon/2) \subset V(f,\varepsilon)$. □

► EXERCISE 45. *Compare the topologies defined in 1 and 3.*

PROOF. It is evident that for every $U(f,F,\delta) \in \mathcal{B}_f$, there exists $V(f,\delta) \in \mathcal{V}_f$ such that $V(f,\delta) \subset U(f,F,\delta)$. Hence, the topology in 1 is weaker than in 3 by Hausdorff criterion. □

2.3 Bases and Subbases

5D. No Axioms for Subbase

► EXERCISE 46. *Any family of subsets of a set X is a subbase for some topology on X and the topology which results is the smallest topology containing the given collection of sets.*

PROOF. Let $\mathcal{S}$ be a family of subsets of X. Let $\tau(\mathcal{S})$ be the intersection of all topologies containing $\mathcal{S}$. Such topologies exist, since 2^X is one such. Also $\tau(\mathcal{S})$ is a topology. It evidently satisfies the requirements "unique" and "smallest."

The topology $\tau(\mathcal{S})$ can be described as follows: It consists of $\varnothing$, X, all finite intersections of the $\mathcal{S}$-sets, and all arbitrary unions of these finite intersections. To verify this, note that since $\mathcal{S} \subset \tau(\mathcal{S})$, then $\tau(\mathcal{S})$ must contain all the sets listed. Conversely, because $\bigcup$ distributes over $\bigcap$, the sets listed actually do from a topology containing $\mathcal{S}$, and which therefore contains $\tau(\mathcal{S})$. □

5E. Bases for the Closed Sets

► EXERCISE 47. *$\mathcal{F}$ is a base for the closed sets in X iff the family of complements of members of $\mathcal{F}$ is a base for the open sets.*

PROOF. Let G be an open set in X. Then $G = X \smallsetminus E$ for some closed subset E. Since $E = \bigcap_{F \in \mathcal{G} \subset \mathcal{F}} F$, we obtain

$$G = X \smallsetminus \left(\bigcap_{F \in \mathcal{G} \subset \mathcal{F}} F \right) = \bigcup_{F \in \mathcal{G} \subset \mathcal{F}} F^c.$$

Thus, $\{F^c : F \in \mathcal{F}\}$ forms a base for the open sets. The converse direction is similar. □

► EXERCISE 48. *$\mathcal{F}$ is a base for the closed sets for some topology on X iff (a) whenever F_1 and F_2 belong to $\mathcal{F}$, $F_1 \cup F_2$ is an intersection of elements of $\mathcal{F}$, and (b) $\bigcap_{F \in \mathcal{F}} F = \varnothing$.*

PROOF. If $\mathcal{F}$ is a base for the closed sets for some topology on X, then (a) and (b) are clear. Suppose, on the other hand, X is a set and $\mathcal{F}$ a collection of subsets of X with (a) and (b). Let $\mathcal{T}$ be all intersections of subcollections from $\mathcal{F}$. Then any intersection of members of $\mathcal{T}$ certainly belongs to $\mathcal{T}$, so $\mathcal{T}$ satisfies (F-a). Moreover, if $\mathcal{F}_1 \subset \mathcal{F}$ and $\mathcal{F}_2 \subset \mathcal{F}$, so that $\bigcap_{E \in \mathcal{F}_1} E$ and $\bigcap_{F \in \mathcal{F}_2} F$ are elements of $\mathcal{T}$, then

$$\left(\bigcap_{E \in \mathcal{F}_1} E \right) \cup \left(\bigcap_{F \in \mathcal{F}_2} F \right) = \bigcap_{E \in \mathcal{F}_1} \bigcap_{F \in \mathcal{F}_2} (E \cup F).$$

But by property (a), the union of two elements of $\mathcal{F}$ is an intersection of elements of $\mathcal{F}$, so $(\bigcap_{E \in \mathcal{F}_1} E) \cup (\bigcap_{F \in \mathcal{F}_2} F)$ is an intersection of elements of $\mathcal{F}$, and hence belongs to $\mathcal{T}$. Thus $\mathcal{T}$ satisfies (F-b). Finally, $\varnothing \in \mathcal{T}$ by (b) and $X \in \mathcal{T}$ since X is the intersection of the empty subcollection from $\mathcal{F}$. Hence $\mathcal{T}$ satisfies (F-c). This completes the proof that $\mathcal{T}$ is the collection of closed sets of X. □

3

NEW SPACES FROM OLD

3.1 Subspaces

3.2 Continuous Functions

7A. Characterization of Spaces Using Functions

▶ Exercise 49. *The characteristic function of A is continuous iff A is both open and closed in X.*

Proof. Let $\mathbb{1}_A \colon X \to \mathbb{R}$ be the characteristic function of A, which is defined by

$$\mathbb{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A. \end{cases}$$

First suppose that $\mathbb{1}_A$ is continuous. Then, say, $\mathbb{1}_A^{-1}\left((1/2, 2)\right) = A$ is open, and $\mathbb{1}_A^{-1}\left((-1, 1/2)\right) = X \smallsetminus A$ is open. Hence, A is both open and closed in X.

Conversely, suppose that A is both open and closed in X. For any open set $U \subset \mathbb{R}$, we have

$$\mathbb{1}_A^{-1}(U) = \begin{cases} A & \text{if } 1 \in U \text{ and } 0 \notin U \\ X \smallsetminus A & \text{if } 1 \notin U \text{ and } 0 \in U \\ \varnothing & \text{if } 1 \notin U \text{ and } 0 \notin U \\ X & \text{if } 1 \in U \text{ and } 0 \in U. \end{cases}$$

Then $\mathbb{1}_A$ is continuous. □

▶ Exercise 50. *X has the discrete topology iff whenever Y is a topological space and $f \colon X \to Y$, then f is continuous.*

Proof. Let Y be a topological space and $f \colon X \to Y$. It is easy to see that f is continuous if X has the discrete topology, so we focus on the sufficiency direction. For any $A \subset X$, let $Y = \mathbb{R}$ and $f = \mathbb{1}_A$. Then by Exercise 49 A is open. □

7C. Functions Agreeing on A Dense Subset

▶ EXERCISE 51. *If f and g are continuous functions from X to $\mathbb{R}$, the set of points x for which $f(x) = g(x)$ is a closed subset of X. Thus two continuous maps on X to $\mathbb{R}$ which agree on a dense subset must agree on all of X.*

PROOF. Denote $A = \{x \in X : f(x) \neq g(x)\}$. Take a point $y \in A$ such that $f(y) > g(y)$ (if it is not true then let $g(y) > f(y)$). Take an $\varepsilon > 0$ such that $f(y) - \varepsilon \geqslant g(y) + \varepsilon$. Since f and g are continuous, there exist nhoods U_1 and U_2 of y such that $f[U_1] \subset (-\varepsilon + f(y), \varepsilon + f(y))$ and $g[U_2] \subset (-\varepsilon + g(y), \varepsilon + g(y))$. Let $U = U_1 \cap U_2$. Then U is a nhood of x and for every $z \in U$ we have

$$f(z) - g(z) > [f(x) - \varepsilon] - [g(x) + \varepsilon] \geqslant 0.$$

Hence, $U \subset A$; that is, U is open, and so $\{x \in X : f(x) = g(x)\} = X \smallsetminus U$ is closed.

Now suppose that $D := \{x \in X : f(x) = g(x)\}$ is dense. Take an arbitrary $x \in X$. Since f and g are continuous, for each $n \in \mathbb{N}$, there exist nhoods V_f and V_g such that $|f(y) - f(x)| < 1/n$ for all $y \in V_f$ and $|g(y) - g(x)| < 1/n$ for all $y \in V_g$. Let $V_n = V_f \cap V_g$. Then there exists $x_n \in V_n \cap D$ with $|f(x_n) - f(x)| < 1/2n$ and $|g(x_n) - g(x)| < 1/2n$. Since $f(x_n) = g(x_n)$, we have

$$\begin{aligned} |f(x) - g(x)| \leqslant |f(x) - f(x_n)| + |f(x_n) - g(x)| &= |f(x) - f(x_n)| + |g(x_n) - g(x)| \\ &< 1/n. \end{aligned}$$

Therefore, $f(x) = g(x)$. □

7E. Range Immaterial

▶ EXERCISE 52. *If $Y \subset Z$ and $f : X \to Y$, then f is continuous as a map from X to Y iff f is continuous as a map from X to Z.*

PROOF. Let $f : X \to Z$ be continuous. Let U be open in Y. Then $U = Y \cap V$ for some V which is open in Z. Therefore,

$$f^{-1}(U) = f^{-1}(Y \cap V) = f^{-1}(Y) \cap f^{-1}(V) = X \cap f^{-1}(V) = f^{-1}(V)$$

is open in X, and so f is continuous as a map from X to Y.

