Unauthorized Solution Manual: Topology, 2nd Edition, James Munkres.

Shane:
https://github.com/Shena4746/Exercise-Munkres

June 18, 2022

Contents

About This Document

This is a set of solutions to the exercises in Chapter 1, Sections 1 through 11 of the book "Topology, 2nd Edition" by James Munkres.

All statements of the exercises in this doc are borrowed from the book, and the solutions are taken from my old TeX note written around 2010 when I was an undergraduate student. At the time of writing, unfortunately, these solutions have not undergone reliable review. Therefore, they should be examined by the readers themselves. Any suggestions for improvements would be appreciated.

Shena.
June 16, 2022,

Notations

We summarize some of the frequently used notation in this document. Those already introduced in Mukres [1] are not included here.

Let P and Q be formulas, let A and B be sets, and let f be a function.

- $\neg$: negation.
- $\vee$: disjunction.
- $\wedge$: conjunction.
- $\Rightarrow$: implication.
- $\Leftrightarrow$: equivalence. Formally, $P \Leftrightarrow Q$ is defined as $(P \Rightarrow Q) \wedge (Q \Rightarrow P)$.
- $\equiv$: the same meaning as "$\Leftrightarrow$."
- ${}^{\forall}$: universal quantifier.
- ${}^{\exists}$: existential quantifier.
- ${}^{\exists 1}$: "${}^{\exists 1}x\,[P(x)]$" means "$P(x)$ holds for exactly one x." Formally, ${}^{\exists 1}x\,[P(x)]$ is defined as ${}^{\exists}x\,\big[P(x) \wedge {}^{\forall}y\,[P(y) \Rightarrow x = y]\big]$.
- T : short for "true."
- F : short for "false."
- $:=$: "is defined to be." $a := b$ means that a is defined to be b, while $a =: b$ means that b is defined to be a.

- $\setminus$: the difference of two sets.
- $\subsetneq$: proper inclusion.
- $\mapsto$: Assignment. We often write "$f : A \ni a \mapsto b \in B$" to mean f is a function from A to B assigning, to each element a of A, an element b of B.
- $\mathrm{Func}(A, B)$: the set of all functions from A to B.
- $\mathrm{Inj}(A, B)$: the set of all injective functions from A to B.
- $\mathrm{Surj}(A, B)$: the set of all surjective functions from A to B.
- $\mathrm{Bij}(A, B)$: the set of all bijective functions from A to B.
- $\sim$: equivalence relation.
- $\simeq$: equivalence relation (often about order type in this document).
- $\hookrightarrow$: existence of injection. We write $A \hookrightarrow B$ if there exists an injection from A into B.
- $(A, <)$: the set A equipped with the order relation $<$.
- $\min A$: a smallest element of A.
- $\max A$: a largest element of A.
- $\inf A$: an infimum of A.
- $\sup A$: a supremum of A.

1 Fundamental Concepts

Exercise 1. Check the distribution laws for $\cup$ and $\cap$ and DeMorgan's laws.

Solution. These laws are simple consequence of the corresponding laws for $\wedge, \vee$, as follows:

$$\begin{aligned} A \cap (B \cup C) &= \{x \mid x \in A \wedge (x \in B \vee x \in C)\} \\ &= \{x \mid (x \in A \wedge x \in B) \vee (x \in A \wedge x \in C)\} \\ &= (A \cap B) \cup (A \cap C), \end{aligned}$$

$$\begin{aligned} A \cup (B \cap C) &= \{x \mid x \in A \vee (x \in B \wedge x \in C)\} \\ &= \{x \mid (x \in A \vee x \in B) \wedge (x \in A \vee x \in C)\} \\ &= (A \cup B) \cap (A \cup C), \end{aligned}$$

$$\begin{aligned} A \setminus (B \cup C) &= \{x \mid x \in A \wedge \neg(x \in B \vee x \in C)\} \\ &= \{x \mid x \in A \wedge (x \notin B \wedge x \notin C)\} \\ &= \{x \mid (x \in A \wedge x \notin B) \wedge (x \in A \wedge x \notin C)\} \\ &= (A \setminus B) \cap (A \setminus C), \end{aligned}$$

$$\begin{aligned} A \setminus (B \cap C) &= \{x \mid x \in A \wedge \neg(x \in B \wedge x \in C)\} \\ &= \{x \mid x \in A \wedge (x \notin B \vee x \notin C)\} \\ &= \{x \mid (x \in A \wedge x \notin B) \vee (x \in A \wedge x \notin C)\} \\ &= (A \setminus B) \cup (A \setminus C). \end{aligned}$$

$\square$

Proposition 1.1 (Basic property of set operation)**.**
Let A, B, C, D, etc be sets. We have the following fact.

(a) $A \subset A \cup B$, $B \subset A \cup B$, $B \cup A = A \cup B$.

(b) $A \cap B \subset A$, $A \cap B \subset B$, $B \cap A = A \cap B$.

(c) $A \subset C \wedge B \subset C \equiv A \cup B \subset C$.

(d) $D \subset A \wedge D \subset B \equiv D \subset A \cap B$.

(e) $A \setminus B \subset A$.

(f) $A \cup A = A$.

(g) $A \cup \emptyset = A$.

(h) $A \cap A = A$.

(i) $A \cap \emptyset = \emptyset$.

(j) $A \setminus A = \emptyset$.

(k) $A \setminus \emptyset = A$.

(l) $\emptyset \setminus A = \emptyset$.

(m) $A = (A \setminus B) \cup (A \cap B)$.

(n) $A \cup B = (A \setminus B) \cup B$.

(o) $B \cap (A \setminus B) = \emptyset$.

(p) $A' \subset A \Rightarrow A' \setminus B \subset A \setminus B$.

(q) $B' \subset B \Rightarrow A \setminus B \subset A \setminus B'$.

(r) $A \setminus B = A \equiv A \cap B = \emptyset$.

(s) *(i)-(v) below are all equivalent to* $A \subset B$.

   (i) $A \cup B = B$.
   (ii) $A \cap B = A$.
   (iii) $A \setminus B = \emptyset$.
   (iv) $A \cup (B \setminus A) = B$.
   (v) $A = B \setminus (B \setminus A)$.

(t) $(A \cup B) \setminus C = (A \setminus C) \cup (B \setminus C)$.

(u) $(A \cap B) \setminus C = (A \setminus C) \cap (B \setminus C)$.

(v) $(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$.

(w) $(A \setminus B) \times (C \setminus D) = (A \times C \setminus B \times C) \setminus A \times D$.

Proof. (a) and (b) are obvious from the property of $\wedge, \vee$.

(c) ($\Leftarrow$) part is obvious by (a) . Conversely, if $x \in A \cup B$, then $x \in A \vee x \in B$, but in any case we have $x \in C$.

(d) ($\Leftarrow$) part comes from (b). Argue as (c) to show the converse.

(e) Observe $x \in A \wedge \neg(x \in B) \Rightarrow x \in A$.

(f) $A \subset A \Rightarrow A \subset A \cup A \subset A$ by (a) and (c).

(g) $A \subset A \wedge \emptyset \subset A \Rightarrow A \subset A \cup \emptyset \subset A$.

(h) $A \subset A \Rightarrow A \subset A \cap A \subset A$ by (b) and (d).

(i) $A \cap \emptyset \subset \emptyset$.

(j) The statement $x \in A \setminus A \Rightarrow x \in \emptyset$ is vacuously true.

(k) Note ${}^{\forall}x\,[x \notin \emptyset]$, which implies $A \subset A \setminus \emptyset$.

(l) $\emptyset \setminus A \subset \emptyset$.

(m) Distribution law for $\wedge, \vee$ gives

$$\begin{aligned} x \in A &\equiv x \in A \wedge (x \in B \vee x \notin B) \\ &\equiv (x \in A \wedge x \in B) \vee (x \in A \wedge x \notin B) \\ &\equiv x \in (A \cap B) \cup (A \setminus B). \end{aligned}$$

(n) Distribution law for $\wedge, \vee$ yields

$$\begin{aligned} x \in (A \setminus B) \cup B &\equiv (x \in A \wedge x \notin B) \vee x \in B \\ &\equiv x \in A \vee x \in B \\ &\equiv x \in A \cup B. \end{aligned}$$

(o) The statement $x \in B \cap (A \setminus B) \Rightarrow x \in \emptyset$ is vacuously true.

(p) Just check the definition.

(q) Note $B' \subset B \equiv {}^{\forall}x\,[x \in B' \Rightarrow x \in B] \equiv {}^{\forall}x\,[x \notin B \Rightarrow x \notin B']$, and argue as (p).

(r) Use (m) and (o).

(s) It is easy to see $B \setminus (B \setminus A) = A \cap B$, and so (ii) and (v) are equivalent. We show (iii) $\Rightarrow$ (ii) $\Rightarrow$ (iv) $\Rightarrow$ $(A \subset B)$ $\Rightarrow$ (i) $\Rightarrow$ (iii).

(m) gives (iii) $\Rightarrow$ (ii) $\Rightarrow$ (iv), and (iv) $\Rightarrow$ $(A \subset B)$ is trivial. $(A \subset B) \Rightarrow$ (i) is verified by seeing $A \subset B \Rightarrow B \subset A \cup B \subset B$. For the last implication, observe $A \setminus B = A \setminus (A \cup B) \subset A \setminus A = \emptyset$.

(t), (u), (v) and (w) are verified by chasing definition.

Exercise 1. Determine which of the following statements are true for all sets A, B, C and D. If a double implication fails, determine whether one or the other of the possible implications holds. If an equality fails, determine whether the statement becomes true if the "equals" symbols is replaced by one or the other of the inclusion symbols $\subset$ or $\supset$.

(a) $A \subset B$ and $A \subset C \Leftrightarrow A \subset (B \cup C)$.

(b) $A \subset B$ or $A \subset C \Leftrightarrow A \subset (B \cup C)$.

(c) $A \subset B$ and $A \subset C \Leftrightarrow A \subset (B \cap C)$.

(d) $A \subset B$ or $A \subset C \Leftrightarrow A \subset (B \cap C)$.

(e) $A \setminus (A \setminus B) = B$.

(f) $A \setminus (B \setminus A) = A \setminus B$.

(g) $A \cap (B \setminus C) = (A \cap B) \setminus (A \cap C)$.

(h) $A \cup (B \setminus C) = (A \cup B) \setminus (A \cup C)$.

(i) $(A \cap B) \cup (A \setminus B) = A$.

(j) $A \subset C$ and $B \subset D \Rightarrow (A \times B) \subset (C \times D)$.

(k) The converse of (j).

(l) The converse of (j), assuming that A and B are nonempty.

(m) $(A \times B) \cup (C \times D) = (A \cup C) \times (B \cup D)$.

(n) $(A \times B) \cap (C \times D) = (A \cap C) \times (B \cap D)$.

(o) $A \times (B \setminus C) = (A \times B) \setminus (A \times C)$.

(p) $(A \setminus B) \times (C \setminus D) = (A \times C \setminus B \times C) \setminus A \times D$.

(q) $(A \times B) \setminus (C \times D) = (A \setminus C) \times (B \setminus D)$.

Solution.

$\boxed{\text{(a)}}$ $(\Rightarrow)$ part is true by Proposition 1.1 and $B \cap C \subset B \cup C$.

$(\Leftarrow)$ part fails for the case, for instance, $A = \{2, 3\}, B = \{1, 2\}, C = \{2, 3\}$.

$\boxed{\text{(b)}}$ $(\Rightarrow)$ part is obviously true while the other part fails by the same counterexample given at (a).

$\boxed{\text{(c)}}$ This is true by Proposition 1.1 (d).

$\boxed{\text{(d)}}$ $(\Rightarrow)$ fails. A counterexample could be $A = \{1\}, B = \{1, 2\}, C = \emptyset$. The converse is true by Proposition 1.1 (d).

$\boxed{\text{(e)}}$ Proposition 1.1 (e) tells us that only $(\subset)$ part holds.

$\boxed{\text{(f)}}$ Proposition 1.1 (o) implies $A \cap (B \setminus A) = \emptyset$, which is equivalent to $A \setminus (B \setminus A) = A$ by Proposition 1.1 (r). Note that $A = A \setminus B$ if and only if $A \cap B = \emptyset$. Thus, only ($\supset$) is valid in general.

$\boxed{\text{(g)}}$ Equality holds as we see that

$$\begin{aligned} (A \cap B) \setminus (A \cap C) &= \{x \mid (x \in A \wedge x \in B) \wedge \neg(x \in A \wedge x \in C)\} \\ &= \{x \mid (x \in A \wedge x \in B) \wedge (x \notin A \vee x \notin C)\} \\ &= \{x \mid x \in A \wedge (x \in B \wedge x \notin C)\} \\ &= A \cap (B \setminus C). \end{aligned}$$

$\boxed{\text{(h)}}$ Only ($\supset$)holds. We deduce that

$$A \cup (B \setminus C) = (A \cup B) \setminus (C \setminus A).$$

It follows from the fact $C \setminus A \subset A \cup C$ and Proposition 1.1 (q) that $(A \cup B) \setminus (C \setminus A) \supset (A \cup B) \setminus (A \cup C)$.

A counterexample for ($\subset$) could be $A = B = C = \{1\}$.

$\boxed{\text{(i)}}$ Equality holds by Proposition 1.1 (m).

$\boxed{\text{(j)}}$ This is almost obvious as follows:

$$\begin{aligned} A \times B &= \{(a, b) \mid a \in A \wedge b \in B\} \\ &\subset \{(a, b) \mid a \in C \wedge b \in D\} \\ &= C \times D. \end{aligned}$$

$\boxed{\text{(k)}}$ The converse of (j) is false. Consider the case where B and C are empty while A and D are nonempty.

$\boxed{(\ell)}$ If $a \in A$ and $b \in B$, or equivalently, if $(a, b) \in A \times B$, then, by assumption, there holds $(a, b) \in C \times D$, from which we conclude $a \in C$ and $b \in D$.

$\boxed{\text{(m)}}$ Only ($\subset$) is the case. If $(x, y) \in (A \times B) \cup (C \times D)$, or equivalently, if

$$(x \in A \wedge y \in B) \vee (x \in C \wedge y \in D),$$

then, we deduce that $x \in A \cup C$ and $y \in B \cup D$, that is, $(x, y) \in (A \cup C) \times (B \cup D)$.

A counterexample for the other part could be: $A = B = \{1\}, C = D = \{2\}$. In that case, $(1, 2) \notin (A \times B) \cup (C \times D)$ but $(1, 2) \in (A \cup C) \times (B \cup D)$.

$\boxed{\text{(n)}}$ This is true by Proposition 1.1 (v).

$\boxed{\text{(o)}}$ This is true by Proposition 1.1 (w).

$\boxed{\text{(p)}}$ This is true by Proposition 1.1 (w).

$\boxed{\text{(q)}}$ Only ($\supset$) is valid. (p) gives

$$(A \setminus C) \times (B \setminus D) = ((A \times B) \setminus (A \times D)) \setminus (C \times D),$$

from which the ($\supset$) part follows.

As a counterexample, consider the case $A = B = \{1, 2\}, C = D = \{1\}$, where we see $(1, 2) \in (A \times B) \setminus (C \times D)$, but $(1, 2) \notin (A \setminus C) \times (B \setminus D)$. $\square$

Exercise 2.

(a) Write the contrapositive and converse of the following statement: "If $x < 0$, then $x^2 - x > 0$," and determine which (if any) of the three statements are true.

(b) Do the same thing for the statement "If $x > 0$, then $x^2 - x > 0$."

Solution.

$\boxed{\text{(a)}}$ The contrapositive statement is "If $x^2 - x \leq 0$, then $x \geq 0$." This is true, since we have $x^2 - x \leq 0$ if and only if $0 \leq x \leq 1$.

The converse is: "If $x^2 - x > 0$, then $x < 0$." This is false, since we see $x^2 - x > 0$ if and only if $x < 0$ or $1 < x$.

$\boxed{\text{(b)}}$ The contrapositive statement is "If $x^2 - x \leq 0$, then $x \leq 0$," and the converse is: "If $x^2 - x > 0$, then $x > 0$." They are both false for the same reason as (a). $\square$

Exercise 3. Let A and B be sets of real numbers. Write the negation of each of the following statements:

(a) For every $a \in A$, it is true that $a^2 \in B$.