Conversely, let $f : X \to Y$ be continuous and V be open in Z. Then $f^{-1}(V) = f^{-1}(Y \cap V)$. Since $Y \cap V$ is open in Y and f is continuous from X to Y, the set $f^{-1}(Y \cap V)$ is open in X and so f is continuous as a map from X to Z. □

7G. Homeomorphisms within the Line

▶ Exercise 53. *Show that all open intervals in* $\mathbb{R}$ *are homeomorphic.*

Proof. We have

- $(a, b) \sim (0, 1)$ by $f_1(x) = (x-a)/(b-a)$.
- $(a, \infty) \sim (1, \infty)$ by $f_2(x) = x - a + 1$.
- $(1, \infty) \sim (0, 1)$ by $f_3(x) = 1/x$.
- $(-\infty, -a) \sim (a, \infty)$ by $f_4(x) = -x$.
- $(-\infty, \infty) \sim (-\pi/2, \pi/2)$ by $f_5(x) = \arctan x$.

Therefore, by compositing, every open interval is homeomorphic to $(0, 1)$. □

▶ Exercise 54. *All bounded closed intervals in* $\mathbb{R}$ *are homeomorphic.*

Proof. $[a, b] \sim [0, 1]$ by $f(x) = (x-a)/(b-a)$. □

▶ Exercise 55. *The property that every real-valued continuous function on* X *assumes its maximum is a topological property. Thus,* $\mathbb{I} := [0, 1]$ *is not homeomorphic to* $\mathbb{R}$*.*

Proof. Every continuous function assumes its maximum on $[0, 1]$; however, x^2 has no maximum on $\mathbb{R}$. Therefore, $\mathbb{I} \not\sim \mathbb{R}$. □

7K. Semicontinuous Functions

▶ Exercise 56. *If* f_α *is a lower semicontinuous real-valued function on* X *for each* $\alpha \in A$*, and if* $\sup_\alpha f_\alpha(x)$ *exists at each* $x \in X$*, then the function* $f(x) = \sup_\alpha f_\alpha(x)$ *is lower semicontinuous on* X*.*

Proof. For an arbitrary $a \in \mathbb{R}$, we have $f(x) \leqslant a$ iff $f_\alpha(x) \leqslant a$ for all $\alpha \in A$. Hence,

$$\{x \in X : f(x) \leqslant \alpha\} = \bigcap_{\alpha \in A} \{x \in X : f_\alpha(x) \leqslant a\},$$

and so $f^{-1}(-\infty, a]$ is closed; that is, f is lower semicontinuous. □

▶ Exercise 57. *Every continuous function from* X *to* $\mathbb{R}$ *is lower semicontinuous. Thus the supremum of a family of continuous functions, if it exists, is lower semicontinuous. Show by an example that "lower semicontinuous" cannot be replaced by "continuous" in the previous sentence.*

Proof. Suppose that $f : X \to \mathbb{R}$ is continuous. Since $(-\infty, x]$ is closed in $\mathbb{R}$, the set $f^{-1}(-\infty, x]$ is closed in X; that is, f is lower semicontinuous.

To construct an example, let $f : [0, \infty) \to \mathbb{R}$ be defined as follows:

$$f_n(x) = \begin{cases} nx & \text{if } 0 \leqslant x \leqslant 1/n \\ 1 & \text{if } x > 1/n. \end{cases}$$

Then

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

and f is not continuous. □

► Exercise 58. *The characteristic function of a set A in X is lower semicontinuous iff A is open, upper semicontinuous iff A is closed.*

Proof. Observe that

$$\mathbb{1}_A^{-1}(-\infty, a] = \begin{cases} \varnothing & \text{if } a < 0 \\ X \smallsetminus A & \text{if } 0 \leqslant a < 1 \\ X & \text{if } a \geqslant 1. \end{cases}$$

Therefore, $\mathbb{1}_A$ is LSC iff A is open. Similarly for the USC case. □

► Exercise 59. *If X is metrizable and f is a lower semicontinuous function from X to $\mathbb{I}$, then f is the supremum of an increasing sequence of continuous functions on X to $\mathbb{I}$.*

Proof. Let d be the metric on X. First assume f is nonnegative. Define

$$f_n(x) = \inf_{z \in X} \{f(z) + nd(x, z)\}.$$

If $x, y \in X$, then $f(z) + nd(x, z) \leqslant f(z) + nd(y, z) + nd(x, y)$. Take the inf over z (first on the left side, then on the right side) to obtain $f_n(x) \leqslant f_n(y) + nd(x, y)$. By symmetry,

$$|f_n(x) - f_n(y)| \leqslant nd(x, y);$$

hence, f_n is uniformly continuous on X. Furthermore, since $f \geqslant 0$, we have $0 \leqslant f_n(x) \leqslant f(x) + nd(x, x) = f(x)$. By definition, f_n increases with n; we must show that $\lim_n f_n$ is actually f.

Given $\varepsilon > 0$, by definition of $f_n(x)$ there is a point $z_n \in X$ such that

$$f_n(x) + \varepsilon > f(z_n) + nd(x, z_n) \geqslant nd(x, z_n) \tag{3.1}$$

since $f \geqslant 0$. But $f_n(x) + \varepsilon \leqslant f(x) + \varepsilon$; hence $d(x, z_n) \to 0$. Since f is LSC, we have $\liminf_n f(z_n) \geqslant f(x)$ (Ash, 2009, Theorem 8.4.2); hence

$$f(z_n) > f(x) - \varepsilon \quad \text{ev.} \tag{3.2}$$

By (3.1) and (3.2),

$$f_n(x) > f(z_n) - \varepsilon + nd(x, z_n) \geqslant f(z_n) - \varepsilon > f(x) - 2\varepsilon$$

for all sufficiently large n. Thus, $f_n(x) \to f(x)$.

If $|f| \leqslant M < \infty$, then $f + M$ is LSC, finite-valued, and nonnegative. If $0 \leqslant g_n \uparrow (f + M)$, then $f_n = (g_n - M) \uparrow f$ and $|f_n| \geqslant M$. □

7M. $C(X)$ and $C^*(X)$

► EXERCISE 60. *If f and g belong to $C(X)$, then so do $f + g$, $f \cdot g$ and $a \cdot f$, for $a \in \mathbb{R}$. If, in addition, f and g are bounded, then so are $f + g$, $f \cdot g$ and $a \cdot f$.*

PROOF. We first do $f + g$. Since $f, g \in C(X)$, for each $x \in X$ and each $\varepsilon > 0$, there exist nhoods U_1 and U_2 of x such that $f[U_1] \subset (-\varepsilon/2 + f(x), \varepsilon/2 + f(x))$ and $g[U_2] \subset (-\varepsilon/2 + g(x), \varepsilon/2 + g(x))$. Let $U = U_1 \cap U_2$. Then U is a nhood of x, and for every $y \in U$, we have

$$|[f(y) + g(y)] - [f(x) + g(x)]| \leqslant |f(y) - f(x)| + |g(y) - g(x)| < \varepsilon;$$

that is, $f + g$ is continuous.

We then do $a \cdot f$. We suppose that $a > 0$ (all other cases are similar). For each $x \in X$ and $\varepsilon > 0$, there exists a nhood U of x such that $f[U] \subset (-\varepsilon/a + f(x), \varepsilon/a + f(x))$. Then $(a \cdot f)[U] \in (-\varepsilon + a \cdot f(x), \varepsilon + a \cdot f(x))$. So $a \cdot f \in C(X)$.

Finally, to do $f \cdot g$, we first show that $f^2 \in C(X)$ whenever $f \in C(X)$. For each $x \in X$ and $\varepsilon > 0$, there is a nhood U of x such that $f[U] \subset (-\sqrt{\varepsilon} + f(x), \sqrt{\varepsilon} + f(x))$. Then $f^2[U] \subset (-\varepsilon + f^2(x), \varepsilon + f^2(x))$, i.e., $f^2 \in C(X)$. Since

$$f(x) \cdot g(x) = \frac{1}{4}\left[\big(f(x) + g(x)\big)^2 - \big(f(x) - g(x)\big)^2\right],$$

we know that $f \cdot g \in C(X)$ from the previous arguments. □

► EXERCISE 61. *$C(X)$ and $C^*(X)$ are algebras over the real numbers.*

PROOF. It follows from the previous exercise that $C(X)$ is a vector space on $\mathbb{R}$. So everything is easy now. □

► EXERCISE 62. *$C^*(X)$ is a normed linear space with the operations of addition and scalar multiplication given above and the norm $\|f\| = \sup_{x \in X} |f(x)|$.*

PROOF. It is easy to see that $C^*(X)$ is a linear space. So it suffices to show that $\|\cdot\|$ is a norm on $C^*(X)$. We focus on the triangle inequality. Let $f, g \in C^*(X)$. Then for every $x \in X$, we have $|f(x) + g(x)| \leqslant |f(x)| + |g(x)| \leqslant \|f\| + \|g\|$; hence, $\|f + g\| \leqslant \|f\| + \|g\|$. □

3.3 Product Spaces, Weak Topologies

8A. Projection Maps

► Exercise 63. *The βth projection map π_β is continuous and open. The projection $\pi_1 : \mathbb{R}^2 \to \mathbb{R}$ is not closed.*

Proof. Let U_β be open in X_β. Then $\pi_\beta^{-1}(U_\beta)$ is a subbasis open set of the Tychonoff topology on $\bigtimes_\alpha X_\alpha$, and so is open. Hence, π_β is continuous.