(b) For at least one $a \in A$, it is true that $a^2 \in B$.

(c) For every $a \in A$, it is true that $a^2 \notin B$.

(d) For at least one $a \notin A$, it is true that $a^2 \in B$.

Solution.

$\boxed{\text{(a)}}$ For at least one $a \in A$, it is true that $a^2 \notin B$.

$\boxed{\text{(b)}}$ For every $a \in A$, it is true that $a^2 \notin B$.

$\boxed{\text{(c)}}$ For at least one $a \in A$, it is true that $a^2 \in B$.

$\boxed{\text{(d)}}$ For every $a \notin A$, it is true that $a^2 \notin B$. $\square$

Exercise 4. Let $\mathcal{A}$ be a nonempty collection of sets. Determine the truth of each of the following statements and of their converses:

(a) $x \in \bigcup_{A\in\mathcal{A}} A \Rightarrow x \in A$ for at least one $A \in \mathcal{A}$.

(b) $x \in \bigcup_{A\in\mathcal{A}} A \Rightarrow x \in A$ for every $A \in \mathcal{A}$.

(c) $x \in \bigcap_{A\in\mathcal{A}} A \Rightarrow x \in A$ for at least one $A \in \mathcal{A}$.

(d) $x \in \bigcap_{A\in\mathcal{A}} \Rightarrow x \in A$ for every $A \in \mathcal{A}$.

Solution.

$\boxed{\text{(a)}}$ The statement and its converse are both true by definition.

$\boxed{\text{(b)}}$ The converse is true by definition while the original statement is false in the case where $x \in A$ for only one $A \in \mathcal{A}$ and $x \notin A$ for others.

$\boxed{\text{(c)}}$ The original is trivially true by definition while the converse fails for the same reason as (b).

$\boxed{\text{(d)}}$ The statement and its converse are both true by definition. $\square$

Exercise 5. Write the contrapositive of each of the statements of Exercise 5.

Solution.

$\boxed{\text{(a)}}$ $x \notin A$ for every $A \in \mathcal{A} \Rightarrow x \notin \bigcup_{A\in\mathcal{A}} A$.

$\boxed{\text{(b)}}$ $x \notin A$ for at least one $A \in \mathcal{A} \Rightarrow x \notin \bigcup_{A\in\mathcal{A}} A$.

$\boxed{\text{(c)}}$ $x \notin A$ for every $A \in \mathcal{A} \Rightarrow x \notin \bigcap_{A\in\mathcal{A}} A$.

$\boxed{\text{(d)}}$ $x \notin A$ for at least one $A \in \mathcal{A} \Rightarrow x \notin \bigcap_{A\in\mathcal{A}} A$. $\square$

Exercise 6. Given sets A, B, and C, express each of the following sets in terms of A, B, and C, using the symbols $\cup$, $\cap$, and $\setminus$.

$$\begin{aligned} D &= \{x \mid x \in A \text{ and } (x \in B \text{ or } x \in C)\}, \\ E &= \{x \mid (x \in A \text{ and } x \in B) \text{ or } x \in C\}, \\ F &= \{x \mid x \in A \text{ and } (x \in B \Rightarrow x \in C)\}. \end{aligned}$$

Solution. It is obvious that we have

$$D = A \cap (B \cup C),$$

and

$$E = (A \cap B) \cup C.$$

For F, we deduce that

$$\begin{aligned} F &= \{x \mid x \in A \wedge (\neg(x \in B) \vee x \in C)\} \\ &= \{x \mid (x \in A \wedge x \notin B) \vee (x \in A \wedge x \in C)\} \\ &= (A \setminus B) \cup (A \cap C). \end{aligned}$$

$\square$

Exercise 7. If a set A has two elements, show that $\mathcal{P}(A)$ has four elements. How many elements does $\mathcal{P}(A)$ have if A has one element. Three elements? No elements? Why is $\mathcal{P}(A)$ called the power set of A?

Solution. It is easy to check that $\mathcal{P}(\{a\}) = \{\emptyset, \{a\}\}$, $\mathcal{P}(\{a, b\}) = \{\emptyset, \{a\}, \{b\}, \{a, b\}\}$.

In general, we claim that if a set A has n elements, then $\mathcal{P}(A)$ has 2^n elements. (This is why we call $\mathcal{P}(A)$ the "power" set.) Given a set A that has n elements, there are

$$\binom{n}{k} \left(:= \frac{n!}{(n-k+1)!k!} \right)$$

subsets of A that have k elements, from which we deduce that $\mathcal{P}(A)$ has

$$\sum_{k=1}^{n} \binom{n}{k} = (1+1)^n = 2^n$$

elements. $\square$

Exercise 8. Formulate and prove DeMorgan's laws for arbitrary unions and intersections.

Solution. Given a set X and a collection $\mathcal{A}$ of subsets of X, DeMorgan's laws are formulated as:

$$X \setminus \bigcup_{A \in \mathcal{A}} A = \bigcap_{A \in \mathcal{A}} (X \setminus A), \quad X \setminus \bigcap_{A \in \mathcal{A}} A = \bigcup_{A \in \mathcal{A}} (X \setminus A).$$

The proof is essentially the same as Exercise 1.

$$\begin{aligned} x \in X \setminus \bigcup_{A \in \mathcal{A}} A &\equiv x \in X \wedge x \notin \bigcup_{A \in \mathcal{A}} A \\ &\equiv x \in X \wedge \neg \left({}^{\exists} A \in \mathcal{A} \, [x \in A] \right) \\ &\equiv x \in X \wedge {}^{\forall} A \in \mathcal{A} \, [x \notin A] \\ &\equiv {}^{\forall} A \in \mathcal{A} \, [x \in X \setminus A] \\ &\equiv x \in \bigcap_{A \in \mathcal{A}} (X \setminus A). \end{aligned}$$

Similar argument works for the other implication. $\square$

Exercise 9. Let $\mathbb{R}$ denote the set of real numbers . For each of the following subsets of $\mathbb{R} \times \mathbb{R}$, determine whether it is equal to the cartesian product of two subsets of $\mathbb{R}$.

(a) $\{(x, y) \mid x \in \mathbb{Z}\}$.

(b) $\{(x, y) \mid 0 < y \leq 1\}$.

(c) $\{(x, y) \mid y > x\}$.

(d) $\{(x, y) \mid x \notin \mathbb{Z} \wedge y \in \mathbb{Z}\}$.

(e) $\{(x, y) \mid x^2 + y^2 < 1\}$.

Solution.

$\boxed{\text{(a)}}$ It is equal to $\mathbb{Z} \times \mathbb{R}$.

$\boxed{\text{(b)}}$ It is equal to $\mathbb{R} \times \{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$.

$\boxed{\text{(c)}}$ It is not equal to any cartesian product. Observe $(-1, 0)$ and $(0, 1)$ belong to the given set, but $(0, 0)$ does not, which violates an obvious necessary condition for the set to be a cartesian product.

$\boxed{\text{(d)}}$ $(\mathbb{R} \setminus \mathbb{Z}) \times \mathbb{Z}$.

$\boxed{\text{(e)}}$ Not equal to any cartesian product. $(3/4, 0)$ and $(0, 3/4)$ belong to the given set whereas $(3/4, 3/4)$ does not. $\square$

The necessary condition we use at Exercise 10(c) and (e) turns out to be a sufficient condition as follows:

Proposition 1.2. *Let X and Y be sets, and let A be a subset of $X \times Y$. A necessary and sufficient condition for A to be the cartesian product of a subset of X and that of Y is that there holds*

$$(x_1, y_1), (x_2, y_2) \in A \Rightarrow (x_i, y_j) \in A$$

for any $i, j = 1, 2$.

Proof. Necessity is obvious. To prove sufficiency, suppose the given condition holds. Define two sets X_1 and Y_1 by setting

$$\begin{aligned} X_1 &:= \left\{x \in X \mid {}^{\exists} y \in Y \, [(x, y) \in A]\right\}, \\ Y_1 &:= \left\{y \in Y \mid {}^{\exists} x \in X \, [(x, y) \in A]\right\}. \end{aligned}$$

We show $A = X_1 \times Y_1$. It is easy to check that $A \subset X_1 \times Y_1$. Let $(x, y) \in X_1 \times Y_1$, or equivalently, let $x \in X_1$ and $y \in Y_1$. By definition, there exist $y_0 \in Y$ such that $(x, y_0) \in A$, and $x_0 \in X$ such that $(x_0, y) \in A$. Then, it follows from the assumption that $(x, y) \in A$, which implies $A \supset X_1 \times Y_1$. Thus, $A = X_1 \times Y_1$.

2 Functions

Remark 2.1 (Empty Function)**.** You can skip this note. Nothing proved in this note is used when working on exercises in §2. (We need the concept of empty function to let several exercise make sense; principle of recursive definition is one such example; see §8 Exercise 8, §10 Exercise 10.) We consider here whether we are able to define a function from or to empty set $\emptyset$. Remembering a function is defined to be a certain subset of the catersian product of two sets, the question is reduced to the one that "could empty set qualify as a function, and if so, in what cases?" The answer is partially "yes"; there exist a unique function from $\emptyset$ to $\emptyset$, and a unique function from $\emptyset$ to nonempty set, each called an *empty function*, but no way could there be a function from nonempty set to $\emptyset$.

First of all, note that the definition of function is expressed as follows:

$$f \in \mathrm{Func}(X, Y) \equiv [f \subset X \times Y] \wedge {}^{\forall}x \in X^{\exists 1}y \in Y\,[(x, y) \in f]\,,$$

where X and Y are sets, of course. Suppose $X = \emptyset$. We claim $\mathrm{Func}(\emptyset, Y) = \{\emptyset\}$. In fact, we see that

$$f \subset \emptyset \times Y \equiv f = \emptyset,$$

and that the statements

$${}^{\forall}x \in X^{\exists 1}y \in Y\,[(x, y) \in f] \equiv {}^{\forall}x\left[x \in \emptyset \Rightarrow {}^{\exists 1}y \in Y\,[(x, y) \in \emptyset]\right]$$

are true since the latter is vacuously true. Note that this argument is valid if $Y = \emptyset$. Thus, the claim follows.

On the other hand, suppose $X \neq \emptyset$ and $Y = \emptyset$. We now insist $\mathrm{Func}(X, \emptyset) = \emptyset$. It suffices to show that $f = \emptyset$ does not qualify as a function. Indeed, the statements

$${}^{\forall}x \in X^{\exists 1}y \in Y\,[(x, y) \in f] \equiv {}^{\forall}x \in X^{\exists 1}y\,[y \in \emptyset \wedge (x, y) \in \emptyset]$$

are both false since, in general, there holds

$${}^{\forall}z\,[z \notin \emptyset]$$

and so, $y \in \emptyset$ is false. $\square$

Remark 2.2 (Equivalent conditions for injectivity and surjectivity)**.** We summarize here some equivalent expressions for injectivity and surjectivity, which we frequently exploit for establishing results that are related to those property. Let $f \in \mathrm{Func}(A, B)$.

For injectivity, we claim that

$$\begin{aligned} f \in \mathrm{Inj}(A,B) &\equiv {}^{\forall}a_1 \in A\, {}^{\forall}a_2 \in A\,[a_1 \neq a_2 \Rightarrow f(a_1) \neq f(a_2)] \\ &\equiv {}^{\forall}a_1 \in A\, {}^{\forall}a_2 \in A\,[f(a_1) = f(a_2) \Rightarrow a_1 = a_2] \\ &\equiv {}^{\forall}a \in A\, {}^{\forall}A_0 \subset A\,[f(a) \in f(A_0) \Leftrightarrow a \in A_0] \\ &\equiv {}^{\forall}a \in A\, {}^{\forall}A_0 \subset A\,[f(a) \notin f(A_0) \Leftrightarrow a \notin A_0]\,. \end{aligned}$$

First two equivalence are obvious by definition. We prove the third. Suppose $f \in \mathrm{Inj}(A,B)$. Let $a \in A$ and let A_0 be a subset of A. It is obvious that we have

$$a \in A_0 \Rightarrow f(a) \in f(A_0).$$

So, suppose $f(a) \in f(A_0)$. Then, there exists $x \in A_0$ such that $f(a) = f(x)$. But injectivity gives $x = a$ and so $a \in A_0$. Conversely, suppose (2.1) holds. Let $a, a' \in A$ with $f(a) = f(a')$. Setting $A_0 := \{a'\}$ in (2.1) yields $a \in \{a'\}$, which means $a = a'$ establishing the injectivity.

For surjectivity, we insist that

$$\begin{aligned} f \in \mathrm{Surj}(A,B) &\equiv f(A) = B \\ &\equiv {}^{\forall}b \in B\, {}^{\exists}a \in A\,[f(a) = b] \\ &\equiv {}^{\forall}b \in B\,\left[f^{-1}(b) \neq \emptyset\right]. \end{aligned}$$

First equivalence is due to the definition. Second owes to the definition of image, third to that of preimage.

We can see more equivalent conditions as we proceed in this section. □

Remark 2.3 (Injectivity of Empty function)**.** Let $f_\emptyset : \emptyset \to Y$ be an empty function. $f_\emptyset$ is injective, but not surjective. In Fact, we see that injectivity is vacuously satisfied, and that there holds

$${}^{\forall}y \in Y\,\left[f_\emptyset^{-1}(y) = \emptyset\right]$$

since, in general, preimage is a subset of the domain of a function. Thus, $f_\emptyset$ is not surjective. □

Exercise 1. Let $f : A \to B$. Let $A_0 \subset A$ and $B_0 \subset B$.

(a) Show that $A_0 \subset f^{-1}(f(A_0))$ and that equality holds if f is injective.

(b) Show that $f(f^{-1}(B_0)) \subset B_0$ and that equality holds if f is surjective.

Solution.

$\boxed{\text{(a)}}$ Let $a \in A_0$. It follows that $f(a) \in f(A_0)$, and that $f^{-1}(f(a)) \subset f^{-1}(f(A_0))$. It is clear that $a \in f^{-1}(f(a))$, and so $a \in f^{-1}(f(A_0))$. Thus, $A_0 \subset f^{-1}(f(A_0))$.

We prove the second implication. In general, we trivially have

$$^{\forall}a \in f^{-1}(f(A_0))^{\exists}a' \in A_0 \, [f(a) = f(a')] \, .$$

So, if we assume $f \in \mathrm{Inj}(A, B)$, then $a = a' \in A_0$. Thus, $f^{-1}(f(A_0)) \subset A_0$.

$\boxed{\text{(b)}}$ In general, we have,

$$^{\forall}b \in f\left(f^{-1}(B_0)\right)^{\exists}a \in f^{-1}(B_0) \, [f(a) = b] \, ,$$

which yields, by definition of preimage, $b = f(a) \in B_0$. Thus, $f(f^{-1}(B_0)) \subset B_0$.

Suppose $f \in \mathrm{Surj}(A, B)$. Then we have

$$^{\forall}b \in B_0{}^{\exists}a \in f^{-1}(B_0) \, [f(a) = b] \, .$$

Hence, $b = f(a) \in f(f^{-1}(B_0))$. $\square$

Remark 2.4 (Another equivalent condition for injectivity and surjectivity)**.** We prove the converse of each of second implication of Exercise 1(a) and (b). Let $f \in \mathrm{Func}(A, B)$, and let A_0 be a subset of A, and B_0 be a subset of B.

Suppose $A_0 = f^{-1}(f(A_0))$ holds. For $a_1, a_2 \in A_0$ with $f(a_1) = f(a_2)$, we deduce

$$\begin{aligned} \{a_1\} &= f^{-1}(f(a_1)) \\ &= f^{-1}(f(a_2)) \\ &= \{a_2\} \, . \end{aligned}$$

Thus, f is injective.

To see the converse of Exercise 2(b), Set $B_0 := B$ and obtain $f(f^{-1}(B)) = B$. Thus, f is surjective. $\square$

Exercise 2. Let $f : A \to B$ and let $A_i \subset A$ and $B_i \subset B$ for $i = 0$ and $i = 1$. Show that f^{-1} preserves inclusions, unions , intersections, and differences of sets;

(a) $B_0 \subset B_1 \Rightarrow f^{-1}(B_0) \subset f^{-1}(B_1)$.