Take an arbitrary basis open set U in the Tychonoff topology. Denote $I := \{1, \ldots, n\}$. Then

$$U = \bigtimes_\alpha U_\alpha,$$

where U_α is open in X_α for every $\alpha \in A$, and $U_{\alpha_j} = X_{\alpha_j}$ for all $j \notin I$. Hence,

$$\pi_\beta(U) = \begin{cases} U_\beta & \text{if } \beta = \alpha_i \text{ for some } i \in I \\ X_\beta & \text{otherwise.} \end{cases}$$

That is, $\pi_\beta(U)$ is open in X_β in both case. Since any open set is a union of basis open sets, and since functions preserve unions, the image of any open set under π_β is open.

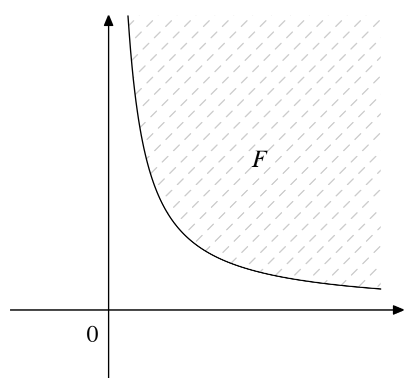


Figure 3.1. $\pi_1(F) = (0, \infty)$

Finally, let $F = \text{epi}(1/x)$. Then F is closed in $\mathbb{R}^2$, but $\pi_1(F) = (0, \infty)$ is open in $\mathbb{R}$; that is, π_1 is not closed. See Figure 3.1. □

► Exercise 64. *Show that the projection of $\mathbb{I} \times \mathbb{R}$ onto $\mathbb{R}$ is a closed map.*

Proof. Let $\pi : \mathbb{I} \times \mathbb{R} \to \mathbb{R}$ be the projection. Suppose $A \subset \mathbb{I} \times \mathbb{R}$ is closed, and suppose $y_0 \in \mathbb{R} \smallsetminus \pi[A]$. For every $x \in \mathbb{I}$, since $(x, y_0) \notin A$ and A is closed, we find a basis open subset $U(x) \times V(x)$ of $\mathbb{I} \times \mathbb{R}$ that contains (x, y_0), and $[U(x) \times V(x)] \cap A = \varnothing$. The collection $\{U(x) : x \in \mathbb{I}\}$ covers $\mathbb{I}$, so finitely many of them cover $\mathbb{I}$ by compactness, say $U(x_1), \ldots, U(x_n)$ do. Now define $V =$

$\bigcap_{i=1}^n V(x_i)$, and note that V is an open nhood of y_0, and $V \cap \pi[A] = \varnothing$. So $\pi[A]$ is closed; that is, π is closed. See Lee (2011, Lemma 4.35, p. 95) for the Tube Lemma.

Generally, if $\pi : X \times Y \to X$ is a projection may where Y is compact, then π is a closed map. □

8B. Separating Points from Closed Sets

► Exercise 65. *If f_α is a map (continuous function) of X to X_α for each $\alpha \in A$, then $\{f_\alpha : \alpha \in A\}$ separates points from closed sets in X iff $\{f_\alpha^{-1}[V] : \alpha \in A, V$ open in $X_\alpha\}$ is a base for the topology on X.*

Proof. Suppose that $\{f_\alpha^{-1}[V] : \alpha \in A, V$ open in $X_\alpha\}$ consists of a base for the topology on X. Let B be closed in X and $x \notin B$. Then $x \in X \setminus B$ and $X \setminus B$ is open in X. Hence there exists $f_\alpha^{-1}[V]$ such that $x \in f_\alpha^{-1}[V] \subset X \setminus B$; that is, $f_\alpha(x) \in V$. Since $V \cap f_\alpha[B] = \varnothing$, i.e., $f_\alpha[B] \subset X_\alpha \setminus V$, and $X_\alpha \setminus V$ is closed, we get $\overline{f_\alpha[B]} \subset X_\alpha \setminus V$. Thus, $f_\alpha(x) \notin \overline{f_\alpha[B]}$.

Next assume that $\{f_\alpha : \alpha \in A\}$ separates points from closed sets in X. Take an arbitrary open subset $U \subset X$ and $x \in U$. Then $B := X \setminus U$ is closed in X, and hence there exists $\alpha \in A$ such that $f_\alpha(x) \notin \overline{f_\alpha[B]}$. Then $f_\alpha(x) \in X_\alpha \setminus \overline{f_\alpha[B]}$ and, since $X_\alpha \setminus \overline{f_\alpha[B]}$ is open in X_α, there exists an open set V of X_α such that $f_\alpha(x) \in V \subset X_\alpha \setminus \overline{f_\alpha[B]}$. Therefore,

$$
\begin{aligned}
x \in f_\alpha^{-1}[V] \subset f_\alpha^{-1}\left[X_\alpha \setminus \overline{f_\alpha[B]}\right] &= X \setminus f_\alpha^{-1}\left[\overline{f_\alpha[B]}\right] \\
&\subset X \setminus f_\alpha^{-1}[f_\alpha[B]] \\
&\subset X \setminus B \\
&= U.
\end{aligned}
$$

Hence, $\{f_\alpha^{-1}[V] : \alpha \in A, V$ open in $X_\alpha\}$ is a base for the topology on X. □

8D. Closure and Interior in Products

Let X and Y be topological spaces containing subsets A and B, respectively. In the product space $X \times Y$:

► Exercise 66. $(A \times B)^\circ = A^\circ \times B^\circ$.

Proof. Since $A^\circ \subset A$ is open in A and $B^\circ \subset B$ is open in B, the set $A^\circ \times B^\circ \subset A \times B$ is open in $A \times B$; hence, $A^\circ \times B^\circ \subset (A \times B)^\circ$.

For the converse inclusion, let $\boldsymbol{x} = (a, b) \in (A \times B)^\circ$. Then there is an basis open set $U_1 \times U_2$ such that $\boldsymbol{x} \in U_1 \times U_2 \subset A \times B$, where U_1 is open in A and U_2 is open in B. Hence, $a \in U_1 \subset A$ and $b \in U_2 \subset B$; that is, $a \in A^\circ$ and $b \in B^\circ$. Then $\boldsymbol{x} \in A^\circ \times B^\circ$. □

► EXERCISE 67. $\overline{A \times B} = \overline{A} \times \overline{B}$.

PROOF. See Exercise 68. □

► EXERCISE 68. *Part 2 can be extended to infinite products, while part 1 can be extended only to finite products.*

PROOF. Assume that $\boldsymbol{y} = (y_\alpha) \in \overline{\bigtimes A_\alpha}$; we show that $y_\alpha \in \overline{A_\alpha}$ for each α; that is, $\boldsymbol{y} \in \bigtimes \overline{A_\alpha}$. Let $y_\alpha \in U_\alpha$, where U_α is open in Y_α; since $\boldsymbol{y} \in \pi_\alpha^{-1}(U_\alpha)$, we must have

$$\varnothing \neq \pi_\alpha^{-1}(U_\alpha) \cap \bigtimes A_\alpha = (U_\alpha \cap A_\alpha) \times \left(\bigtimes_{\beta \neq \alpha} A_\beta \right),$$

and so $U_\alpha \cap A_\alpha \neq \varnothing$. This proves $y_\alpha \in \overline{A_\alpha}$. The converse inclusion is established by reversing these steps: If $\boldsymbol{y} \in \bigtimes \overline{A_\alpha}$, then for any open nhood

$$B := U_{\alpha_1} \times \cdots \times U_{\alpha_n} \times \left(\bigtimes \{ Y_\beta : \beta \neq \alpha_1, \ldots, \alpha_n \} \right),$$

each $U_{\alpha_i} \cap A_{\alpha_i} \neq \varnothing$ so that $B \cap \bigtimes A_\alpha \neq \varnothing$. □

► EXERCISE 69. $\operatorname{Fr}(A \times B) = [\overline{A} \times \operatorname{Fr}(B)] \cup [\operatorname{Fr}(A) \times \overline{B}]$.

PROOF. We have

$$\begin{aligned} \operatorname{Fr}(A \times B) &= \overline{A \times B} \cap \overline{(X \times Y) \smallsetminus (A \times B)} \\ &= (\overline{A} \times \overline{B}) \cap \left[(X \times Y) \smallsetminus (A^\circ \times B^\circ) \right] \\ &= (\overline{A} \times \overline{B}) \cap \left[\left(X \times (Y \smallsetminus B^\circ) \right) \cup \left((X \smallsetminus A^\circ) \times Y \right) \right] \\ &= [\overline{A} \times \operatorname{Fr}(B)] \cup [\operatorname{Fr}(A) \times \overline{B}]. \end{aligned}$$ □

► EXERCISE 70. *If X_α is a nonempty topological space and $A_\alpha \subset X_\alpha$, for each $\alpha \in A$, then $\bigtimes A_\alpha$ is dense in $\bigtimes X_\alpha$ iff A_α is dense in X_α, for each α.*

PROOF. It follows from Exercise 68 that

$$\overline{\bigtimes A_\alpha} = \bigtimes \overline{A_\alpha};$$

that is, $\bigtimes A_\alpha$ is dense in $\bigtimes X_\alpha$ iff A_α is dense in X_α, for each α. □

8E. Miscellaneous Facts about Product Spaces

Let X_α be a nonempty topological space for each $\alpha \in A$, and let $X = \bigtimes X_\alpha$.

► EXERCISE 71. *If V is a nonempty open set in X, then $\pi_\alpha(V) = X_\alpha$ for all but finitely many $\alpha \in A$.*

PROOF. Let $\mathcal{T}_\alpha$ be the topology on X_α for each $\alpha \in A$. Let V be an arbitrary open set in X. Then $V = \bigcup_{k \in K} B_k$, where for each $k \in K$ we have $B_k = \bigtimes_{\alpha \in A} E_{\alpha k}$, and for each $\alpha \in A$ we have $E_{\alpha k} \in \mathcal{T}_\alpha$ while

$$A_k := \{\alpha \in A : E_{\alpha k} \neq X_\alpha\}$$

is finite. Then $\bigcap_{k \in K} A_k$ is finite. If $\alpha_0 \notin \bigcap_{k \in K} A_k$, then there exists $k_0 \in K$ such that $E_{\alpha_0 k_0} = X_{\alpha_0}$. Then

$$\pi_{\alpha_0}^{-1}(B_{k_0}) = \pi_{\alpha_0}^{-1}\left(\bigtimes_{\alpha \in A} E_{\alpha k_0}\right) = X_{\alpha_0},$$

and so $X_{\alpha_0} = \pi_{\alpha_0}^{-1}(B_{k_0}) \subset \pi_{\alpha_0}^{-1}(V)$ implies that $\pi_{\alpha_0}^{-1}(V) = X_{\alpha_0}$. □

► EXERCISE 72. *If b_α is a fixed point in X_α, for each $\alpha \in A$, then $X'_{\alpha_0} = \{x \in X : x_\alpha = b_\alpha$ whenever $\alpha \neq \alpha_0\}$ is homeomorphic to X_{α_0}.*

PROOF. Write an element in X'_{α_0} as $(x_{\alpha_0}, \boldsymbol{b}_{-\alpha_0})$. Then consider the mapping $(x_{\alpha_0}, \boldsymbol{b}_{-\alpha_0}) \mapsto x_{\alpha_0}$. □

8G. The Box Topology

Let X_α be a topological space for each $\alpha \in A$.

► EXERCISE 73. *In $\bigtimes X_\alpha$, the sets of the form $\bigtimes U_\alpha$, where U_α is open in X_α for each $\alpha \in A$, form a base for a topology.*

PROOF. Let $\mathcal{B} := \left\{\bigtimes U_\alpha : \alpha \in A, U_\alpha \text{ open in } X_\alpha\right\}$. Then it is clear that $\bigtimes X_\alpha \in \mathcal{B}$ since X_α is open for each $\alpha \in A$. Now take any $B_1, B_2 \in \mathcal{B}$, with $B_1 = \bigtimes U_\alpha^1$ and $B_2 = \bigtimes U_\alpha^2$. Let

$$p = (p_1, p_2, \dots) \in B_1 \cap B_2 = \bigtimes \left(U_\alpha^1 \cap U_\alpha^2\right).$$

Then $p_\alpha \in U_\alpha^1 \cap U_\alpha^2$, and so there exists an open set $B_\alpha \subset X_\alpha$ such that $p_\alpha \in B_\alpha \subset U_\alpha^1 \cap U_\alpha^2$. Hence, $\bigtimes B_\alpha \in \mathcal{B}$ and $p \in B \subset B_1 \cap B_2$. □

8H. Weak Topologies on Subspaces

Let X have the weak topology induced by a collection of maps $f_\alpha : X \to X_\alpha$, for $\alpha \in A$.

► EXERCISE 74. *If each X_α has the weak topology given by a collection of maps $g_{\alpha\lambda} : X_\alpha \to Y_{\alpha\lambda}$, for $\lambda \in \Lambda_\alpha$, then X has the weak topology given by the maps $g_{\alpha\lambda} \circ f_\alpha : X \to Y_{\alpha\lambda}$ for $\alpha \in A$ and $\lambda \in \Lambda_\alpha$.*

PROOF. A subbase for the weak topology on X_α induced by $\{g_{\alpha\lambda} : \lambda \in \Lambda_\alpha\}$ is

$$\left\{g_{\alpha\lambda}^{-1}(U_{\alpha\lambda}) : \lambda \in \Lambda_\alpha, U_{\alpha\lambda} \text{ open in } Y_{\alpha\lambda}\right\}.$$

Then a subbasic open set in X for the weak topology on X induced by $\{f_\alpha : \alpha \in A\}$ is

$$\left\{ f_\alpha^{-1}[g_{\alpha\lambda}^{-1}(U_{\alpha\lambda})] : \alpha \in A, \lambda \in \Lambda_\alpha, U_{\alpha\lambda} \text{ open in } Y_{\alpha\lambda} \right\}.$$

Since $f_\alpha^{-1}(g_{\alpha\lambda}^{-1}(U_{\alpha\lambda})) = (g_{\alpha\lambda} \circ f_\alpha)^{-1}(U_{\alpha\lambda})$, we get the result. □

► Exercise 75. *Any $B \subset X$ has the weak topology induced by the maps $f_\alpha \restriction B$.*

Proof. As a subspace of X, the subbase on B is

$$\left\{ B \cap f_\alpha^{-1}(U_\alpha) : \alpha \in A, U_\alpha \text{ open in } X_\alpha \right\}.$$

On the other hand, $(f_\alpha \restriction B)^{-1}(U_\alpha) = B \cap f_\alpha^{-1}(U_\alpha)$ for every $\alpha \in A$ and U_α open in X_α. Hence, the above set is also the subbase for the weak topology induced by $\{f_\alpha \restriction B : \alpha \in A\}$. □

3.4 Quotient Spaces

9B. Quotients versus Decompositions

► Exercise 76. *The process given in 9.5 for forming the topology on a decomposition space does define a topology.*

Proof. Let $(X, \mathcal{T})$ be a topological space; let $\mathcal{D}$ be a decomposition of X. Define

$$\mathcal{F} \subset \mathcal{D} \text{ is open in } \mathcal{D} \iff \bigcup \{F : F \in \mathcal{F}\} \text{ is open in } X. \tag{3.3}$$

Let $\mathfrak{T}$ be the collection of open sets defined by (3.3). We show that $(\mathcal{D}, \mathfrak{T})$ is a topological space.

- Take an arbitrary collection $\{\mathcal{F}_i\}_{i \in I} \subset \mathfrak{T}$; then $\bigcup \{F : F \in \mathcal{F}_i\}$ is open in X for each $i \in I$. Hence, $\bigcup_{i \in I} \mathcal{F}_i \in \mathfrak{T}$ since

$$\bigcup_{F \in \bigcup_{i \in I} \mathcal{F}_i} F = \bigcup_{i \in I} \left(\bigcup_{F \in \mathcal{F}_i} F \right)$$

  is open in X.

- Let $\mathcal{F}_1, \mathcal{F}_2 \in \mathfrak{T}$; then $\bigcup_{E \in \mathcal{F}_1} E$ and $\bigcup_{F \in \mathcal{F}_2} F$ are open in X. Therefore, $\mathcal{F}_1 \cap \mathcal{F}_2 \in \mathfrak{T}$ since

$$\bigcup_{F \in \mathcal{F}_1 \cap \mathcal{F}_2} F = \left(\bigcup_{E \in \mathcal{F}_1} E \right) \cap \left(\bigcup_{F \in \mathcal{F}_2} F \right)$$

  is open in X.