(b) $f^{-1}(B_0 \cup B_1) = f^{-1}(B_0) \cup f^{-1}(B_1)$.

(c) $f^{-1}(B_0 \cap B_1) = f^{-1}(B_0) \cap f^{-1}(B_1)$.

(d) $f^{-1}(B_0 \setminus B_1) = f^{-1}(B_0) \setminus f^{-1}(B_1)$.

Show that f preserves inclusions and unions only;

(e) $A_0 \subset A_1 \Rightarrow f(A_0) \subset f(A_1)$.

(f) $f(A_0 \cup A_1) = f(A_0) \cup f(A_1)$.

(g) $f(A_0 \cap A_1) \subset f(A_0) \cap f(A_1)$; show that equality holds if f is injective.

(h) $f(A_0 \setminus A_1) \supset f(A_0) \setminus f(A_1)$; show that equality holds if f is injective.

Solution. We can safely omit and actually do Exercise 2(b), (c) and (f) since we are anyway required to show the generalized versions of them in Exercise 3. The proof there of course proves the (b), (c), (f) of Exercise 2.

Let $f \in \mathrm{Func}(A, B)$, and let A_i and B_i be subsets of A and B, respectively, for $i = 0, 1$.

$\boxed{\text{(a)}}$ Let $a \in f^{-1}(B_0)$. By definition and hypothesis, we have $f(a) \in B_0 \subset B_1$, which means $a \in f^{-1}(B_1)$. Thus, $f^{-1}(B_0) \subset f^{-1}(B_1)$.

$\boxed{\text{(d)}}$ Observe

$$\begin{aligned} a \in f^{-1}(B_0 \setminus B_1) &\equiv f(a) \in B_0 \setminus B_1 \\ &\equiv f(a) \in B_0 \wedge \neg\,(f(a) \in B_1) \\ &\equiv a \in f^{-1}(B_0) \wedge \neg\,\bigl(a \in f^{-1}(B_1)\bigr) \\ &\equiv a \in f^{-1}(B_0) \setminus f^{-1}(B_1), \end{aligned}$$

which completes the proof.

$\boxed{\text{(e)}}$ Let $b \in f(A_0)$. Ther exists $a \in A_0$ such that $b = f(a)$. Since hypothesis gives $a \in A_1$, we conclude $b = f(a) \in f(A_1)$. Thus, $f(A_0) \in f(A_1)$.

$\boxed{\text{(g)}}$ Since $A_0 \cap A_1 \subset A_0$ and $A_0 \cap A_1 \subset A_1$, (e) yields $f(A_0 \cap A_1) \subset f(A_0)$ and $f(A_0 \cap A_1) \subset f(A_1)$, and so $f(A_0 \cap A_1) \subset f(A_0) \cap f(A_1)$.

We proceed to proving the second claim. In general, we have

$$^{\forall} b \in f(A_0) \cap f(A_1)\, ^{\exists} a_i \in A_i\, [b = f(a_i)]$$

for $i = 0, 1$. So, if we assume $f \in \mathrm{Inj}(A, B)$, then $b = f(a_0) = f(a_1)$ implies $a_0 = a_1$. Setting $a := a_0 (= a_1)$, we see that $a \in A_0 \cap A_1$ and that $b = f(a) \in f(A_0 \cap A_1)$. This establishes the claim.

$\boxed{\text{(h)}}$ We generally see that

$$\begin{aligned} b \in f(A_0) \setminus f(A_1) &\equiv {}^{\exists}a\,[a \in A_0 \wedge f(a) = b \wedge f(a) \in f(A_0) \wedge f(a) \notin f(A_1)] \\ &\equiv {}^{\exists}a\,[a \in A_0 \wedge f(a) = b \wedge f(a) \notin f(A_1)] \\ &\Rightarrow {}^{\exists}a\,[a \in A_0 \wedge f(a) = b \wedge a \notin A_1] \\ &\equiv b \in f(A_0 \setminus A_1), \end{aligned}$$

where we note that, in general, there holds

$$[f(x) \notin f(P) \Rightarrow x \notin P] \equiv [x \in P \Rightarrow f(x) \in f(P)]\,,$$

and that RHS is obviously true.

Next we show the second implication. If we assume that f is injective, or equivalently, if we assume

$$f(x) \notin f(P) \Leftrightarrow x \notin P,$$

holds for any subset P of B, then we can replace "$\Rightarrow$" with "$\equiv$" at (2.3), which proves the result. $\square$

Remark 2.5 (Yet another equivalent condition for injectivity)**.** We prove here the converse of second implications of Exercise 2(g) and (h). In other words, we claim that

$$\begin{aligned} f \in \mathrm{Inj}(A, B) &\equiv {}^{\forall}A_0 \subset A\,{}^{\forall}A_1 \subset A\,[f(A_0 \cap A_1) = f(A_0) \cap f(A_1)] \\ &\equiv {}^{\forall}A_0 \subset A\,{}^{\forall}A_1 \subset A\,[f(A_0) \setminus f(A_1) = f(A_0 \setminus A_1)]\,. \end{aligned}$$

We have already proved $[f \in \mathrm{Inj}(A, B) \Rightarrow (2.4)]$ and $[f \in \mathrm{Inj}(A, B) \Rightarrow (2.5)]$. So, we establish each of the converse. Let $a_0, a_1 \in A$ with $f(a_0) = f(a_1)$. If we assume (2.4), then $f(\{a_0\} \cap \{a_1\}) \neq \emptyset$ and so $\{a_0\} = \{a_1\}$; if we assume (2.5), then $f(\{a_0\} \setminus \{a_1\}) = \emptyset$ and so $\{a_0\} \setminus \{a_1\} = \emptyset$, which means $\{a_0\} = \{a_1\}$. $\square$

Exercise 3. Show that (b), (c), (f), and (g) of Exercise 2 hold for arbitrary unions and intersections.

Solution. Let $f \in \mathrm{Func}(A', B')$, and let A be a subset of A' for all $A \in \mathcal{A}$, and B be a subset of B' for all $B \in \mathcal{B}$.

$\boxed{\text{(b)}}$ Observe that $\bigcup_{B \in \mathcal{B}} B \supset B$ for all $B \in \mathcal{B}$, and that, by Exercise 2(a), $f^{-1}\left(\bigcup_{B \in \mathcal{B}} B\right) \supset f^{-1}(B)$ for all $B \in \mathcal{B}$, which yields $f^{-1}\left(\bigcup_{B \in \mathcal{B}} B\right) \supset \bigcup_{B \in \mathcal{B}} f^{-1}(B)$. On the other hand, we have, in general,

$${}^{\forall}a \in f^{-1}\left(\bigcup_{B \in \mathcal{B}} B\right) {}^{\exists}B_0 \in \mathcal{B}\,[f(a) \in B_0]\,,$$

from which it follows that $a \in f^{-1}(f(a)) \subset f^{-1}(B_0) \subset \bigcup_{B \in \mathcal{B}} f^{-1}(B)$. Thus, the result follows.

$\boxed{\text{(c)}}$ As we have done in (b), we deduce $f^{-1}\left(\bigcap_{B \in \mathcal{B}} B\right) \subset \bigcap_{B \in \mathcal{B}} f^{-1}(B)$. We show the opposite inclusion. Let $a \in \bigcap_{B \in \mathcal{B}} f^{-1}(B)$. Definition of intersection gives $a \in f^{-1}(B)$ for all $B \in \mathcal{B}$, or equivalently $f(a) \in B$ for all $B \in \mathcal{B}$. This means $f(a) \in \bigcap_{B \in \mathcal{B}} B$, and hence $a \in f^{-1}\left(\bigcap_{B \in \mathcal{B}} B\right)$. Thus, $\bigcap_{B \in \mathcal{B}} f^{-1}(B) \subset f^{-1}\left(\bigcap_{B \in \mathcal{B}} B\right)$.

$\boxed{\text{(f)}}$ As before, we deduce that $f\left(\bigcup_{A \in \mathcal{A}} A\right) \supset \bigcup_{A \in \mathcal{A}} f(A)$.

On the other hand, we see, in general, that

$$^{\forall} b \in f\left(\bigcup_{A \in \mathcal{A}} A\right) {}^{\exists} A_0 \in \mathcal{A} {}^{\exists} a \in A_0 \, [b = f(a)],$$

and so $b = f(a) \in f(A_0) \subset \bigcup_{A \in \mathcal{A}} f(A)$.

$\boxed{\text{(g)}}$ The fact $\bigcap_{A \in \mathcal{A}} A \subset A$ for all $A \in \mathcal{A}$ implies $f\left(\bigcap_{A \in \mathcal{A}} A\right) \subset f(A)$. The same argument as Exercise 2 shows the equality holds if f is injective. □

Exercise 4. Let $f : A \to B$ and $g : B \to C$.

(a) If $C_0 \subset C$, show that $(g \circ f)^{-1}(C_0) = f^{-1}(g^{-1}(C_0))$.

(b) If f and g are injective, show that $g \circ f$ is injective.

(c) If $g \circ f$ is injective, what can you say about injectivity of f and g.

(d) If f and g are surjective, show that $g \circ f$ is surjective.

(e) If $g \circ f$ is surjective, what can you say about surjectivity of f and g.

(f) Summarize your answers to (b)–(e) in the form of a theorem.

Solution. Let $f \in \mathrm{Func}(A, B)$ and $g \in \mathrm{Func}(B, C)$.

$\boxed{\text{(a)}}$ Observe that

$$\begin{aligned}
(g \circ f)^{-1}(C_0) &= \{a \in A \mid (g \circ f)(a) \in C_0\} \\
&= \left\{a \in A \mid f(a) \in g^{-1}(C_0)\right\} \\
&= f^{-1}\left(g^{-1}(C_0)\right).
\end{aligned}$$

$\boxed{\text{(b)}}$ Suppose f and g are injective.

For any $a, a' \in A$ with $a \neq a'$, injectivity gives $f(a) \neq f(a')$, and $g(f(a)) \neq g(f(a'))$. Hence, $g \circ f \in \mathrm{Inj}(A, C)$.

$\boxed{\text{(c)}}$ Suppose $g \circ f \in \operatorname{Inj}(A, C)$. We claim in this case that f is injective. For $a, a' \in A$ with $f(a) = f(a')$, we have $g \circ f(a) = g \circ f(a')$, which means, by assumption, $a = a'$. Thus, f is injective.

$\boxed{\text{(d)}}$ Suppose f and g are surjective. In general, (a) implies that for every $c \in C$, we have $(g \circ f)^{-1}(c) = (f^{-1} \circ g^{-1}(c)) = (f^{-1}(g^{-1}(c)))$. Assumption then yields $g^{-1}(c) \neq \emptyset$ and $f^{-1}(g^{-1}(c)) \neq \emptyset$. Thus, $g \circ f$ is surjective.

$\boxed{\text{(e)}}$ Suppose $g \circ f \in \operatorname{Surj}(A, C)$. We insist that g is surjective. For every $c \in C$, assumption implies that $(g \circ f)^{-1}(c) = f^{-1}(g^{-1}(c))$ is nonempty, which necessarily means that $g^{-1}(c) \neq \emptyset$. Thus, g is surjective.

$\boxed{\text{(f)}}$ We summarize (a)-(e) to state the following:

Theorem 2.6. *Let $f \in \operatorname{Func}(A, B)$ and $g \in \operatorname{Func}(B, C)$. If both f and g are injective (or surjective), so is $g \circ f$. Conversely, If $g \circ f$ is injective, so is f; and if $g \circ f$ is surjective, so is g.* □

Exercise 5. In general, let us denote the ***identity function*** for a set C by i_C. That is, define $i_C : C \to C$ to be the function given by the rule $i_C(x) = x$ for all $x \in C$. Given $f : A \to B$, we say that a function $g : B \to A$ is a ***left inverse*** for f if $g \circ f = i_A$; and we say that $h : B \to A$ is a ***right inverse*** for f if $f \circ h = i_B$.

(a) Show that if f has a left inverse, f is injective; and if f has a right inverse, f is surjective.

(b) Give an example of a function that has a a left inverse but no right inverse.

(c) Give an example of a function that has a a right inverse but no left inverse.

(d) Can a function have more than one left inverse? More than one right inverse?

(e) Show that if f has both a left inverse g and a right inverse h, then f is bijective and $g = h = f^{-1}$.

Solution.

$\boxed{\text{(a)}}$ This is a direct consequence of Exercise 4(c) and (e).

$\boxed{\text{(b)}}$ A function given by

$$f : \mathbb{R}_+ \ni x \mapsto \exp x \in \mathbb{R}_+$$

admits a left inverse

$$\ell : \mathbb{R}_+ \to \mathbb{R}_+ : x \mapsto \begin{cases} 1 & : \ x \in (0,1) \\ \log x & : \ x > 1, \end{cases}$$

but no right inverse.

$\boxed{\text{(c)}}$ A function given by

$$g : \mathbb{R} \to \{0,1\} : x \mapsto \begin{cases} 1 & : \ x \in \mathbb{Q} \\ 0 & : \ x \notin \mathbb{Q} \end{cases}$$

has a right inverse

$$r : \{0,1\} \to \mathbb{R} : x \mapsto \begin{cases} 1 & : \ x = 1 \\ e & : \ x = 0, \end{cases}$$

but no left inverse.

$\boxed{\text{(d)}}$ Yes. For instance, setting $\ell := 1/2$ on $[0,1)$ in (b), and $r(0) := \pi$ in (c) does not spoil their inverse property.

$\boxed{\text{(e)}}$ Apply Lemma 2.1. □

Exercise 6. Let $f : \mathbb{R} \to \mathbb{R}$ be the function $f(x) = x^3 - x$. By restricting the domain and range of f appropriately, obtain from f a bijective function g. Draw the graphs of g and g^{-1}. There are several possible choices for g.)

Solution. Consider a strictly increasing function given by

$$g : (1, \infty) \ni x \mapsto x^3 - x \in \mathbb{R}_+.$$

g possesses a inverse

$$g^{-1} : \mathbb{R}_+ \ni x \mapsto \frac{1}{x^3 - x} \in (1, \infty).$$

The drawing part is left to readers. □

3 Relations

Equivalence Relations

Exercise 1. Define two points (x_0, y_0) and (x_1, y_1) of the plane to be equivalent if $y_0 - x_0^2 = y_1 - x_1^2$. Check that this is an equivalence relation and describe the equivalence classes.

Solution. Concerning set here is a subset of $\mathbb{R} \times \mathbb{R}$. So, it is a relation. reflexivity and symmetry are obviously satisfied. If $(x_0, y_0) \sim (x_1, y_1)$ and $(x_1, y_1) \sim (x_2, y_2)$, then $y_0 - {x_0}^2 = y_1 - {x_1}^2 = y_2 - {x_2}^2$, that is, $(x_0, y_0) \sim (x_2, y_2)$. Thus, the given relation satisfies transitivity, and is an equivalence relation.

The collection of equivalence classes consists of all parabolas that are symmetric with respect to y-axis. □

Exercise 2. Let C be a relation on a set A. If $A_0 \subset A$, define the ***restriction*** of C to A_0 to be the relation $C \cap (A_0 \times A_0)$. Show that the restriction of an equivalence relation is an equivalence relation.

Solution. Observe $D := C \cap (A_0 \times A_0)$ is a subset of $A_0 \times A_0$ and so a relation on A_0.

(Reflexivity) If $x \in A_0$, then $(x, x) \in D$ since $(x, x) \in C$. Hence, xDx.

(Symmetry) If xDy, or equivalently, if $(x, y) \in D$, then $(x, y) \in C$ and $(x, y) \in A_0 \times A_0$. Symmetry of C gives $(y, x) \in C$. It follows from $(y, x) \in A_0 \times A_0$ that $(y, x) \in C \cap (A_0 \times A_0) = D$.

(Transitivity) If xDy and yDz, or equivalently, if $(x, y) \in D$ and $(y, z) \in D$, similar argument shows that $(x, z) \in C$ Since $(x, z) \in A_0 \times A_0$, we have $(x, z) \in D$, that is, xDy. □

Exercise 3. Here is an "proof" that every relation C that is both symmetric and transitive is also reflexive: "Since C is symmetric, aCb implies bCa. Since C is a transitive, aCb and bCa together imply aCa, as desired." Find the flaw in this argument.