- $\varnothing \in \mathfrak{T}$ since $\bigcup \varnothing = \varnothing$ is open in X; finally, $\mathcal{D} \in \mathfrak{T}$ since $\bigcup \mathcal{D} = X$. □

► EXERCISE 77. *The topology on a decomposition space $\mathcal{D}$ of X is the quotient topology induced by the natural map $P: X \to \mathcal{D}$. (See 9.6.)*

PROOF. Let $\mathfrak{T}$ be the decomposition topology of $\mathcal{D}$, and let $\mathfrak{T}_P$ be the quotient topology induced by P. Take an open set $\mathcal{F} \in \mathfrak{T}$; then $\bigcup_{F \in \mathcal{F}} F$ is open in X. Hence,

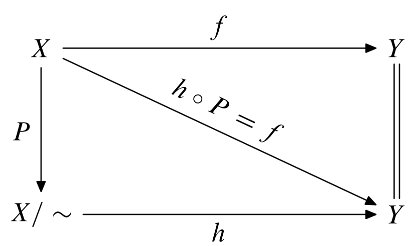
$$P^{-1}(\mathcal{F}) = P^{-1}\left(\bigcup_{F \in \mathcal{F}} F\right) = \bigcup_{F \in \mathcal{F}} P^{-1}(F) = \bigcup_{F \in \mathcal{F}} F$$

is open in X, and so $\mathcal{F} \in \mathfrak{T}_P$. We thus proved that $\mathfrak{T} \subset \mathfrak{T}_P$.

Next take an arbitrary $\mathcal{F} \in \mathfrak{T}_P$. By definition, we have $P^{-1}(\mathcal{F}) = \bigcup_{F \in \mathcal{F}} F$ is open in X. But then $\mathcal{F} \in \mathfrak{T}$.

We finally prove Theorem 9.7 (McCleary, 2006, Theorem 4.18): *Suppose $f: X \to Y$ is a quotient map. Suppose $\sim$ is the equivalence relation defined on X by $x \sim x'$ if $f(x) = f(x')$. Then the quotient space $X/\sim$ is homeomorphic to Y.*

By the definition of the equivalence relation, we have the diagram:



Define $h: X/\sim \to Y$ by letting $h([x]) = f(x)$. It is well-defined. Notice that $h \circ P = f$ since for each $x \in X$ we obtain

$$(h \circ P)(x) = h(P(x)) = h([x]) = f(x).$$

Both f and P are quotient maps so h is continuous by Theorem 9.4. We show that h is injective, subjective and h^{-1} is continuous, which implies that h is a homeomorphism. If $h([x]) = h([x'])$, then $f(x) = f(x')$ and so $x \sim x'$; that is, $[x] = [x']$, and h is injective. If $y \in Y$, then $y = f(x)$ since f is surjective and $h([x]) = f(x) = y$ so h is surjective. To see that h^{-1} is continuous, observe that since f is a quotient map and P is a quotient map, this shows $P = h^{-1} \circ f$ and Theorem 9.4 implies that h^{-1} is continuous. □

4

CONVERGENCE

4.1 Inadequacy of Sequences

10B. Sequential Convergence and Continuity

► Exercise 78. *Find spaces X and Y and a function $F\colon X \to Y$ which is not continuous, but which has the property that $F(x_n) \to F(x)$ in Y whenever $x_n \to x$ in X.*

Proof. Let $X = \mathbb{R}^{\mathbb{R}}$ and $Y = \mathbb{R}$. Define $F\colon \mathbb{R}^{\mathbb{R}} \to \mathbb{R}$ by letting $F(f) = \sup_{x \in \mathbb{R}} |f(x)|$. Then F is not continuous: Let

$$E = \left\{ f \in \mathbb{R}^{\mathbb{R}} : f(x) = 0 \text{ or } 1 \text{ and } f(x) = 0 \text{ only finitely often} \right\},$$

and let $g \in \mathbb{R}^{\mathbb{R}}$ be the function which is 0 everywhere. Then $g \in \overline{E}$. However, $0 \in F[\overline{E}]$ since $F(g) = 0$, and $\overline{F[E]} = \{1\}$. □

10C. Topology of First-Countable Spaces

Let X and Y be first-countable spaces.

► Exercise 79. *$U \subset X$ is open iff whenever $x_n \to x \in U$, then (x_n) is eventually in U.*

Proof. If U is open and $x_n \to x \in U$, then x has a nhood V such that $x \in V \subset U$. By definition of convergence, there is some positive integer n_0 such that $n \geqslant n_0$ implies $x_n \in V \subset U$; hence, (x_n) is eventually in U.

Conversely, suppose that whenever $x_n \to x \in U$, then (x_n) is eventually in U. If U is not open, then there exists $x \in U$ such that for every nhood of V of x we have $V \cap (X \setminus U) \neq \varnothing$. Since X is first-countable, we can pick a countable nhood base $\{V_n : n \in \mathbb{N}\}$ at x. Replacing $V_n = \bigcap_{i=1}^{n} V_i$ where necessary, we may assume that $V_1 \supset V_2 \supset \cdots$. Now $V_n \cap (X \setminus U) \neq \varnothing$ for each n, so we can pick $x_n \in V_n \cap (X \setminus U)$. The result is a sequence (x_n) contained in $X \setminus U$

which converges to $x \in U$; that is, $x_n \to x$ but (x_n) is not eventually in U. A contradiction. □

► EXERCISE 80. *$F \subset X$ is closed iff whenever (x_n) is contained in F and $x_n \to x$, then $x \in F$.*

PROOF. Let F be closed; let (x_n) be contained in F and $x_n \to x$. Then $x \in \bar{F} = F$.

Conversely, assume that whenever (x_n) is contained in F and $x_n \to x$, then $x \in F$. It follows from Theorem 10.4 that $x \in \bar{F}$ with the hypothesis; therefore, $\bar{F} \subset F$, i.e., $\bar{F} = F$ and so F is closed. □

► EXERCISE 81. *$f: X \to Y$ is continuous iff whenever $x_n \to x$ in X, then $f(x_n) \to f(x)$ in Y.*

PROOF. Suppose f is continuous and $x_n \to x$. Since f is continuous at x, for every nhood V of $f(x)$ in Y, there exists a nhood U of x in X such that $f(U) \subset V$. Since $x_n \to x$, there exists n_0 such that $n \geqslant n_0$ implies that $x_n \in U$. Hence, for every nhood V of $f(x)$, there exists n_0 such that $n \geqslant n_0$ implies that $f(x_n) \in V$; that is, $f(x_n) \to f(x)$.

Conversely, let the criterion hold. Suppose that f is not continuous. Then there exists $x \in X$ and a nhood V of $f(x)$, such that for every nhood base U_n, $n \in \mathbb{N}$, of x, there is $x_n \in U_n$ with $f(x_n) \notin V$. By letting $U_1 \supset U_2 \supset \cdots$, we have $x_n \to x$ and so $f(x_n) \to f(x)$; that is, eventually, $f(x_n)$ is in V. A contradiction. □

4.2 NETS

11A. Examples of Net Converence

► EXERCISE 82. *In $\mathbb{R}^{\mathbb{R}}$, let*

$$E = \left\{ f \in \mathbb{R}^{\mathbb{R}} : f(x) = 0 \text{ or } 1, \text{ and } f(x) = 0 \text{ only finitely often} \right\},$$

and g be the function in $\mathbb{R}^{\mathbb{R}}$ which is identically 0. Then, in the product topology on $\mathbb{R}^{\mathbb{R}}$, $g \in \bar{E}$. Find a net (f_λ) in E which converges to g.

PROOF. Let $\mathcal{U}_g = \{U(g, F, \varepsilon) : \varepsilon > 0, F \subset \mathbb{R} \text{ a finite set}\}$ be the nhood base of g. Order $\mathcal{U}_g$ as follows:

$$\begin{aligned} U(g, F_1, \varepsilon_1) \leqslant U(g, F_2, \varepsilon_2) &\iff U(g, F_2, \varepsilon_2) \subset U(g, F, \varepsilon_2) \\ &\iff F_1 \subset F_2 \text{ and } \varepsilon_2 \leqslant \varepsilon_1. \end{aligned}$$

Then $\mathcal{U}_g$ is a directed set. So we have a net $(f_{F,\varepsilon})$ converging to g. □

11B. Subnets and Cluster Points

▶ EXERCISE 83. *Every subnet of an ultranet is an ultranet.*

PROOF. Take an arbitrary subset $E \subset X$. Let (x_λ) be an ultranet in X, and suppose that (x_λ) is residually in E, i.e., there exists some $\lambda_0 \in \Lambda$ such that $\lambda \geqslant \lambda_0$ implies that $x_\lambda \in E$. If (x_{λ_μ}) is a subnet of (x_λ), then there exists some μ_0 such that $\lambda_{\mu_0} \geqslant \lambda_0$. Then for every $\mu \geqslant \mu_0$, we have $\lambda_\mu \geqslant \lambda_0$, and so $\mu \geqslant \mu_0$ implies that $x_{\lambda_\mu} \in E$; that is, (x_{λ_μ}) is residually in E. □

▶ EXERCISE 84. *Every net has a subnet which is an ultranet.*

PROOF. See Adamson (1996, Exercise 127, p. 40). □

▶ EXERCISE 85. *If an ultranet has x as a cluster point, then it converges to x.*

PROOF. Let (x_λ) be an ultranet, and x be a cluster point of (x_λ). Let U be a nhood of x. Then (x_λ) lies in U eventually since for any λ_0 there exists $\lambda \geqslant \lambda_0$ such that $x_\lambda \in U$. □

11D. Nets Describe Topologies

▶ EXERCISE 86. *Nets have the following four properties:*

a. *if $x_\lambda = x$ for each $\lambda \in \Lambda$, then $x_\lambda \to x$,*

b. *if $x_\lambda \to x$, then every subnet of (x_λ) converges to x,*

c. *if every subnet of (x_λ) has a subnet converging to x, then (x_λ) converges to x,*

d. *(Diagonal principal) if $x_\lambda \to x$ and, for each $\lambda \in \Lambda$, a net $(x^\lambda_\mu)_{u \in M_\lambda}$ converges to x_λ, then there is a diagonal net converging to x; i.e., the net $(x^\lambda_\mu)_{\lambda \in \Lambda, \mu \in M_\lambda}$, ordered lexicographically by Λ, then by M_λ, has a subnet which converges to x.*

PROOF. **(a)** If the net (x_λ) is trivial, then for each nhood U of x, we have $x_\lambda \in U$ for all $\lambda \in \Lambda$. Hence, $x_\lambda \to x$.

(b) Let $(x_{\varphi(\mu)})_{\mu \in M}$ be a subnet of (x_λ). Take any nhood U of x. Then there exists $\lambda_0 \in \Lambda$ such that $\lambda \geqslant \lambda_0$ implies that $x_\lambda \in U$ since $x_\lambda \to x$. Since φ is cofinal in Λ, there exists $\mu_0 \in M$ such that $\varphi(\mu_0) \geqslant \lambda_0$; since φ is increasing, $\mu \geqslant \mu_0$ implies that $\varphi(\mu) \geqslant \varphi(\mu_0) \geqslant \lambda_0$. Hence, there exists $\mu_0 \in M$ such that $\mu \geqslant \mu_0$ implies that $x_{\varphi(\mu)} \in U$; that is, $x_{\varphi(\mu)} \to x$.

(c) Suppose by way of contradiction that (x_λ) does not converge to x. Then there exists a nhood U of x such that for any $\lambda \in \Lambda$, there exists some $\varphi(\lambda) \geqslant \lambda$ with $x_{\varphi(\lambda)} \notin U$. Then $(x_{\varphi(\lambda)})$ is a subnet of (x_λ), but which has no converging subnets.

(d) Order $\{(\lambda,\mu) : \lambda \in \Lambda, \mu \in M_\lambda\}$ as follows:

$$(\lambda_1,\mu_1) \leqslant (\lambda_2,\mu_2) \iff \lambda_1 \leqslant \lambda_2, \text{ or } \lambda_1 = \lambda_2 \text{ and } \mu_1 \leqslant \mu_2.$$

Let $\mathcal{U}$ be the nhood system of x which is ordered by $U_1 \leqslant U_2$ iff $U_2 \subset U_1$ for all $U_1, U_2 \in \mathcal{U}$. Define

$$\Gamma = \left\{(\lambda, U) : \lambda \in \Lambda, U \in \mathcal{U} \text{ such that } x^\lambda \in U\right\}.$$

Order Γ as follows: $(\lambda_1, U_1) \leqslant (\lambda_2, U_2)$ iff $\lambda_1 \leqslant \lambda_2$ and $U_2 \subset U_1$. For each $(\lambda, U) \in \Gamma$ pick $\mu_\lambda \in M_\lambda$ so that $x^\lambda_\mu \in U$ for all $\mu \geqslant \mu_\lambda$ (such a μ_λ exists since $x^\lambda_\mu \to x^\lambda$ and $x^\lambda \in U$). Define $\varphi\colon (\lambda, U) \mapsto x^\lambda_{\mu_\lambda}$ for all $(\lambda, U) \in \Gamma$. It now easy to see that this subnet converges to x. □

4.3 Filters

12A. Examples of Filter Convergence

► Exercise 87. *Show that if a filter in a metric space converges, it must converge to a unique point.*

Proof. Suppose a filter $\mathcal{F}$ in a metric space (X, d) converges to $x, y \in X$. If $x \neq y$, then there exists $r > 0$ such that $\mathbb{B}(x,r) \cap \mathbb{B}(y,r) = \varnothing$. But since $\mathcal{F} \to x$ and $\mathcal{F} \to y$, we must have $\mathbb{B}(x,r) \in \mathcal{F}$ and $\mathbb{B}(y,r) \in \mathcal{F}$. This contradicts the fact that the intersection of every two elements in a filter is nonempty. Thus, $x = y$. □

12C. Ultrafilters: Uniqueness

► Exercise 88. *If a filter $\mathcal{F}$ is contained in a unique ultrafilter $\mathcal{F}'$, then $\mathcal{F} = \mathcal{F}'$.*

Proof. We first show: *Every filter $\mathcal{F}$ on a non-empty set X is the intersection of the family of ultrafilters which include $\mathcal{F}$.*

Let E be a set which does not belong to $\mathcal{F}$. Then for each set $F \in \mathcal{F}$ we cannot have $F \subset E$ and hence we must have $F \cap E^c \neq \varnothing$. So $\mathcal{F} \cup \{E^c\}$ generates a filter on X, which is included in some ultrafilter $\mathcal{F}_E$. Since $E^c \in \mathcal{F}_E$ we must have $E \notin \mathcal{F}_E$. Thus E does not belong to the intersection of the set of all ultrafilters which include $\mathcal{F}$. Hence this intersection is just the filter $\mathcal{F}$ itself.

Now, if $\mathcal{F}$ is contained in a unique ultrafilter $\mathcal{F}'$, we must have $\mathcal{F} = \mathcal{F}'$. □

12D. Nets and Filters: The Translation Process

► EXERCISE 89. *A net (x_λ) has x as a cluster point iff the filter generated by (x_λ) has x as a cluster point.*

PROOF. Suppose x is a cluster point of the net (x_λ). Then for every nhood U of x, we have $x_\lambda \in U$ i. o. But then U meets every $B_{\lambda_0} := \{x_\lambda : \lambda \geqslant \lambda_0\}$, the filter base of the filter $\mathscr{F}$ generated by (x_λ); that is, x is a cluster point of $\mathscr{F}$. The converse implication is obvious. □

► EXERCISE 90. *A filter $\mathscr{F}$ has x as a cluster point iff the net based on $\mathscr{F}$ has x as a cluster point.*

PROOF. Suppose x is a cluster point of $\mathscr{F}$. If U is a nhood of x, then U meets every $F \in \mathscr{F}$. Then for an arbitrary $(p, F) \in \Lambda_{\mathscr{F}}$, pick $q \in F \cap U$ so that $(q, F) \in \Lambda_{\mathscr{F}}$, $(q, F) \geqslant (p, F)$, and $P(p, F) = p \in U$; that is, x is a cluster point of the net based on $\mathscr{F}$.