Solution. Suppose C is a relation on A. Note that C is reflexive if

$$^{\forall}a\,[a \in A \Rightarrow aCa]$$

is true while what is shown in the given "proof" is

$$^{\forall}a\,[aCb \Rightarrow aCa]\,.$$

□

Exercise 4. Let $f : A \to B$ be a surjective function. Let us define a relation on A by setting $a_0 \sim a_1$ if

$$f(a_0) = f(a_1).$$

(a) Show that this is an equivalence relation.

(b) Let A^* be the set of equivalence classes. Show that there is a bijective correspondence of A^* with B.

Solution. Let $f \in \mathrm{Surj}(A, B)$.

$\boxed{\text{(a)}}$ $C := \{(a_0, a_1) \in A \times A \mid f(a_0) = f(a_1)\}$ is a subset of $A \times A$, that is, a relation on A. We show that C satisfies reflexivity, symmetry, transitivity.

(Reflexivity) ${}^{\forall}a \in A\,[f(a) = f(a)]$ means ${}^{\forall}a \in A\,[aCa]$

(Symmetry) For $a_0, a_1 \in A$, we have

$$a_0 C a_1 \equiv f(a_0) = f(a_1) \equiv f(a_1) = f(a_0) \equiv a_1 C a_0.$$

(Transitivity) For $a_0, a_1, a_2 \in A$, we have

$$a_0 C a_1 \wedge a_1 C a_2 \equiv f(a_0) = f(a_1) \wedge f(a_1) = f(a_2) \Rightarrow f(a_0) = f(a_2) \equiv a_0 C a_2.$$

$\boxed{\text{(b)}}$ Let A^* be the set of equivalence classes. The definition of equivalence class leads to

$$ {}^{\forall}E \in A^{*\,\exists 1}b \in B^{\forall}a \in E\,[b = f(a)] \tag{3.1}$$

We first establish this. Let $E \in A^*$. By definition, we have

$$ {}^{\forall}a \in E^{\forall}a' \in E\,[f(a) = f(a')]\,.$$

Setting this common values as b, that is, setting $b := f(a)$, this b has the required property. (3.1) allows us to define a function $g : A^* \ni E \mapsto b \in B$. Since each of equivalence classes is disjoint one another, g is injective. Moreover, surjectivity of f gives

$$ {}^{\forall}b \in B^{\exists}a \in A\,[f(a) = b]\,,$$

which yields

$$ {}^{\forall}b \in B^{\exists}E \in A^{*\,\forall}a \in E\,[f(a) = b]\,,$$

from which we conclude

$$ {}^{\forall}b \in B^{\exists}E \in A^*\,[g(E) = b]\,.$$

Thus, g is surjective. □

Exercise 5. Let S and S' be the following subsets of the plane.

$$\begin{aligned} S &= \{(x, y) \mid y = x + 1 \text{ and } 0 < x < 2\}\,, \\ S' &= \{(x, y) \mid y - x \in \mathbb{Z}\}\,. \end{aligned}$$

(a) Show that S' is an equivalence relation on the real line and $S' \supset S$. Describe the equivalence classes.

(b) Show that given any collection of equivalence relations on a set A, their intersection is an equivalence relation on A.

(c) Describe the equivalence relation T on the real line that is the intersection of all equivalence relation on the real line that contains S. Describe the equivalence classes of T.

Solution.

$\boxed{\text{(a)}}$ We first claim that $S \subset S'$ since for every $(x, y) \in S$, we deduce that

$$(x, y) \in S \equiv (y = x + 1) \wedge (0 < x < 2) \Rightarrow y - x = 1 \Rightarrow (x, y) \in S'.$$

S' is clearly a relation by definition. We show that S' is an equivalence relation.

(Reflexivity) This is obvious since we see that $x - x = 0$ is an integer for every $x \in \mathbb{R}$, and so $(x, x) \in S'$.

(Symmetry) This is also obvious since if $y - x$ is an integer, so is $x - y$, establishing the fact that $(x, y) \in S'$ implies $(y, x) \in S'$.

(Transitivity) Observe that if $y - x$ and $z - y$ are integers, so is $z - x = (z - y) + (y - x)$. This proves the transitivity.

$\boxed{\text{(b)}}$ Let $\mathcal{C}$ be a collection of equivalence relations on A. Note that $\bigcap_{C \in \mathcal{C}} C$ is a relation on A since every member of $\mathcal{C}$ is a subset of $A \times A$. We proceed to showing that $\bigcap_{C \in \mathcal{C}} C$ is an equivalence relation.

(Reflexivity) For every $a \in A$, we have $(a, a) \in C$ for all $C \in \mathcal{C}$. Hence, we conclude $(a, a) \in \bigcap_{C \in \mathcal{C}} C$.

(Symmetry) Let $(a, a') \in \bigcap_{C \in \mathcal{C}} C$. It follows that $(a, a') \in C$ and $(a', a) \in C$ for all $C \in \mathcal{C}$ by symmetry of each C. Thus, have $(a', a) \in \bigcap_{C \in \mathcal{C}} C$.

(Transitivity) Apply similar argument to prove the validity of transitivity.

$\boxed{\text{(c)}}$ T is the smallest equivalence relation on the real line that containis S, which is given by

$$T = T_2 \cup T_1 \cup T_0 \cup T_{-1} \cup T_{-2},$$

where

$$\begin{aligned} T_2 &:= \{(x, y) \mid y = x + 2 \wedge 0 < x < 1\} \\ T_1 &:= \{(x, y) \mid y = x + 1 \wedge 0 < x < 2\} \\ T_0 &:= \{(x, x) \mid x \in \mathbb{R}\} \\ T_{-1} &:= \{(x, y) \mid y = x - 1 \wedge 1 < x < 3\} \\ T_{-2} &:= \{(x, y) \mid y = x - 2 \wedge 2 < x < 3\}. \end{aligned}$$

The equivalence classes are: T_0 and $T_{-1} \cup T_1$ and $T_{-2} \cup T_2$. $\square$

Order Relations

Let us introduce a notation; Let $(A, <)$ denote the ordered set A equipped with order relation $<$. If we say "let $(A, <)$ be a ordered set", we mean that a ordered set A is given whose order relation is defined to be $<$.

Definition 3.1 (Order isomorphism)**.** *Let A and B be simply ordered sets, and let $f : A \to B$ be a function. We say that f is an order isomorphism from A to B if it is surjective and order preserving.*

Proposition 3.2 (Sufficient condition for order isomorphism)**.**
Let $(A, <_A)$ and $(B, <_B)$ be ordered sets, and let $f : A \to B$ be a bijective function. We claim that f is order-preserving, that is, an order isomorphism if and only if f^{-1} is.

Proof. Note that there holds

$$\begin{aligned}
& {}^\forall a_1 \in A {}^\forall a_2 \in A \left[a_1 <_A a_2 \Rightarrow f(a_1) <_B f(a_2)\right] \\
\equiv\ & {}^\forall a_1 \in A {}^\forall a_2 \in A \left[f(a_1) \geq_B f(a_2) \Rightarrow a_1 \geq_A a_2\right] \\
\equiv\ & {}^\forall a_1 \in A {}^\forall a_2 \in A \left[f(a_1) >_B f(a_2) \Rightarrow a_1 >_A a_2\right].
\end{aligned}$$

So, for $f \in \mathrm{Bij}(A, B)$, we conclude that f preserves order if and only if f^{-1} does.

We summarize here the property of order isomorphism, which insists, roughly speaking, that an order isomorphism preserves the property of an ordered set such as an immediate predecessor (successor), a smallest (largest) element, an upper (lower) bound, a supremum (infimum), and a section (see §4).

Proposition 3.3 (Property of sets of the same order type)**.**
Let $(A, <_A)$ and $(B, <_B)$ be an ordered sets of the same order type with an order isomorphism $f : A \to B$. We insist that

(a) A subset A_0 of A is bounded above (below) by an element u of A if and only if $f(A_0)$ is bounded above (below) by $f(u)$.

(b) An element a of A is a smallest(largest) element of subset A_1 of A if and only if $f(a)$ is that of $f(A_1)$.

(c) A subset P of A has a supremum s if and only if $f(P)$ has the supremum $f(s)$.

(d) An element i of A has an immediate predecessor (successor) if and only if $f(i)$ does.

(e) *Let α be an element of A; let $S_\alpha(A)$ denote the set of the elements of A less than α. We call $S_\alpha(A)$ a section of A by α. For every α, we claim $f(S_\alpha(A)) = S_{f(\alpha)}(B)$, that is,*

the image of a section of ordered set by an element under an order isomorphism is the section of the range by the value of the isomorphism at that element.

Proof. (a) and (e) are direct consequences of Proposition 3.2.

(b) is easy to show; just note

$$\begin{aligned} {}^\forall x \in A_1\,[a \leq x] &\equiv {}^\forall x \in A_1\,[f(a) \leq f(x)] \\ &\equiv {}^\forall y \in f(A_1)\,[f(a) \leq y]\,. \end{aligned}$$

The proof for a largest element is now trivial.

To prove (c), note that s is a smallest element of the set of all upper bounds for P, which is equivalent, by (a) and (b), to the statement that $f(s)$ is a smallest element of the set of all upper bounds for $f(P)$. Thus, (c) follows.

Lastly, consider (d). suppose $i \in A$ has an immediate successor i'. Note that

$$\begin{aligned} (i, i') = \emptyset &\equiv {}^\forall x \in A\,[x \notin (i, i')] \\ &\equiv {}^\forall x \in A\,[f(x) \notin f((i, i'))] \\ &\equiv {}^\forall x \in A\,[f(x) \notin (f(i), f(i'))] \\ &\equiv {}^\forall y \in B\,[y \notin (f(i), f(i'))] \\ &\equiv (f(i), f(i')) = \emptyset. \end{aligned}$$

Second equivalence comes from $f \in \mathrm{Inj}(A, B)$, third from the assumption that f is an order isomorphism, fourth from $f \in \mathrm{Surj}(A, B)$. Thus, we complete the proof for (d).

Exercise 6. Define a relation on the plane by setting

$$(x_0, y_0) < (x_1, y_1)$$

if either $y_0 - x_0^2 < y_1 - x_1^2$, or $y_0 - x_0^2 = y_1 - x_1^2$ and $x_0 < x_1$. Show that this is an order relation on the plane, and describe it geometrically.

Solution. (Comparability) Let $(x_0, y_0), (x_1, y_1) \in \mathbb{R}^2$ with $(x_0, y_0) \neq (x_1, y_1)$. If $y_0 - {x_0}^2 \neq y_1 - {x_1}^2$ holds, comparability is trivial. If, on the other hand, $y_0 - {x_0}^2 = y_1 - {x_1}^2$ holds, then we necessarily have $x_0 \neq x_1$, which gives

$x_0 < x_1$ or $x_0 > x_1$ and hence $(x_0, y_0) < (x_1, y_1)$ or $(x_0, y_0) > (x_1, y_1)$ respectively.

(Nonreflexivity) Observe the following equivalence:

$$\begin{aligned}
& {}^\exists(x, y) \in \mathbb{R}^2 \, [(x, y) < (x, y)] \\
\equiv \; & {}^\exists(x, y) \in \mathbb{R}^2 \left[\left(y - x^2 < y - x^2 \right) \vee \left(\left(y - x^2 = y - x^2 \right) \wedge (x < x) \right) \right] \\
\equiv \; & {}^\exists(x, y) \in \mathbb{R}^2 \, [F \vee (T \wedge F)] \\
\equiv \; & F
\end{aligned}$$

(Transitivity) Suppose $(x_0, y_0) < (x_1, y_1)$ and $(x_1, y_1) < (x_2, y_2)$. If $y_{i-1} - {x_{i-1}}^2 < y_i - {x_i}^2$ for some $i = 1, 2$, then $(x_0, y_0) < (x_2, y_2)$ is obvious. So, we consider the case where

$$y_0 - {x_0}^2 = y_1 - {x_1}^2 \wedge x_0 < x_1 \wedge y_1 - {x_1}^2 = y_2 - {x_2}^2 \wedge x_1 < x_2$$

holds. But it is also clear for this case that $(x_0, y_0) < (x_2, y_2)$.

This order relation makes $(x_0, y_0) < (x_1, y_1)$ if (x_1, y_1) is on a "higher" parabola than (x_0, y_0), or, if both of them happen to be on the same parabola and if (x_1, y_1) is located more right than (x_0, y_0). □

Exercise 7. Show that the restriction of an order relation is an order relation.

Solution. Analogical argument of §3 Exercise 2 works. □

Exercise 8. Check that the relation defined in Example 7 is an order relation.

Solution. Comparability is obvious, and we can show nonreflexivity exactly as we have done in Exercise 6. So, we establish transitivity. Suppose $x_0 C x_1$ and $x_1 C x_2$. If $x_{i-1}^2 < x_i^2$ for some i, then transitivity is obvious. But, transitivity is also obvious in the case $x_0^2 = x_1^2 \wedge x_0 < x_1 \wedge x_1^2 = x_2^2 \wedge x_1 < x_2$. □

Exercise 9. Check that the dictionary order is an order relation.

Solution. (Comparability) Let $(a_0, b_0), (a_1, b_1) \in A \times B$ with $(a_0, b_0) \neq (a_1, b_1)$. It follows that $a_0 \neq a_1$ or "$a_0 = a_1$ and $b_0 \neq b_1$", which implies there hold either "$a_0 <_A a_1$ or $a_0 >_A a_1$" or "$a_0 = a_1$ and ($b_0 <_B b_1$ or $b_0 >_B b_1$)." since each of A and B is equipped with order relation and hence both of them satisfy comparability. Thus, (a_0, b_0) and (a_1, b_1) are comparable.

(Nonreflexivity) Noting that A and B satisfy nonereflexivity,

$$\begin{aligned}
{}^\exists(a, b) \in A \times B \, [(a, b) < (a, b)] & \equiv {}^\exists(a, b) \in A \times B \, [a < a \vee (a = a \wedge b < b)] \\
& \equiv {}^\exists(a, b) \in A \times B \, [F \vee (T \wedge F)] \\
& \equiv F.
\end{aligned}$$

(Transitivity) Suppose $(a_0, b_0) < (a_1, b_1)$ and $(a_1, b_1) < (a_2, b_2)$. If $a_{i-1} < a_i$ for some i, then it is obvious that $(a_0, b_0) < (a_2, b_2)$. If, on the other hand, $a_0 = a_1 \wedge b_0 < b_1 \wedge a_1 = a_2 \wedge b_1 < b_2$ holds, then it follows that $a_0 = a_2 \wedge b_0 < b_2$, that is, $(a_0, b_0) < (a_2, b_2)$. $\square$

Exercise 10.

(a) Show that the map $f : (-1, 1) \to \mathbb{R}$ of Example 9 is order preserving.

(b) Show that the equation $g(y) = 2y/\left[1 + (1 + 4y^2)^{1/2}\right]$ defines a function $g : (-1, 1) \to \mathbb{R}$ that is both a left and right inverse of f.

Solution.

$\boxed{\text{(a)}}$ Note that $f(x) = -f(-x)$ and that it suffices to show the claim for f restricted on $[0, 1)$. From the fact that $y^2(1 - x^2) - x^2(1 - y^2) = y^2 - x^2$, we deduce that

$$\begin{aligned} x < y \wedge (x, y \in [0, 1)) &\equiv x^2 < y^2 \wedge (x, y \in [0, 1)) \\ &\equiv x^2(1 - y^2) < y^2(1 - x^2) \wedge (x, y \in [0, 1)) \\ &\equiv \frac{x^2}{1 - x^2} < \frac{y^2}{1 - y^2} \wedge (x, y \in [0, 1)). \end{aligned}$$

$\boxed{\text{(b)}}$ Note that if f admits left and right inverse, then §2 Exercise 5 implies f is bijective and these inverses are equal to f^{-1}.

To verify the inverse property of g is reduced to an elementary calculation. So, the proof is left to readers. $\square$

Exercise 11. Show that an element in an ordered set has at most one immediate successor and at most one immediate predecessor. Show that a subset of an ordered set has at most one smallest element and at most one largest element.

Solution. Consider the first statement; suppose $a < b$ and $a' < b$. we need to establish one of the following equivalent statements:

$$[(a, b) = \emptyset \wedge (a', b) = \emptyset \Rightarrow a = a'] \equiv [a \neq a' \Rightarrow (a, b) \neq \emptyset \vee (a', b) \neq \emptyset].$$

The latter is obvious as we may assume $a < a'$ there, and in this case $a' \in (a, b)$. Thus, first statement is valid for immediate predecessors. Similar argument proves the claim for immediate successors.

We now turn to the second statement. Let a and a' be largest elements of a subset A_0 of an ordered set A. Largestness gives $a \geq x$ for all $x \in A_0$, in particular, $a \geq a'$. Exchanging the role of a and a' yields $a' \geq a$. Thus, $a = a'$. The same argument works for a smallest element. $\square$

Exercise 12. Let $\mathbb{Z}_+$ denote the set of positive integers. Consider the following order relations on $\mathbb{Z}_+ \times \mathbb{Z}_+$:

(i) The dictionary order.

(ii) $(x_0, y_0) < (x_1, y_1)$ if either $y_0 - x_0 < y_1 - x_1$, or $y_0 - x_0 = y_1 - x_1$ and $x_0 < x_1$.

(iii) $(x_0, y_0) < (x_1, y_1)$ if either $y_0 + x_0 < y_1 + x_1$, or $y_0 + x_0 = y_1 + x_1$ and $x_0 < x_1$.

In these order relations, which elements have immediate predecessors? Does the set have a smallest element? Show that all three order types are different.

Solution. In dictionary order (i), every element except that on the line $y = 1$ has a immediate predecessor, and $\mathbb{Z}_+ \times \mathbb{Z}_+$ has a smallest element $(1, 1)$.

In the order relation (ii), every element except that on the lines $x = 1$ and $y = 1$ has a immediate predecessor, and $\mathbb{Z}_+ \times \mathbb{Z}_+$ admits no smallest element.

In the order relation (iii), $\mathbb{Z}_+ \times \mathbb{Z}_+$ has a smallest element $(1, 1)$. and every element but $(1, 1)$ has a immediate predecessor.

To prove that (i)(ii)(iii) all have different order types, We need Proposition 3.3, from which it follows that sets of the same order type have the same number of smallest elements, and the same number of elements which admit (no) immediate predecessors. (We need the concept of cardinality in order to rigorously describe what is meant by "the same number" here, but we can safely omit this problem since situation is not so complicated to require a further definition.) So, we deduce from this fact that $\mathbb{Z}_+ \times \mathbb{Z}_+$ in (ii) has different order type from that in (i) and (iii) since (ii), contrary to the others, allows us to have no smallest element. Moreover, observe that $\mathbb{Z}_+ \times \mathbb{Z}_+$ in (i) has infinitely many elements which admit no immediate predecessors while there is only one such element in (iii), from which we conclude that (i) and (iii) have different order type. $\square$

Exercise 13. Prove the following:
Theorem. If an ordered set A has the least upper bound property, then it has the greatest lower bound property.

Solution. Let A be an ordered set that has the largest upper bound property, and let A_0 be a nonempty subset of A that is bounded below. Let B the set of all lower bounds for A_0, say, $B := \{a \in A \mid {}^\forall a_0 \in A_0\, [a \leq a_0]\}$. Observe that B is nonempty and bounded above by every member of A_0.

The largest upper bound property of A allows us to have a supremum s of B. It is obvious that $s \leq a_0$ for all $a_0 \in A_0$, and so $s \in B$. This means $s = \max B$. Thus, $s = \inf A_0$. $\square$

Exercise 14. If C is a relation on a set A, define a new relation D on A by letting $(b, a) \in D$ if $(a, b) \in C$.

(a) Show that C is symmetric if and only if $C = D$.

(b) Show that if C is an order relation, D is also an order relation.

(c) Prove the converse of the theorem in Exercise 13.

Solution.

$\boxed{\text{(a)}}$ if we assume that C is symmetric, then it follows that

$$(a, b) \in C \equiv (b, a) \in C \equiv (a, b) \in D,$$

and that $C = D$. Conversely, if we assume $C = D$, then we have

$$(a, b) \in C \equiv (b, a) \in D \equiv (b, a) \in C,$$

from which we conclude that C is symmetric.

$\boxed{\text{(b)}}$ Let C be an order relation. It is clear that D inherits comparability and nonereflexivity from C. This is also the case for transitivity. In fact, if we let $(x, y) \in D$ and $(y, z) \in D$, or equivalently, if $(y, x) \in C$ and $(z, y) \in C$, then, by transitivity $(z, x) \in C$, that is, $(x, z) \in D$.

$\boxed{\text{(c)}}$ Let $(A, <_C)$ and $(A, <_D)$ be ordered set equipped with order relation C and D, respectively, and let A_0 be a nonempty subset of A. (b) makes it easy to check that $(A_0, <_C)$ has an upper bound if and only if $(A_0, <_D)$ has an lower bound, and that m is a largest element of $(A_0, <_C)$ if and only if it a smallest element of $(A_0, <_D)$, and that s is a supremum of $(A_0, <_C)$ if and only if it is a infimum of $(A_0, <_D)$. As a result, $(A_0, <_C)$ has the greatest lower bound property if and only if $(A_0, <_D)$ has the least upper bound property.

If we assume that $(A_0, <_C)$ has the greatest lower bound property, or equivalently, if $(A_0, <_D)$ has the least upper bound property, then Exercise 13 implies that $(A_0, <_D)$ has the greatest lower bound property, to which it is equivalent to state that $(A_0, <_C)$ has the least upper bound property. $\square$

Exercise 15. Assume that real line has the least upper bound property.

(a) Show that the sets

$$[0, 1] = \{x \mid 0 \leq x \leq 1\},$$
$$[0, 1) = \{x \mid 0 \leq x < 1\}$$

have the least upper bound property.

(b) Does $[0,1] \times [0,1]$ in the dictionary order have the least upper bound property? What about $[0,1] \times [0,1)$? What about $[0,1) \times [0,1]$?

Solution.

$\boxed{\text{(a)}}$ We establish the greatest lower bound property instead. Let A be a nonempty subset of $[0,1)$ that is bounded below in $[0,1)$. The fact that A has 0 as its lower bound implies A has an infimum, denoted by s, in $\mathbb{R}$ since $\mathbb{R}$ is assumed to have the greatest lower bound property.

It is clear that $r \in [0,1)$, and so r is a lower bound for A in $[0,1)$. It is straightforward to check that a largest element of the set of all lower bounds for A in $\mathbb{R}$ is equal to that in $[0,1)$, which implies $\inf A = s$. Thus, $[0,1)$ has the required property. We can repeat the same argument if we replace $[0,1)$ with $[0,1]$.

$\boxed{\text{(b)}}$ $[0,1] \times [0,1)$ fails to have the property. Indeed, the set $\{0\} \times [0,1)$ is bounded above by every element whose first coordinate is greater than 0, but the set admits no supremum. $\square$

4 The Integers and Real Numbers

Remark 4.1 (Direct proof of Strong induction principle)**.** Here we provide another proof of Theorem 4.2, Strong induction principle, that includes no apagogical argument. You can safely skip here if you are not interested. This note is just an illustration of a possible way of verifying the result, and nothing given here is needed later.

Logical rule allows us to deduce that following statements are all equivalent to Theorem 4.2:

$$\begin{aligned}
& A \subset \mathbb{Z}_+ \Rightarrow \left[{}^{\forall} n \in \mathbb{Z}_+ \left[S_n \subset A \Rightarrow n \in A\right] \Rightarrow A = \mathbb{Z}_+\right] \\
\equiv\ & \left[A \subset \mathbb{Z}_+ \Rightarrow \left[A \subsetneq \mathbb{Z}_+ \Rightarrow \neg {}^{\forall} n \in \mathbb{Z}_+ \left[S_n \subset A \Rightarrow n \in A\right]\right]\right] \\
\equiv\ & \left[A \subset \mathbb{Z}_+ \Rightarrow \left[A \subsetneq \mathbb{Z}_+ \Rightarrow \neg {}^{\forall} n \in \mathbb{Z}_+ \left[\neg\left(S_n \subset A\right) \vee n \in A\right]\right]\right] \\
\equiv\ & \left[A \subset \mathbb{Z}_+ \Rightarrow \left[A \subsetneq \mathbb{Z}_+ \Rightarrow {}^{\exists} n \in \mathbb{Z}_+ \left[S_n \subset A \wedge n \notin A\right]\right]\right].
\end{aligned}$$

We, of course, prove the last one. Suppose $A \subsetneq \mathbb{Z}_+$. Then, Theorem 4.1 lets us have a smallest element n of $\mathbb{Z}_+ \setminus A$. It turns out that so derived n has the required property. In fact, smallest property implies $S_n \subset A$, and $n \in \mathbb{Z}_+ \setminus A$ does $n \notin A$. Thus, Theorem 4.2 follows.

Remark 4.2 (Direct proof of Archimedean ordering property)**.** We continue to apply an analogous argument. This time, however, the proof establishes some useful fact that is actually exploited later.(see Exrceise 8 and 9, for instance.)

It is clear that following statements are equivalent to Archimedean ordering property:

$$\begin{aligned} \neg {}^{\exists}b \in \mathbb{R} {}^{\forall}n \in \mathbb{Z}_+ \, [n \leq b] &\equiv {}^{\forall}b \in \mathbb{R} \neg {}^{\forall}n \in \mathbb{Z}_+ \, [n \leq b] \\ &\equiv {}^{\forall}b \in \mathbb{R} {}^{\exists}n \in \mathbb{Z}_+ \, [n > b] \, . \end{aligned}$$

Let A be a subset of $\mathbb{Z}_+$ such that

$${}^{\forall}b \in \mathbb{R} {}^{\exists}n \in A \, [n > b]$$

holds. It is clear that A satisfies the Hypothesis of Theorem 4.2, from which we conclude $A = \mathbb{Z}_+$. This establishes the result.

It is convenient to use Archimedean ordering property in the form we have just given the proof of, rather than the original stated in the text. This equivalent statement of Archimedean ordering property can be verbalized as: *For any real number, there exists a positive integer larger than the real number.*

Exercise 1. Prove the following "laws of algebra" for $\mathbb{R}$, using only axiom (1)–(5):

(a) If $x + y = x$, then $y = 0$.

(b) $0 \cdot x = 0$.

(c) $-0 = 0$.

(d) $-(-x) = x$.

(e) $x(-y) = -(xy) = (-x)y$.

(f) $(-1)x = -x$.

(g) $x(y - z) = xy - xz$.

(h) $-(x - y) = -x - y$; $-(x - y) = -x + y$.

(i) If $x \neq 0$ and $x \cdot y = x$ then $y = 1$.

(j) $x/x = 1$ if $x \neq 0$.

(k) $x/1 = x$.

(l) $x \neq 0$ and $y \neq 0 \Rightarrow xy \neq 0$.

(m) $(1/y)(1/z) = 1/(yz)$ if $y, z \neq 0$.

(n) $(x/y)(w/z) = (xw)/(yz)$ if $y, z \neq 0$.

(o) $(x/y) + (w/z) = (xz + wy)/(yz)$ if $y, z \neq 0$.

(p) $x \neq 0 \Rightarrow 1/x \neq 0$.

(q) $1/(w/z) = z/w$ if $w, z, \neq 0$.

(r) $(x/y)/(w/z) = (xz)/(yw)$ if $y, w, z \neq 0$.

(s) $(ax)/y = a(x/y)$ if $y \neq 0$.

(t) $(-x)/y = x/(-y) = -(x/y)$ if $y \neq 0$.

Solution.

$\boxed{\text{(a)}}$ The result is verified by

$$\begin{aligned} 0 &\overset{(4)}{=} x + (-x) = (x + y) + (-x) \overset{(1)}{=} x + (y + (-x)) \\ &\overset{(2)}{=} x + ((-x) + y) \overset{(1)}{=} (x + (-x)) + y \overset{(4)}{=} 0 + y \overset{(3)}{=} y. \end{aligned}$$

$\boxed{\text{(b)}}$ It is easy to see that

$$x \cdot x + 0 \cdot x \overset{(2)}{=} x \cdot x + x \cdot 0 \overset{(5)}{=} x \cdot (x + 0) \overset{(3)}{=} x \cdot x.$$

Thus, (a) implies the result.

$\boxed{\text{(c)}}$ (3) yields $0 + 0 = 0$. Hence, $-0 = 0$ by (4).

$\boxed{\text{(d)}}$ Observe $x + (-x) = 0$ and use (4).

$\boxed{\text{(e)}}$ It follows that

$$0 \overset{(b)}{=} x0 \overset{(4)}{=} x(y + (-y)) \overset{(5)}{=} xy + x(-y).$$

So, by (4), we have $x(-y) = -(xy)$. Similar argument shows the other equality.

$\boxed{\text{(f)}}$ We deduce that

$$0 \overset{(b)}{=} (1 + (-1))\, x \overset{(1)}{=} x + (-1)x,$$

and that $(-1)x = -x$ by (4).

$\boxed{\text{(g)}}$ Check that

$$x(y-z) = x(y+(-z)) \overset{(5)}{=} xy + x(-z) \overset{(e)}{=} xy - xz.$$

$\boxed{\text{(h)}}$ Observe

$$(x+y) - x - y = (y+x) - x - y = y + (x + (-x)) - y = y - y = 0,$$

which implies $-(x+y) = -x-y$. Similar argument proves the other equality.

$\boxed{\text{(i)}}$ Just use (4).

$\boxed{\text{(j)}}$ This is an immediate consequence from the the definition of reciprocal and quotient.

$\boxed{\text{(k)}}$ $1 \cdot 1 = 1$ implies $1/1 = 1$, from which it follows that

$$\frac{x}{1} = x \cdot \frac{1}{1} = x \cdot 1 = x.$$

$\boxed{(\ell)}$ Note that the statement is equivalent to each of the following:

$$\begin{aligned} & {}^{\forall}x \in \mathbb{R}\,{}^{\forall}y \in \mathbb{R}\,[x \neq 0 \land y \neq 0 \Rightarrow xy \neq 0] \\ \equiv\ & {}^{\forall}x \in \mathbb{R}\,{}^{\forall}y \in \mathbb{R}\,[xy = 0 \Rightarrow x = 0 \lor y = 0] \\ \equiv\ & {}^{\forall}x \in \mathbb{R}\,{}^{\forall}y \in \mathbb{R}\,[xy = 0 \Rightarrow (x \neq 0 \Rightarrow y = 0)]\,. \end{aligned}$$

Of course, we establish the last statement. If $xy = 0$ and $x \neq 0$, then we have

$$y = \left(x \cdot \frac{1}{x}\right) y = xy \cdot \frac{1}{x} = 0 \cdot \frac{1}{x} = 0.$$

$\boxed{\text{(m)}}$ Observe

$$yz \cdot \left(\frac{1}{y}\right)\left(\frac{1}{z}\right) = y\left(z \cdot \frac{1}{z}\right)\frac{1}{y} = y \cdot \frac{1}{y} = 1,$$

which implies

$$\frac{1}{yz} = \left(\frac{1}{y}\right)\left(\frac{1}{z}\right).$$

$\boxed{\text{(n)}}$ Use (m).

$\boxed{\text{(o)}}$ Letting, in (n), $w = z$ and $x = y$ yields

$$\frac{xz}{yz} = \frac{x}{y} \cdot \frac{w}{w} = \frac{x}{y}$$

and

$$\frac{wy}{yz} = \frac{w}{z}$$

respectively, from which the result follows.

$\boxed{\text{(p)}}$ Note that

$$^{\forall}x \in \mathbb{R}\left[x \neq 0 \Rightarrow \frac{1}{x} \neq 0\right] \equiv {}^{\forall}x \in \mathbb{R}\left[\frac{1}{x} = 0 \Rightarrow x = 0\right],$$

and that the latter is vacuously true.

$\boxed{\text{(q)}}$ Observe

$$\frac{w}{z} \cdot \frac{z}{w} = w \cdot \frac{1}{z} \cdot z\frac{1}{w} = w \cdot \frac{1}{w} = 1,$$

from which the result follows.