Conversely, suppose the net based on $\mathscr{F}$ has x as a cluster point. Let U be a nhood of x. Then for every $(p_0, F_0) \in \Lambda_{\mathscr{F}}$, there exists $(p, F) \geqslant (p_0, F_0)$ such that $p \in U$. Then $F_0 \cap U \neq \varnothing$, and so x is a cluster point of $\mathscr{F}$. □

► EXERCISE 91. *If (x_{λ_μ}) is a subnet of (x_λ), then the filter generated by (x_{λ_μ}) is finer than the filter generated by (x_λ).*

PROOF. Suppose (x_{λ_μ}) is a subnet of (x_λ). Let $\mathscr{F}_{\lambda_\mu}$ is the filter generated by (x_{λ_μ}), and $\mathscr{F}_\lambda$ be the filter generated by (x_λ). Then the base generating $\mathscr{F}_{\lambda_\mu}$ is the sets $B_{\lambda_{\mu_0}} = \{x_{\lambda_\mu} : \mu \geqslant \mu_0\}$, and the base generating $\mathscr{F}_\lambda$ is the sets $B_{\lambda_0} = \{x_\lambda : \lambda \geqslant \lambda_0\}$. For each such a B_{λ_0}, there exists μ_0 such that $\lambda_{\mu_0} \geqslant \lambda_0$; that is, $B_{\lambda_{\mu_0}} \subset B_{\lambda_0}$. Therefore, $\mathscr{F}_\lambda \subset \mathscr{F}_{\lambda_\mu}$. □

► EXERCISE 92. *The net based on an ultrafilter is an ultranet and the filter generated by an ultranet is an ultrafilter.*

PROOF. Suppose $\mathscr{F}$ is an ultrafilter. Let $E \subset X$ and we assume that $E \in \mathscr{F}$. Pick $p \in E$. If $(q, F) \geqslant (p, E)$, then $q \in E$; that is, $P(p, F) \in E$ ev. Hence, the net based on $\mathscr{F}$ is an ultranet.

Conversely, suppose (x_λ) is an ultranet. Let $E \subset X$ and we assume that there exists λ_0 such that $x_\lambda \in E$ for all $\lambda \geqslant \lambda_0$. Then $B_{\lambda_0} = \{x_\lambda : \lambda \geqslant \lambda_0\} \subset E$ and so $E \in \mathscr{F}$, where $\mathscr{F}$ is the filter generated by (x_λ). Hence, $\mathscr{F}$ is an ultrafilter. □

► EXERCISE 93. *The net based on a free ultrafilter is a nontrivial ultranet. Hence, assuming the axiom of choice, there are nontrivial ultranets.*

PROOF. Let $\mathscr{F}$ be a free ultrafilter, and (x_λ) be the net based on $\mathscr{F}$. It follows from the previous exercise that (x_λ) is an ultranet. If (x_λ) is trivial, i.e., $x_\lambda = x$ for some $x \in X$ and all $\lambda \in \Lambda_{\mathscr{F}}$, then for all $F \in \mathscr{F}$, we must have $F = \{x\}$. But then $\bigcap \mathscr{F} = \{x\} \neq \varnothing$; that is, $\mathscr{F}$ is fixed. A contradiction.

Now, for instance, the Frechet filter $\mathscr{F}$ on $\mathbb{R}$ is contained in some free ultrafilter $\mathscr{G}$ by Example (b) when the Axiom of Choice is assumed. Hence, the net based on $\mathscr{G}$ is a nontrivial ultranet. □

5

SEPARATION AND COUNTABILITY

5.1 The Separation Axioms

13B. T_0- and T_1-Spaces

▶ Exercise 94. *Any subspace of a T_0- or T_1-space is, respectively, T_0 or T_1.*

Proof. Let X be a T_0-space, and $A \subset X$. Let x and y be distinct points in A. Then, say, there exists an open nhood U of x such that $y \notin U$. Then $U \cap A$ is relatively open in A, contains x, and $y \notin A \cap U$. The T_1 case can be proved similarly. □

▶ Exercise 95. *Any nonempty product space is T_0 or T_1 iff each factor space is, respectively, T_0 or T_1.*

Proof. If X_α is a T_0-space, for each $\alpha \in A$, and $x \neq y$ in $\bigtimes X_\alpha$, then for some coordinate α we have $x_\alpha \neq y_\alpha$, so there exists an open set U_α containing, say, x_α but not y_α. Now $\pi_\alpha^{-1}(U_\alpha)$ is an open set in $\bigtimes X_\alpha$ containing x but not y. Thus, $\bigtimes X_\alpha$ is T_0.

Conversely, if $\bigtimes X_\alpha$ is a nonempty T_0-space, pick a fixed point $b_\alpha \in X_\alpha$, for each $\alpha \in A$. Then the subspace $B_\alpha := \{x \in \bigtimes X_\alpha : x_\beta = b_\beta \text{ unless } \beta = \alpha\}$ is T_0, by Exercise 94, and is homeomorphic to X_α under the restriction to B_α of the projection map. Thus X_α is T_0, for each $\alpha \in A$. The T_1 case is similar. □

13C. The T_0-Identification

For any topological space X, define $\sim$ by $x \sim y$ iff $\overline{\{x\}} = \overline{\{y\}}$.

▶ Exercise 96. *$\sim$ is an equivalence relation on X.*

Proof. Straightforward. □

▶ Exercise 97. *The resulting quotient space $X/\sim = \widetilde{X}$ is T_0.*

PROOF. We first show that X is T_0 iff whenever $x \neq y$ then $\overline{\{x\}} \neq \overline{\{y\}}$. If X is T_0 and $x \neq y$, then there exists an open nhood U of x such that $y \notin U$; then $y \notin \overline{\{x\}}$. Since $y \in \overline{\{y\}}$, we have $\overline{\{x\}} \neq \overline{\{y\}}$. Conversely, suppose that $x \neq y$ implies that $\overline{\{x\}} \neq \overline{\{y\}}$. Take any $x \neq y$ in X and we show that there exists an open nhood of one of the two points such that the other point is not in U. If not, then $y \in \overline{\{x\}}$; since $\overline{\{x\}}$ is closed, we have $\overline{\{y\}} \subset \overline{\{x\}}$; similarly, $\overline{\{x\}} \subset \overline{\{y\}}$. A contradiction.

Now take any $\overline{\{x\}} \neq \overline{\{y\}}$ in $X/\sim$. Then $\overline{\{x\}} = \overline{\overline{\{x\}}} \neq \overline{\overline{\{y\}}} = \overline{\{y\}}$. Hence, $X/\sim$ is T_0. □

13D. The Zariski Topology

For a polynomial P in n real variables, let $Z(P) = \{(x_1, \ldots, x_n) \in \mathbb{R}^n : P(x_1, \ldots, x_n) = 0\}$. Let $\mathcal{P}$ be the collection of all such polynomials.

► EXERCISE 98. *$\{Z(P) : P \in \mathcal{P}\}$ is a base for the closed sets of a topology (the* Zariski topology*) on $\mathbb{R}^n$.*

PROOF. Denote $\mathcal{Z} := \{Z(P) : P \in \mathcal{P}\}$. If $Z(P_1)$ and $Z(P_2)$ belong to $\mathcal{Z}$, then $Z(P_1) \cup Z(P_2) = Z(P_1 \cdot P_2) \in \mathcal{Z}$ since $P_1 \cdot P_2 \in \mathcal{P}$. Further, $\bigcap_{P \in \mathcal{P}} Z(P) = \varnothing$ since there are $P \in \mathcal{P}$ with $Z(P) = \varnothing$ (for instance, $P = 1 + X_1^2 + \cdots + X_n^2$). It follows from Exercise 48 that $\mathcal{Z}$ is a base for the closed sets of the Zariski topology on $\mathbb{R}^n$. □

► EXERCISE 99. *The Zariski topology on $\mathbb{R}^n$ is T_1 but not T_2.*

PROOF. To verify that the Zariski topology is T_1, we show that every singleton set in $\mathbb{R}^n$ is closed (by Theorem 13.4). For each $(x_1, \ldots, x_n) \in \mathbb{R}^n$, define a polynomial $P \in \mathcal{P}$ as follows:

$$P = (X_1 - x_1)^2 + \cdots (X_n - x_n)^2.$$

Then $Z(P) = \{(x_1, \ldots, x_n)\}$; that is, $\{(x_1, \ldots, x_n)\}$ is closed.