$\boxed{\text{(r)}}$ For $x \neq 0$, replacing, in (m), y with y/x and z with w/z allows us to have, along with (q),

$$\frac{x/y}{w/z} = \left(\frac{1}{y/x}\right)\left(\frac{1}{w/z}\right) = \frac{1}{\frac{yw}{xz}} = \frac{xz}{yw}.$$

This is also valid for $x = 0$.

$\boxed{\text{(s)}}$ For $a \neq 0$, setting $w = 1$ and $z = a$ in (r), lets us to obtain

$$\frac{x/y}{1/a} = \frac{ax}{y}.$$

Then, (q) implies

$$\frac{(ax)}{y} = a\left(\frac{x}{y}\right),$$

which is also valid for $a = 0$.

$\boxed{\text{(t)}}$ In (o), choosing $z = y$ and $w = -x$ gives

$$\frac{x}{y} + \frac{(-x)}{y} = \frac{xy + (-x)y}{y^2} = 0,$$

from which it follows that

$$-\left(\frac{x}{y}\right) = \frac{(-x)}{y}.$$

Similar argument establishes the other equality. $\square$

Exercise 2. Prove the following "laws of inequalities" for $\mathbb{R}$, using only axiom (1)–(6) along with the result of Exercise 1:

(a) $x > y$ and $w > z \Rightarrow x + w > y + z$.

(b) $x > 0$ and $y > 0 \Rightarrow x + y > 0$ and $x \cdot y > 0$.

(c) $x > 0 \Leftrightarrow -x < 0$.

(d) $x > y \Leftrightarrow -x < -y$.

(e) $x > y$ and $z < 0 \Rightarrow xz < yz$.

(f) $x \neq 0 \Rightarrow x^2 > 0$, where $x^2 = x \cdot x$.

(g) $-1 < 0 < 1$.

(h) $xy > 0 \Leftrightarrow x$ and y are both positive or both negative.

(i) $x > 0 \Rightarrow 1/x > 0$.

(j) $x > y > 0 \Rightarrow 1/x < 1/y$.

(k) $x < y \Rightarrow x < (x + y)/2 < y$.

Solution.

$\boxed{\text{(a)}}$ Use (6) successively to gain $x + y > y + w > y + z$.

$\boxed{\text{(b)}}$ The statement immediately follows from (a) and (6) respectively.

$\boxed{\text{(c)}}$ First, suppose $x > 0$. Letting in (6) $y = 0$ and $z = -x$ yields

$$0 = x + (-x) > -x.$$

Suppose $-x < 0$. It follows trivially from (6) that

$$0 > y \Rightarrow z > y + z.$$

Taking $y = -x$ and $z = x$ here proves the remaining implication.

$\boxed{\text{(d)}}$ Use (6),(c),(h) of the previous exercise, and (6) in order, to deduce

$$x > y \equiv x - y > 0 \equiv -(x - y) < 0 \equiv -x + y < 0 \equiv -x < -y.$$

$\boxed{\text{(e)}}$ Since $-z > 0$ by (c), it follows from (6) that

$$x(-z) > y(-z) \equiv xz < yz.$$

$\boxed{\text{(f)}}$ If $x > 0$, then (6) implies $x \cdot x > 0$. If otherwise, say, $x < 0$, then $-x > 0$. So, it follows from (2),(6), and (d) of the previous exercise that $x \cdot x > 0$.

$\boxed{\text{(g)}}$ Note that $1 \neq 0$. So, it follows from (f) that $1 = 1 \cdot 1 > 0$. Then, (c) implies $-1 < 0$.

$\boxed{\text{(h)}}$ We claim that RHS on the statement of (h) is equivalent to:

$$\begin{aligned}(x > 0 \wedge y > 0) \vee (x < 0 \wedge y < 0) &\equiv [(x \leq 0 \vee y \leq 0) \Rightarrow (x < 0 \wedge y < 0)] \\ &\equiv xy > 0\end{aligned}$$

The first equivalence is a result of logic. We need to prove the second. But, ($\Leftarrow$) part is obvious, so we establish ($\Rightarrow$) part. If $x \leq 0 \vee y \leq 0$ holds, then hypothesis yields $x < 0 \wedge y < 0$, which implies, with the help of (b),

$$xy = (-x)(-y) > 0$$

as required. On the other hand, If $x \leq 0 \vee y \leq 0$ does not hold, that is, if $x > 0 \wedge y > 0$ is the case, then (b) again gives the required reslut.

$\boxed{\text{(i)}}$ Noting $x > 0$ and

$$x \cdot \frac{1}{x} = 1 > 0,$$

apply (h) to get $1/x > 0$.

$\boxed{\text{(j)}}$ Observe that $xy > 0$ by assumption and (h), and that $1/(xy) > 0$ by (i). Then, exploit (o) of the previous exercise and (h) to gain

$$\frac{1}{y} - \frac{1}{x} = \frac{x-y}{xy} > 0.$$

$\boxed{\text{(k)}}$ Use (6) successively to obtain $2x < x + y < 2y$. Since $0 < 1$ by (g), we have that $2 > 0$ by (a), and that, by (i), $1/2 > 0$, which, combined with the first inequality, implies the result. $\square$

Exercise 3.

(a) Show that if $\mathcal{A}$ is a collection of inductive sets, then the intersection of the elements of $\mathcal{A}$ is a inductive set.

(b) Prove the basic properties of (1) and (2) of $\mathbb{Z}_+$.

Solution.

$\boxed{\text{(a)}}$ Obviously $1 \in \cap\mathcal{A}$. the proof then completes by seeing

$$x \in \cap\mathcal{A} \equiv {}^{\forall}A \in \mathcal{A}\,[x \in A] \Rightarrow {}^{\forall}A \in \mathcal{A}\,[x + 1 \in A] \equiv x + 1 \in \cap\mathcal{A}.$$

$\boxed{\text{(b)}}$ $1 \in \mathbb{Z}_+$ is a direct consequence of (a). To prove principle of induction, note that $A \subset \mathbb{Z}_+$, by definition, and that $\mathbb{Z}_+ = \cap\mathcal{A} \subset A$ since A itself is inductive. Thus, $A = \mathbb{Z}_+$. □

Exercise 4.

(a) Prove by induction that given $n \in \mathbb{Z}_+$, every nonempty subset of $\{1, \cdots, n\}$ has a largest element.

(b) Explain why you cannot conclude from (a) that every nonempty subset of $\mathbb{Z}_+$ has a largest element.

Solution.

$\boxed{\text{(a)}}$ Let A be a subset of $\mathbb{Z}_+$ such that given $n \in A$, every nonempty subset of $\{1, \cdots, n\}$ has a largest element. It is clear that $1 \in A$. Assume $x \in A$ and let $B \subset \{1, \cdots, x+1\}$. We prove $x + 1 \in A$.

If $x + 1 \notin B$, then B possesses a largest element since there hold $B \subset \{1, \cdots, x\}$ and $x \in A$. So, suppose $x+1 \in B$. Let $B_0 := B \setminus \{x + 1\}$, and let k be a largest element of B_0. Then, the existence of a largest element of B is reduced to that of $\{k, x + 1\}$. But the set admits a largest element since comparability of $\mathbb{Z}_+$ allows us to have either $k > x + 1$ or $k < x + 1$. This means $x + 1 \in A$ and so, A is inductive. Thus, $A = \mathbb{Z}_+$.

$\boxed{\text{(b)}}$ (a) does not deal with the sets like $\mathbb{Z}_+$ and $\mathbb{Z}_+ \setminus \{1, \cdots, n\}$. □

Exercise 5. Prove the following properties of $\mathbb{Z}$ and $\mathbb{Z}_+$:

(a) $a, b \in \mathbb{Z}_+ \Rightarrow a + b \in \mathbb{Z}_+$.

(b) $a, b \in \mathbb{Z}_+ \Rightarrow a \cdot b \in \mathbb{Z}_+$.

(c) Show that $a \in \mathbb{Z}_+ \Rightarrow a - 1 \in \mathbb{Z}_+ \cup \{0\}$.

(d) $c, d \in \mathbb{Z} \Rightarrow c + d \in \mathbb{Z}$ and $c - d \in \mathbb{Z}$.

(e) $c, d \in \mathbb{Z} \Rightarrow c \cdot d \in \mathbb{Z}$.

Solution.

$\boxed{\text{(a)}}$ Given $a \in \mathbb{Z}_+$, let $X := \{x \in \mathbb{R} \mid a + x \in \mathbb{Z}_+\}$. The fact that $\mathbb{Z}_+$ is inductive implies $a + 1 \in \mathbb{Z}_+$ and so $1 \in X$. Exactly the same argument shows that if $y \in X$, then $y + 1 \in X$. Hence, X is inductive, which yields

$$^{\forall}a^{\forall}b\left[a \in \mathbb{Z}_+ \wedge b \in \mathbb{R} \Rightarrow a + b \in \mathbb{Z}_+\right].$$

Now it is easy to deduce (a) from this.

$\boxed{\text{(b)}}$ Given $a \in \mathbb{Z}_+$, let $X := \{x \in \mathbb{Z}_+ \mid a \cdot x \in \mathbb{Z}_+\}$. Assume $S_n \subset X$ for every $n \in \mathbb{Z}_+$ where S_n is a section of n. Then, we have $n \in X$ since $a(n-1) \in \mathbb{Z}_+ \wedge a \in \mathbb{Z}_+$ and (a) give $an \in \mathbb{Z}_+$. Thus, X is inductive, which completes the proof.

$\boxed{\text{(c)}}$ We show the set $X := \{x \in \mathbb{Z}_+ \mid x - 1 \in \mathbb{Z}_+ \cup \{0\}\}$ is inductive. It is clear that $1 \in X$ and $\mathbb{Z}_+ \cup \{0\}$ is inductive. It follows from the latter fact that if $x \in X$, then $x + 1 \in X$. Thus, X is inductive.

$\boxed{\text{(d)}}$ It suffices to prove that

$$c \in \mathbb{Z} \wedge d \in \mathbb{Z}_+ \Rightarrow c + d \in \mathbb{Z} \wedge c - d \in \mathbb{Z}. \tag{4.1}$$

Note that (a) and (c) establishes the following special case for this:

$$c \in \mathbb{Z}_+ \wedge d = 1 \Rightarrow c + d \in \mathbb{Z}_+ \wedge c - d \in \mathbb{Z}_+ \cup \{0\},$$

which yields, by the definition of $\mathbb{Z}$,

$$c \in \mathbb{Z} \wedge d = 1 \Rightarrow c + d \in \mathbb{Z} \wedge c - d \in \mathbb{Z}.$$

This implies that if $z \in \mathbb{Z}$, then $z + 1 \in \mathbb{Z} \wedge z - 1 \in \mathbb{Z}$. Hence, (4.1) follows from the induction on d.

$\boxed{\text{(e)}}$ We already know (b), one special case for (e), and the fact $0 \cdot 0 = 0 \in \mathbb{Z}$. With this in mind, we start with considering another special case:

$$c, d \in \mathbb{Z} \setminus \{\mathbb{Z}_+ \cup \{0\}\} \Rightarrow c \cdot d \in \mathbb{Z}.$$

The proof is easy as it is reduced to the case where (b) works. So, it remains to show that

$$c \in \mathbb{Z}_+ \wedge d \in \mathbb{Z} \setminus \{\mathbb{Z}_+ \cup \{0\}\} \Rightarrow c \cdot d \in \mathbb{Z}. \tag{4.2}$$

We prove this by induction on c. (4.2) is obviously valid for $c = 1$. Now assume (4.2) holds for $c = n$. We have $d, dn \in \mathbb{Z}$ and then (d) gives $d(n+1) = d + dn \in \mathbb{Z}$. This means (4.2) holds for $c = n + 1$, from which we conclude (4.2). $\square$

Exercise 6. Let $a \in \mathbb{R}$. Define inductively

$$\begin{aligned} a^1 &= a, \\ a^{n+1} &= a^n \cdot a \end{aligned}$$

for $n \in \mathbb{Z}_+$. Show that for every $n, m \in \mathbb{Z}_+$ and $a, b \in \mathbb{Z}_+$,

$$\begin{aligned} a^n a^m &= a^{n+m}, \\ (a^n)^m &= a^{nm}, \\ a^m b^m &= (ab)^m. \end{aligned}$$

These are called the ***laws of exponents***.

Solution. Note that every term that appears in (4.3)–(4.5) these equality is well-defined by induction. (We need principle of recursive definition to define it rigorously. See §7, §8 for the principle on the set of positive integers.)

First, we prove (4.3). Given $a \in \mathbb{R}$ and $n \in \mathbb{Z}_+$, let

$$X := \left\{x \in \mathbb{Z}_+ \mid a^n a^x = a^{n+x}\right\}.$$

It follows from the definition of a^n that X is inductive. Thus, (4.3) follows.

We proceed to verifying (4.4). For $a \in \mathbb{R}$ and $n \in \mathbb{Z}_+$, set

$$Y := \{y \in \mathbb{Z}_+ \mid (a^n)^y = a^{ny}\}.$$

It is easy to check that Y is inductive, from which we conclude (4.4).

We establish (4.5). Let $a, b \in \mathbb{R}$ and let

$$W := \{m \in \mathbb{Z}_+ \mid a^m b^m = (ab)^m\}.$$

It suffices to show that W is inductive. Observe $1 \in W$ and assume $m \in W$, from which we deduce that

$$a^{m+1}b^{m+1} = a(a^m b^m)b = (ab)^m(ab) = (ab)^{m+1},$$

that is, $m + 1 \in W$. Hence, W is inductive. □

Exercise 7. Let $a \in \mathbb{R}$ and $a \neq 0$. Define $a^0 = 1$, and for $n \in \mathbb{Z}_+$, $a^{-n} = 1/a^n$. Show that the laws of exponents hold for $a, b \neq 0$ and $n, m \in \mathbb{Z}_+$.

Solution. Let $a, b \in \mathbb{R}$ and $a \neq 0, b \neq 0$ throughout this exercise, and let us simply mention (4.3), for instance, to mean the corresponding statement we are concerned here.

Note that, for $n \in \mathbb{Z}_+$, we have

$$a^{-n} = \frac{1}{a^n} = \left(\frac{1}{a}\right)^n.$$

We first prove (4.5) for convenience. It is obvious that (4.5) holds for $m = 0$. For $-m \in \mathbb{Z}_+$, previous exercise gives

$$a^m b^m = \left(\frac{1}{a}\right)^{-m}\left(\frac{1}{b}\right)^{-m} = \left(\frac{1}{ab}\right)^{-m} = (ab)^m.$$

Thus, we complete the proof of (4.5).

Next, we show (4.4). Since Exercise 6 yields $1^\ell = 1$ for all $\ell \in \mathbb{Z}_+$, we deduce that (4.4) is valid for the case where either $n = 0$ or $m = 0$ holds. For $-n, -m \in \mathbb{Z}_+$, it follows from Exercise 6 and Exercise 1(h) that

$$(a^n)^m = \left(\frac{1}{\frac{1}{a^{-n}}}\right)^{-m} = \left(a^{-n}\right)^{-m} = a^{(-n)(-m)} = a^{nm}.$$

So, (4.4) holds for $-n, -m \in \mathbb{Z}_+$. One remaining case is where $n \in \mathbb{Z}_+$ and $-m \in \mathbb{Z}_+$. In this case, we have

$$(a^n)^m = \left(\frac{1}{a^n}\right)^{-m} = \left(\left(\frac{1}{a}\right)^n\right)^{-m} = \left(\frac{1}{a}\right)^{n(-m)} = \left(\frac{1}{a}\right)^{-(nm)} = a^{nm},$$

as required. Confirm that similar argument applied to the case $m \in \mathbb{Z}_+$, $-n \in \mathbb{Z}_+$ also leads us to the required equation. Thus, (4.4) is proved.

Lastly, We turn to (4.3). For $n = 0$ and $m = 0$, (4.3) is trivial. For $-n, -m \in \mathbb{Z}_+$, it is easy to see that

$$a^n a^m = \left(\frac{1}{a}\right)^{-n}\left(\frac{1}{a}\right)^{-m} = \left(\frac{1}{a}\right)^{-(n+m)} = a^{n+m}.$$

Thanks to the symmetry of n and m, we only have to consider the case $n \in \mathbb{Z}_+$, $-m \in \mathbb{Z}_+$. If $n \geq -m$, then

$$a^n a^m = a^{n+m}a^{-m}a^m = a^{n+m}a^0 = a^{n+m}.$$

It is now obvious how to prove if $n < -m$. □

We briefly establish two useful inequalities we often use in real analysis. We exploit these inequalities in Exercise 8 and Proposition 4.4, which is in turn used in a direct proof for existence of squared root (Exercise 10).

Lemma 4.3. *Let x and y be real numbers. We claim that there hold*

$${}^{\forall}\epsilon > 0\,[x < y + \epsilon] \equiv x \leq y,$$

and

$$x \geq 0 \Rightarrow \left[{}^{\forall}\epsilon > 0\,[x < \epsilon] \equiv x = 0\right].$$

Proof. First consider (4.3), which is equivalent to

$${}^{\exists}\epsilon > 0\,[x \geq y + \epsilon] \equiv x > y.$$

This is true since LHS obviously implies RHS, and choosing $\epsilon := (x - y)/2$ proves the converse. (4.3) follows by setting $y := 0$ at (4.3).

Exercise 8.

(a) Show that $\mathbb{R}$ has the greatest lower bound property.

(b) Show that $\inf \{1/n \mid n \in \mathbb{Z}_+\} = 0$.

(c) Show that given a with $0 < a < 1$, $\inf \{a^n \mid n \in \mathbb{Z}_+\} = 0$.

Solution.

$\boxed{\text{(a)}}$ Use (7) and §3 Exercise 14.

$\boxed{\text{(b)}}$ Let $N := \{1/n \mid n \in \mathbb{Z}_+\}$. Note that 0 is a lower bound for N, and so $\inf N \geq 0$. In light of Lemma 4.3 proof is reduced to

$$^\forall \epsilon > 0 \, [\inf N < \epsilon] \, .$$

To this end, it suffices to prove that

$$^\forall \epsilon > 0 ^\exists n \in \mathbb{Z}_+ \left[\frac{1}{n} < \epsilon \right] .$$

Fix $\epsilon > 0$. Archimedean ordering property(see Note 4.2) allows us to have $n \in \mathbb{Z}_+$ such that $1/\epsilon < n$, that is, $1/n < \epsilon$.

$\boxed{\text{(c)}}$ It is easy to deduce from binomial expansion that

$$(1 + h)^n \geq 1 + nh$$

for every $h > 0$ and $n \in \mathbb{Z}_+$. Choosing $h = (1 - a)/a$ yields

$$0 \leq a^n \leq \frac{a}{a + n - an} < \frac{1}{n}.$$

Since it is obvious that $\inf a_n \geq 0$, it remains to show $\inf a_n \leq \inf (1/n)$, that is, monotonicity property of infimum. We establish this in Proposition 4.4.

□

Proposition 4.4 (Property of supremum and infimum)**.**

(a) Let $A := \{a_n \in \mathbb{R} \mid n \in \mathbb{Z}_+\}$ and $B := \{b_n \in \mathbb{R} \mid n \in \mathbb{Z}_+\}$. Suppose that $\sup A$ and $\inf B$ and etc. exist, and that $a_n \leq b_n$ for all $n \in \mathbb{Z}_+$. Then we have

$$\sup A \leq \sup B, \quad \inf A \leq \inf B.$$

(b) *Let $c > 0$, and let X and Y be nonempty subsets of $\mathbb{R}$ for which $\sup X$ and $\inf Y$ etc. exist. Then we have*

$$\sup cX = c \cdot \sup X, \quad \inf cY = c \cdot \inf Y,$$

where $cX := \{c \cdot x \mid x \in X\}$.

Proof. We show (a). Note that, in general, there holds $\inf A \leq a_n$ for all $n \in \mathbb{Z}_+$, which combined with assumption gives $\inf A \leq b_n$ for all $n \in \mathbb{Z}_+$. This means $\inf A$ is a lower bound for B, from which we conclude that $\inf A \leq \inf B$. Similar argument verifies the other implication.

Consider (b). Observe cX is bounded above by $c \cdot \sup X$. Hence, $\sup cX \leq c \cdot \sup X$. For every $\epsilon > 0$ there exists $x \in X$ such that $\sup cX - \epsilon < cx$, or equivalently, $(\sup cX - \epsilon)/c < x$, from which it follows that $(\sup cX - \epsilon)/c < \sup X$, that is, $\sup cX < c \cdot \sup X + \epsilon$. This implies that $\sup cX \leq c \cdot \sup X$.

Exercise 9.

(a) Show that every nonempty subset of $\mathbb{Z}$ that is bounded above has a largest element.

(b) If $x \notin \mathbb{Z}$, show there is exactly one $n \in \mathbb{Z}$ such that $n < x < n + 1$.

(c) If $x - y > 1$, show there is at least one $n \in \mathbb{Z}$ such that $y < n < x$.

(d) If $y < x$, show there is a rational number z such that $y < z < x$.

Solution.

$\boxed{\text{(a)}}$ Let A be a nonempty subset of $\mathbb{Z}$ that is bounded above by $N \in \mathbb{Z}_+$. Assume first that $A \cap \mathbb{Z}_+ \neq \emptyset$. Exercise 4(a) allows us to choose a largest element m of $\{1, \cdots, N\} \cap A$. It is obvious that m is a largest element of A.

For general A, consider an order isomorphism $f : \mathbb{Z} \ni n \to n + K \in \mathbb{Z}$ for sufficiently large $K \in \mathbb{Z}_+$, so that $f(A) \cap \mathbb{Z}_+ \neq \emptyset$. Then the preceding argument guarantees the existence of a largest element of $f(A)$, that is, that of A.

$\boxed{\text{(b)}}$ Let $x \notin \mathbb{Z}$. The set $B := \{b \in \mathbb{Z} \mid b < x\}$ is nonempty and bounded above. (a) implies that B possesses a largest element n of B, for which we have $n < x \leq n + 1$. The required inequality follows from the fact $x \notin \mathbb{Z}$. Uniqueness of such n follows from the property of n as a largest element.

$\boxed{\text{(c)}}$ Let $x - y > 1$. If $x \in \mathbb{Z}$ is the case, then $n := x - 1$ does the job. So, assume $x \notin \mathbb{Z}$. In that case, by (b), there exists n such that $n < x < n + 1$. Since $y + 1 < x < n + 1$, we deduce that $y < n$. Thus, $y < n < x$.

$\boxed{\text{(d)}}$ In light of (c), we may assume that $0 < x - y \leq 1$. Archimedean ordering property(see Note 4.2) lets us pick $m \in \mathbb{Z}_+$ such that $1/(x-y) < m$, for which we have $mx - my > 1$. Then, (c) gives $n \in \mathbb{Z}$ such that $my < n < mx$, that is, $y < n/m < x$. This completes the proof. $\square$

Exercise 10.

Show that every positive number a has exactly one positive square root, as follows:

(a) Show that if $x > 0$ and $0 \leq h < 1$, then

$$\begin{aligned} (x+h)^2 &\leq x^2 + h(2x+1), \\ (x-h)^2 &\leq x^2 - h(2x). \end{aligned}$$

(b) Let $x > 0$. Show that if $x^2 < a$, then $(x+h)^2 < a$ for some $h > 0$; and if if $x^2 > a$, then $(x-h)^2 > a$ for some $h > 0$.

(c) Given $a > 0$, let B be the set of all real numbers x such that $x^2 < a$. Show that B is bounded above and contains at least one positive number. Let $b = \sup B$; show that $b^2 = a$.

(d) Show that if b and c are positive and $b^2 = c^2$, then $b = c$.

Solution.

$\boxed{\text{(a)}}$ Consider binomial expansion as we did in exercise 8(c). note that both inequalities are valid for $h = 1$.

$\boxed{\text{(b)}}$ Let $x > 0$ and $x^2 < a$. If $(x+1)^2 < a$, then $h := 1$ qualifies. If, on the other hand, $(x+1)^2 \geq a$, then we have

$$0 < \frac{a - x^2}{2x + 1} \leq 1.$$

In this case, exercise 9(d) allows us to have a rational number r such that

$$0 < r < \frac{a - x^2}{2x + 1},$$

which leads us to $(x + r)^2 < a$.

To show the other implication, check that

$$\frac{x^2 - a}{2x} < 1$$

holds if $(x-1)^2 \leq a$, and choose a rational number q with

$$\frac{x^2-a}{2x} < q < 1.$$

$\boxed{\text{(c)}}$ Let $a > 0$. Observe $B := \{x \in \mathbb{R} \mid x^2 < a\}$. Since $0^2 = 0 < a$, it follows from (b) that $h^2 < a$ for some $h > 0$, which gives $h \in B$, and $B_0 := B \cap \mathbb{R}_+ \neq \emptyset$.

We claim that B_0 (and so B) is bounded above. Let $x \in B_0$. If $a \leq 1$, then $x^2 < 1$ and so $x < 1$. Next, if $a > 1$, then we see that for every $x > 1$ we have $x < x^2 < a$. Thus, B_0 is bounded above. Note that we have also shown that $b > 0$.

For every $h > 0$, we have $b + h \notin B_0$ and so $(b+h)^2 \geq a$, which yields, by (b), $b^2 \geq a$.

Suppose, to the contrary, that $b^2 > a$ were the case, then (b) lets us to pick $h > 0$ such that $(b+h)^2 < a$. But this contradicts the fact that b is a supremum of B. Thus, $b^2 = a$. (We have not used here the second statement of (b).)

We have shown $b^2 \leq a$ resorting to proof by contradiction since it is straightforward and depends on less information than direct proof.

We provide two notes that proves the fact $b^2 \leq a$ by exploiting (b) or property of supremum. In the first approach, we have to look a bit more into (b) in order to gain stronger result of it. See Note 4.5 and Note 4.7.

$\boxed{\text{(d)}}$ Hypothesis is equivalent to $(b-c)(b+c) = 0$. If b and c are positive, then $b - c = 0$. □

Remark 4.5 (First direct proof for existence of square root)**.** As predicted, we begin with extracting more information from exercise 10(b) in order to construct short direct proof for exercise10(c), existence of square root.

We provide a proposition, whose contraposition is used for our main proof:

Proposition 4.6. *Let $x > 0$ and $\epsilon > 0$. If $x^2 < a$, then $(x+h)^2 < a$ for some $h \in (0, \epsilon)$; and if $x^2 > a$, then $(x-h)^2 > a$ for some $h \in (0, \epsilon)$.*

Proof. As a preparation, first observe the following very intuitive fact:
The mapping $f : \mathbb{R}_+ \cup \{0\} \ni x \mapsto x^2 \in \mathbb{R}_+ \cup \{0\}$ preserves order.
Note that this is a direct consequence of exercise 2(h).

Suppose $x^2 < a$. Exercise 10(b) gives some $d > 0$ such that $(x+d)^2 < a$. Let $\epsilon_0 := \min\{\epsilon, d\}$. It follows from the fact the mapping f defined above preserves the order that for every $h \in (0, \epsilon_0)$ we have $f(x+h) < f(x+d)$, that is, $(x+h)^2 < (x+d)^2 < a$. Similar argument works for verifying the other implication.

Note that the proof says more than Proposition 4.6 states, but for our proof for exercise 10(c), we do not need further result. More general result can be found in theory of continuous function.

Now we provide a predicted direct proof for existence of squared root. It suffices to show that $b^2 \leq a$. Let $\epsilon := \min\{b, 1\}$. For every $h \in (0, \epsilon)$, there exists $x \in B_0$ such that $0 < b - h < x$, which yields $(b-h)^2 < x^2 < a$. Then Proposition 4.6 implies $b^2 \leq a$.

Remark 4.7 (Second direct proof for existence of square root)**.** In the second approach, we do not need such finer information as Proposition 4.6. We instead need an almost obvious property of supremum, which we have established in Proposition 4.4 and an obvious result of exercise 10(b) stated as follows:

Let $x > 0$. If $x^2 > a$, then $(x-h)^2 > a$ for some $0 < h \leq 1$.

We establish $b^2 = a$. Suppose $a > 1$. We already know that B_0 is bounded above by a, which gives $b \leq a$. This implies that if $b \leq 1$, then, $b^2 \leq b \leq a$. If $b > 1$, then for every $0 < h \leq 1$, there exists $x \in B_0$ such that $0 < b - h < x$, from which it follows that $0 < (b-h)^2 < a$. By the italic statement, this yields $b^2 \leq a$. Thus $b^2 = a$ for this case.

Next, Suppose $a \leq 1$. In this case, Archimedean ordering property(see Note 4.2) allows us to choose $k \in \mathbb{Z}_+$ such that $ak^2 > 1$. Define the sets $B_k := \{x \in \mathbb{R} \mid x^2 < ak^2\}$ and $kB := \{kx \mid x \in B\}$. For a fixed $k \in \mathbb{Z}_+$, we see $B_k = kB$, from which we deduce that

$$k^2 b^2 = (k \cdot \sup B)^2 = (\sup kB)^2 = (\sup B_k)^2 = ak^2.$$

Thus, $b^2 = a$. $\square$

Exercise 11. Given $m \in \mathbb{Z}$, we say that m is ***even*** if $m/2 \in \mathbb{Z}$, and m is ***odd*** if otherwise.

(a) Show that if m is odd, $m = 2n + 1$ for some $n \in \mathbb{Z}$.

(b) Show that if p and q are odd, so are $p \cdot q$ and p^n, for any $n \in \mathbb{Z}_+$.

(c) Show that if $a > 0$ is rational, then $a = m/n$ for some $m, n \in \mathbb{Z}_+$ where not both m and n are even.

(d) *Theorem.* $\sqrt{2}$ *is irrational.*

Solution.

$\boxed{\text{(a)}}$ If $m/2 \notin \mathbb{Z}$, then Exercise 9(d) allows us to take n such that $n < m/2 < n + 1$, that is, $2n < m < 2n + 2$. Thus, $m = 2n + 1$.

$\boxed{\text{(b)}}$ Note that

$$\frac{m}{2} \in \mathbb{Z} \equiv {}^{\exists}n \in \mathbb{Z}\,[m = 2n]$$

and that this equivalence establishes the converse of (a). Let p and q are odd. Choose $k, \ell \in \mathbb{Z}$ so that $p = 2k + 1$, $q = 2\ell + 1$. Then, we have $p \cdot q = 2(2k\ell + k + \ell) + 1$. Since $2k\ell + k + \ell \in \mathbb{Z}$, we conclude that $p \cdot q$ is odd.

Let A be the subset of $\mathbb{Z}_+$ consisting of all n for which p^n is odd for every. It is obvious that $1 \in A$ and that if $n \in A$, then $p^{n+1} = p^n \cdot p$ is odd by what we have just verified. This means A is inductive, which completes the proof.

$\boxed{\text{(c)}}$ Let $a > 0$ be a rational. By definition, $a = \ell/k$ for some $k, \ell \in \mathbb{Z}$ with $k \neq 0$. If both k and ℓ are even, then there exists $k', \ell' \in \mathbb{Z}$ with $k' < k$, $\ell' < \ell$ and $a = \ell'/k'$.

Let n be a smallest element of the set $A := \left\{x \in \mathbb{Z}_+ \mid {}^{\exists}m \in \mathbb{Z}_+ \left[a = \frac{n}{m}\right]\right\}$, and let $a = n/m$. Because of the smallest property, we cannot have a smaller n' or m' than n or m respectively. Thus, both m and n are not even.

$\boxed{\text{(d)}}$ Observe the following equivalence:

$$\begin{aligned}
\sqrt{2} \in \mathbb{Q} \quad &\equiv \quad {}^{\exists}n \in \mathbb{Z}_+{}^{\exists}m \in \mathbb{Z}_+ \left[2nn = mm \wedge \neg\left(\frac{n}{2} \in \mathbb{Z}_+ \wedge \frac{m}{2} \in \mathbb{Z}_+\right)\right] \\
&\equiv \quad {}^{\exists}n \in \mathbb{Z}_+{}^{\exists}m \in \mathbb{Z}_+ \left[2nn = mm \wedge \frac{n}{2} \notin \mathbb{Z}_+ \wedge \frac{m}{2} \in \mathbb{Z}_+\right]
\end{aligned}$$

For second equivalence, note that, m needs to be even since $2nn$ are even, from which it follows that n are odd. Each of three statements turns out to be false since, in (4.7), we have $2nn/4 \notin \mathbb{Z}_+$ while $mm/4 \in \mathbb{Z}_+$ and hence $2nn \neq mm$. Thus, the statement $\sqrt{2} \in \mathbb{Q}$ is false.