To see the Zariski topology is not T_2, consider the $\mathbb{R}$ case. In $\mathbb{R}$, the Zariski topology coincides with the cofinite topology (see Exercise 100). It is well know that the cofinite topology is not Hausdorff (Example 13.5(a)). □

► EXERCISE 100. *On $\mathbb{R}$, the Zariski topology coincides with the cofinite topology; in $\mathbb{R}^n$, $n > 1$, they are different.*

PROOF. On $\mathbb{R}$, every $Z(P)$ is finite. So on $\mathbb{R}$ every closed set in the Zariski topology is finite since every closed set is an intersection of some subfamily of $\mathcal{Z}$. However, if $n > 1$, then $Z(P)$ can be infinite: for example, consider the polynomial $X_1 X_2$ (let $X_1 = 0$, then all $X_2 \in \mathbb{R}$ is a solution). □

13H. Open Images of Hausdorff Spaces

► EXERCISE 101. *Given any set X, there is a Hausdorff space Y which is the union of a collection $\{Y_x : x \in X\}$ of disjoint subsets, each dense in Y.*

PROOF. □

5.2 Regularity and Complete Regularity

THEOREM 5.1 (Dugundji 1966). a. *Let $P : X \to Y$ be a closed map. Given any subset $S \subset Y$ and any open U containing $P^{-1}(S)$, there exists an open $V \supset S$ such that $P^{-1}(V) \subset U$.*

b. *Let $P : X \to Y$ be an open map. Given any subset $S \subset Y$, and any closed A containing $P^{-1}S$, there exists a closed $B \supset S$ such that $P^{-1}(B) \subset A$.*

PROOF. It is enough to prove (a). Let $V = Y \smallsetminus P(X \smallsetminus U)$. Then

$$\begin{aligned} P^{-1}(S) \subset U &\implies X \smallsetminus U \subset X \smallsetminus P^{-1}(S) = P^{-1}(Y \smallsetminus S) \\ &\implies P(X \smallsetminus U) \subset P[P^{-1}(Y \smallsetminus S)] \\ &\implies Y \smallsetminus P[P^{-1}(Y \smallsetminus S)] \subset V. \end{aligned}$$

Since $P[P^{-1}(Y \smallsetminus S)] \subset Y \smallsetminus S$, we obtain

$$S = Y \smallsetminus (Y \smallsetminus S) \subset Y \smallsetminus P[P^{-1}(Y \smallsetminus S)] \subset V;$$

that is, $S \subset V$. Because P is closed, V is open in Y. Observing that

$$P^{-1}(V) = X \smallsetminus P^{-1}[P(X \smallsetminus U)] \subset X \smallsetminus (X \smallsetminus U) = U$$

completes the proof. □

THEOREM 5.2 (Theorem 14.6). *If X is T_3 and f is a continuous, open and closed map of X onto Y, then Y is T_2.*

PROOF. By Theorem 13.11, it is sufficient to show that the set

$$A := \{(x_1, x_2) \in X \times X : f(x_1) = f(x_2)\}$$

is closed in $X \times X$. If $(x_1, x_2) \notin A$, then $x_1 \notin f^{-1}[f(x_2)]$. Since a T_3-space is T_1, the singleton set $\{x_2\}$ is closed in X; since f is closed, $\{f(x_2)\}$ is closed in Y; since f is continuous, $f^{-1}[f(x_2)]$ is closed in X. Because X is T_3, there are disjoint open sets U and V with

$$x_1 \in U, \quad \text{and} \quad f^{-1}[f(x_2)] \subset V.$$

Since f is closed, it follows from Theorem 5.1 that there exists open set $W \subset Y$ such that $\{f(x_2)\} \subset W$, and $f^{-1}(W) \subset V$; that is,

$$f^{-1}[f(x_2)] \subset f^{-1}(W) \subset V.$$

Then $U \times f^{-1}(W)$ is a nhood of (x_1, x_2). We finally show that $[U \times f^{-1}(W)] \cap A = \varnothing$. If there exists $(y_1, y_2) \in A$ such that $(y_1, y_2) \in U \times f^{-1}(W)$, then $y_1 \in f^{-1}[f(y_2)] \subset f^{-1}(W)$; that is, $y_1 \in U \times f^{-1}(W)$. However, $U \cap V = \varnothing$ and $f^{-1}(W) \subset V$ imply that $U \cap f^{-1}(W) = \varnothing$. A contradiction. □

DEFINITION 5.3. If X is a space and $A \subset X$, then X/A denotes the quotient space obtained via the equivalence relation whose equivalence classes are A and the single point sets $\{x\}$, $x \in X \setminus A$.

THEOREM 5.4. *If X is T_3 and Y is obtained from X by identifying a single closed set A in X with a point, then Y is T_2.*

PROOF. Let A be a closed subset of a T_3-space X. Then $X \setminus A$ is an open subset in both X and X/A and its two subspace topologies agree. Thus, points in $X \setminus A \subset X/A$ are different from $[A]$ and have disjoint nhoods as X is Hausdorff. Finally, for $x \in X \setminus A$, there exist disjoint open nhoods $V(x)$ and $W(A)$. Their images, $f(V)$ and $f(W)$, are disjoint open nhoods of x and $[A]$ in X/A, because $V = f^{-1}[f(V)]$ and $W = f^{-1}[f(W)]$ are disjoint open sets in X. □

5.3 NORMAL SPACES

15B. Completely Normal Spaces

► EXERCISE 102. *X is completely normal iff whenever A and B are subsets of X with $A \cap \overline{B} = \overline{A} \cap B = \varnothing$, then there are disjoint open sets $U \supset A$ and $V \supset B$.*

PROOF. Suppose that whenever A and B are subsets of X with $A \cap \overline{B} = \overline{A} \cap B = \varnothing$, then there are disjoint open sets $U \supset A$ and $V \supset B$. Let $Y \subset X$, and $C, D \subset Y$ be disjoint closed subsets of Y. Hence,

$$\varnothing = \mathrm{cl}_Y(C) \cap \mathrm{cl}_Y(D) = [\overline{C} \cap Y] \cap [\overline{D} \cap Y] = \overline{C} \cap [\overline{D} \cap Y].$$

Since $D \subset \mathrm{cl}_Y(D)$, we have $\overline{C} \cap D = \varnothing$. Similarly, $C \cap \overline{D} = \varnothing$. Hence there are disjoint open sets U' and V' in X such that $C \subset U'$ and $D \subset V'$. Let $U = U' \cap Y$ and $V = V' \cap Y$. Then U and V are open in Y, $C \subset U$, and $D \subset V$; that is, Y is normal, and so X is completely normal.

Now suppose that X is completely normal and consider the subspace $Y := X \setminus (\overline{A} \cap \overline{B})$. We first show that $A, B \subset Y$. If $A \not\subset Y$, then there exists $x \in A$ with $x \notin Y$; that is, $x \in \overline{A} \cap \overline{B}$. But then $x \in A \cap \overline{B}$. A contradiction. Similarly for B. In the normal space Y, we have

$$\text{cl}_Y(A) \cap \text{cl}_Y(B) = [\bar{A} \cap Y] \cap [\bar{B} \cap Y] = (\bar{A} \cap \bar{B}) \cap [X \smallsetminus (\bar{A} \cap \bar{B})] = \varnothing.$$

Therefore, there exist disjoint open sets $U \supset \text{cl}_Y(A)$ and $V \supset \text{cl}_Y(B)$. Since $A \subset \text{cl}_Y(A)$ and $B \subset \text{cl}_Y(B)$, we get the desired result. □

▶ EXERCISE 103. *Why can't the method used to show every subspace of a regular space is regular be carried over to give a proof that every subspace of a normal space is normal?*

PROOF. In the first proof, if $A \subset Y \subset X$ is closed in Y and $x \in Y \smallsetminus A$, then there must exists closed set B in X such that $x \notin B$. This property is not applied if $\{x\}$ is replaced a general closed set B in Y. □

▶ EXERCISE 104. *Every metric space is completely normal.*

PROOF. Every subspace of a metric space is a metric space; every metric space is normal Royden and Fitzpatrick (2010, Proposition 11.7). □

5.4 COUNTABILITY PROPERTIES

16A. First Countable Spaces

▶ EXERCISE 105. *Every subspace of a first-countable space is first countable.*

PROOF. Let $A \subset X$. If $x \in A$, then V is a nhood of x in A iff $V = U \cap A$, where U is a nhood of $x \in X$ (Theorem 6.3(d)). □

▶ EXERCISE 106. *A product $\bigtimes X_\alpha$ of first-countable spaces is first countable iff each X_α is first countable, and all but countably many of the X_α are trivial spaces.*

PROOF. If $\bigtimes X_\alpha$ is first-countable, then each X_α is first countable since it is homeomorphic to a subspace of $\bigtimes X_\alpha$. If the number of the family of untrivial sets $\{X_\alpha\}$ is uncountable, then for $x \in \bigtimes X_\alpha$ the number of nhood bases is uncountable. □

▶ EXERCISE 107. *The continuous image of a first-countable space need not be first countable; but the continuous open image of a first-countable space is first countable.*

PROOF. Let X be a discrete topological space. Then any function defined on X is continuous.

Now suppose that X is first countable, and f is a continuous open map of X onto Y. Pick an arbitrary $y \in Y$. Let $x \in f^{-1}(y)$, and $\mathcal{U}_x$ be a countable nhood base of x. If W is a nhood of y, then there is a nhood V of x such that

$f(V) \subset W$ since f is continuous. So there exists $U \in \mathcal{U}_x$ with $f(U) \subset W$. This proves that $\{f(U) : U \in \mathcal{U}_x\}$ is a nhood base of y. Since $\{f(U) : U \in \mathcal{U}_x\}$ is □