Put in a language of number theory, we can simply argue in (4.6) that $2nn$ has odd number of prime factors whereas mm has even, which implies $2nn \neq mm$. □

5 Cartesian Product

Note that §2 Exercise 5 serves as an useful criterion to show that a given function is injective or surjective, or bijective. We exploit it throughout this section, and rarely prove a function is, for instance, bijective by chasing definition.

Exercise 1. Show there is a bijective correspondence of $A \times B$ and $B \times A$.

Solution. Observe each of the following functions

$$\begin{aligned}
f &: A \times B \ni (a, b) \mapsto (b, a) \in B \times A \\
g &: B \times A \ni (b, a) \mapsto (a, b) \in A \times B
\end{aligned}$$

is the inverse of the other, and hence is bijective. $\square$

Exercise 2.

(a) Show that if $n > 1$ there is bijective correspondence of

$$A_1 \times \cdots \times A_n$$

with

$$(A_1 \times \cdots \times A_{n-1}) \times A_n$$

(b) Given the indexed family $\{A_1, A_2 \cdots\}$, let $B_i = A_{2i-1} \times A_{2i}$ for each positive integer i. Show there is bijective correspondence of $A_1 \times A_2 \times \cdots$ with $B_1 \times B_2 \times \cdots$.

Solution.

$\boxed{\text{(a)}}$ Check that each of the following functions is the inverse of the other:

$$f : \prod_{i=1}^{n} A_i \ni (a_1, \cdots, a_n) \mapsto ((a_1, \cdots, a_{n-1}), a_n) \in \left(\prod_{i=1}^{n-1} A_i \right) \times A_n$$

$$g : \left(\prod_{i=1}^{n-1} A_i \right) \times A_n \ni ((a_1, \cdots, a_{n-1}), a_n) \mapsto (a_1, \cdots, a_n) \in \prod_{i=1}^{n} A_i.$$

$\boxed{\text{(b)}}$ Apply the same argument to:

$$f : \prod_{i \in \mathbb{Z}_+} A_i \ni (a_i)_{i \in \mathbb{Z}_+} \mapsto ((a_{2i-1}, a_{2i}))_{i \in \mathbb{Z}_+} \in \prod_{i \in \mathbb{Z}_+} B_i$$

$$g : \prod_{i \in \mathbb{Z}_+} B_i \ni ((a_{2i-1}, a_{2i}))_{i \in \mathbb{Z}_+} \mapsto (a_i)_{i \in \mathbb{Z}_+} \in \prod_{i=1}^{n} A_i.$$

$\square$

Exercise 3. Let $A = A_1 \times A_2 \times \cdots$ and $B = B_1 \times B_2 \times \cdots$.

(a) Show that if $B_i \subset A_i$ for all i, then $B \subset A$. (Strictly speaking, if we are given a function mapping the index set $\mathbb{Z}_+$ into the union of the sets B_i, we must change its range before it can be considered as a function mapping $\mathbb{Z}_+$ into the union of the sets A_i. We shall ignore this technicality when dealing with cartesian product.)

(b) Show the converse of (a) holds if B is nonempty.

(c) Show that if A is nonempty, each A_i is nonempty. Does the converse hold?

(d) What is the relation between the set $A \cup B$ and the cartesian product of the sets $A_i \cup B_i$? What is the relation between the set $A \cap B$ and the cartesian product of the sets $A_i \cap B_i$?

Solution.

$\boxed{\text{(a)}}$ Let $(x_i)_{i \in \mathbb{Z}_+} \in B$. Hypothesis gives $x_i \in B_i \subset A_i$ for every i, which means $(x_i)_{i \in \mathbb{Z}_+} \in A$.

$\boxed{\text{(b)}}$ Fix $i_0 \in \mathbb{Z}_+$ and $b_0 \in B_{i_0}$. Hypothesis allows us to have $(b'_i)_{i \in \mathbb{Z}_+} \in B$. Then, define a new ω-tuple $(b_i)_{i \in \mathbb{Z}_+}$ by setting

$$b_i := \begin{cases} b_0 & : \ i = i_0 \\ b'_i & : \ i \in \mathbb{Z}_+ \setminus \{i_0\}, \end{cases}$$

from which it follow that $(b_i)_{i \in \mathbb{Z}_+} \in B \subset A$. Hence, we have, in particular, $b_0 \in A_{i_0}$.

$\boxed{\text{(c)}}$ If $(a_i)_{i \in \mathbb{Z}_+} \in A$, then for every i we have $a_i \in A_i$ and hence, $A_i \neq \emptyset$.

Converse does not hold in general. We cannot conclude from the assumption that existence of ω-tuple which qualifies as a element of Cartesian product A.

$\boxed{\text{(d)}}$ For intersection, confirm the following equivalence:

$$\begin{aligned} (x_i)_{i \in \mathbb{Z}_+} \in \left(\prod_{i \in \mathbb{Z}_+} A_i \right) \cap \left(\prod_{i \in \mathbb{Z}_+} B_i \right) \quad &\equiv \quad {}^{\forall} i \in \mathbb{Z}_+ \, [x_i \in A_i] \wedge {}^{\forall} i \in \mathbb{Z}_+ \, [x_i \in B_i] \\ &\equiv \quad {}^{\forall} i \in \mathbb{Z}_+ \, [x_i \in A_i \cap B_i] \\ &\equiv \quad (x_i)_{i \in \mathbb{Z}_+} \in \prod_{i \in \mathbb{Z}_+} (A_i \cap B_i) \,. \end{aligned}$$

For union, however, we can only insist, as we have seen in §1 Exercise 2(m), that

$$\left(\prod_{i \in \mathbb{Z}_+} A_i \right) \cup \left(\prod_{i \in \mathbb{Z}_+} B_i \right) \subset \prod_{i \in \mathbb{Z}_+} (A_i \cup B_i) \,.$$

Proof and counterexample can be derived by trivial modification of those found there. $\square$

Exercise 4. Let $m, n \in \mathbb{Z}_+$. Let $X \neq \emptyset$.

(a) If $m \le n$, find an injective map $f : X^m \to X^n$.

(b) Find a bijective map $g : X^m \times X^n \to X^{m+n}$.

(c) Find a bijective map $h : X^n \to X^\omega$.

(d) Find a bijective map $k : X^n \times X^\omega \to X^\omega$.

(e) Find a bijective map $\ell : X^\omega \times X^\omega \to X^\omega$.

(f) If $A \subset B$, find an injective map $f : X^A \to X^B$.

Solution.

$\boxed{\text{(a)}}$ Let $m \le n$. Fix $x_0 \in X$ and consider a function f given by

$$f : X^m \ni (x_1, \cdots, x_m) \mapsto (x_1, \cdots, x_m, \underbrace{x_0, x_0, \cdots, x_0}_{n-m}) \in X^n,$$

and note that f has a left inverse

$$f_\ell : X^n \ni (x_1, \cdots, x_n) \mapsto (x_1, \cdots, x_m) \in X^m.$$

Thus, §2 Exercise 5 implies f is injective.

Note that f is a trivial "inclusion" function from X^m into X^n while f_ℓ is a "cut-off" function of X^n to X^m.

$\boxed{\text{(b)}}$ Confirm that

$$g : X^m \times X^n \ni ((x_1, \cdots, x_m), (y_1, \cdots, y_n)) \mapsto (x_1, \cdots, x_m, y_1, \cdots, y_n) \in X^{m+n}$$

is bijective since it admits the inverse

$$g_1 : X^{m+n} \ni (x_1, \cdots, x_{n+m}) \mapsto ((x_1, \cdots, x_m), (x_{m+1}, \cdots, x_{n+m})) \in X^m \times X^n.$$

$\boxed{\text{(c)}}$ Similar argument for (a) works here. Fix $x_0 \in X$ and define a "inclusion" function

$$h : X^m \ni (x_1, \cdots, x_m) \mapsto (x_1, \cdots, x_m, x_0, x_0, \cdots) \in X^\omega.$$

h admits a left inverse, which is given as a"cut-off" function by the rule

$$h_\ell : X^\omega \ni (x_i)_{i \in \mathbb{Z}_+} \mapsto (x_1, \cdots, x_n) \in X^n.$$

So, f is injective.

$\boxed{\text{(d)}}$ Make sure that a function given by

$$k : X^n \times X^\omega \ni ((x_1, \cdots, x_n), (y_1, y_2, \cdots)) \mapsto (x_1, \cdots, x_n, y_1, y_2, \cdots) \in X^\omega$$

is bijective since it has the inverse

$$k_1 : X^\omega \ni (x_i)_{i\in\mathbb{Z}_+} \mapsto ((x_1, \cdots, x_n), (x_{n+1}, x_{n+2}, \cdots)) \in X^n \times X^\omega.$$

$\boxed{\text{(e)}}$ Do the same thing as (d) to the following:

$$\ell : X^\omega \times X^\omega \ni \big((x_i)_{i\in\mathbb{Z}_+}, (y_i)_{i\in\mathbb{Z}_+}\big) \mapsto (x_1, y_1, x_2, y_2, \cdots) \in X^\omega$$
$$\ell_1 : X^\omega \ni (x_i)_{i\in\mathbb{Z}_+} \mapsto \Big((x_{2i-1})_{i\in\mathbb{Z}_+}, (x_{2i})_{i\in\mathbb{Z}_+}\Big) \in X^\omega \times X^\omega.$$

$\boxed{\text{(f)}}$ Successive application of (e) implies a function

$$\ell_0 : (A^\omega)^n \ni \Big((a_i^{(1)})_{i\in\mathbb{Z}_+}, \cdots, (a_i^{(n)})_{i\in\mathbb{Z}_+}\Big) \mapsto \Big(a_1^{(1)}, a_1^{(2)}, \cdots, a_1^{(n)}, a_2^{(1)}, a_2^{(2)}, \cdots\Big) \in A^\omega$$

is bijective. So,

$$m := i \circ \ell_0$$

is injective, where

$$i : A^\omega \ni (a_i)_{i\in\mathbb{Z}_+} \mapsto (a_i)_{i\in\mathbb{Z}_+} \in B^\omega,$$

which is obviously injective. □

Exercise 5. Which of the following subsets of $\mathbb{R}^\omega$ can be expressed as the cartesian product of subsets of $\mathbb{R}$?

(a) $\{\boldsymbol{x} \mid x_i \text{ is an integer for all } i\}$.

(b) $\{\boldsymbol{x} \mid x_i \geq i \text{ for all } i\}$.

(c) $\{\boldsymbol{x} \mid x_i \text{ is an integer for all } i \geq 100\}$.

(d) $\{\boldsymbol{x} \mid x_2 = x_3\}$.

Solution.
$\boxed{\text{(a)}}$ $\mathbb{Z}^\omega$. $\boxed{\text{(b)}}$ $[1,\infty)\times[2,\infty)\times[3,\infty)\times\cdots$. $\boxed{\text{(c)}}$ $\underbrace{\mathbb{R}\times\cdots\times\mathbb{R}}_{99}\times\mathbb{Z}\times\mathbb{Z}\times\cdots$.

$\boxed{\text{(d)}}$ Let $X := \{x \mid x_2 = x_3\}$. X is not the cartesian product of subsets of $\mathbb{R}$. Observe $(0, 0, \cdots), (1, 1, \cdots) \in X$, but $(0, 1, 0, \cdots) \notin X$, which violates an obvious necessary condition for X to be the cartesian product. □

We end this section by introducing an extension of the proposition we have proved in §1, which can be used to work on Exercise 5(d).

Proposition 5.1 (Equivalent condition for cartesian product)**.** *Let X_i be a set for every $i \in \mathbb{Z}_+$, and let A be a subset of $\prod_{i\in\mathbb{Z}_+} X_i$. A necessary and sufficient condition for A to be the cartesian product of subsets of X_i is that there holds*

$$\left(x_i^{(1)}\right)_{i\in\mathbb{Z}_+}, \left(x_i^{(2)}\right)_{i\in\mathbb{Z}_+}, \cdots \in A \Rightarrow (z_i)_{i\in\mathbb{Z}_+} \in A,$$

where $z_i = x_i^{(j)}$ for some $j \in \mathbb{Z}_+$.

Proof. Sufficiency is proved as before. We show necessity part. Suppose that A itself is the cartesian product, and let $\left(x_i^{(1)}\right)_{i\in\mathbb{Z}_+}, \left(x_i^{(2)}\right)_{i\in\mathbb{Z}_+}, \cdots \in A$. Let I be a subset of $\mathbb{Z}_+$ such that we have $(z_i)_{i\in\mathbb{Z}_+} \in A$, where each z_i is given by

$$z_i := \begin{cases} x_i^{(j)} & : i \in I \\ x_i^{(1)} & : i \notin I \end{cases}$$

for arbitrary chosen $j \in \mathbb{Z}_+$. Since A is assumed to be a cartesian product, it is easy to see that I is inductive, and so $I = \mathbb{Z}_+$. This proves the necessity.

6 Finite Sets

Remark 6.1 (Direct proof of Corollary 6.3)**.** Corollary 6.3. is equivalent to the following statements:

$$\begin{aligned} & {}^{\exists}n \in \mathbb{Z}_+ \left[\mathrm{Bij}(A, S_{n+1}) \neq \emptyset\right] \Rightarrow \left[B \subsetneq A \Rightarrow \mathrm{Bij}(A, B) = \emptyset\right] \\ \equiv \;& \left[{}^{\exists}n \in \mathbb{Z}_+ \left[\mathrm{Bij}(A, S_{n+1}) \neq \emptyset\right] \Rightarrow \left[\mathrm{Bij}(A, B) \neq \emptyset \Rightarrow \neg\left(B \subsetneq A\right)\right]\right] \\ \equiv \;& \left[{}^{\exists}n \in \mathbb{Z}_+ \left[\mathrm{Bij}(A, S_{n+1}) \neq \emptyset\right] \Rightarrow \left[\mathrm{Bij}(B, S_{n+1}) \neq \emptyset \Rightarrow \neg\left(B \subsetneq A\right)\right]\right] \\ \equiv \;& \left[{}^{\exists}n \in \mathbb{Z}_+ \left[\mathrm{Bij}(A, S_{n+1}) \neq \emptyset\right] \Rightarrow \left[B \subsetneq A \Rightarrow \mathrm{Bij}(B, S_{n+1}) = \emptyset\right]\right]. \end{aligned}$$

The last statement is exactly what Theorem 6.2 insists. □

Remark 6.2 (Direct proof of Corollary 6.5.)**.** It suffices to show that

$$\mathrm{Bij}(A, S_n) \neq \emptyset \wedge \mathrm{Bij}(A, S_m) \neq \emptyset \Rightarrow n = m.$$

Note that we have

$$\mathrm{Bij}(A, S_n) \neq \emptyset \wedge \mathrm{Bij}(A, S_m) \neq \emptyset \equiv \mathrm{Bij}(A, S_n) \neq \emptyset \wedge \mathrm{Bij}(S_n, S_m) \neq \emptyset,$$

and that one of S_n and S_m is a subset of the other. But, $\mathrm{Bij}(S_n, S_m) \neq \emptyset$ and contrapositive of Theorem 6.2 together implies it cannot be a proper subset, which gives $S_n = S_m$. Thus, $n = m$.

(Strictly speaking, we prove $S_n = S_m$ *implies* $n = m$ as follows: $E \setminus S_n$ and $E \setminus S_m$ coincide and both sets have smallest element n and m respectively. From the uniqueness of a smallest element, we conclude that $n = m$.) $\square$

Remark 6.3 (Cardinality of the set of injective maps)**.** Let $A := \{a_1, a_2, \ldots, a_m\}$ and $B := \{b_1, b_2, \ldots, b_n\}$, and suppose $1 \leq m \leq n$. We claim that cardinality of $\mathrm{Inj}(A, B)$ is equal to $m \cdot (m-1) \cdot \cdots \cdot (m-n+1)$.

$\mathrm{Inj}(A, B)$ is nonempty since identity function is injective. Let $f \in \mathrm{Inj}(A, B)$. There are n possible choices for the value of $f(a_1)$ since $f(a_1)$ can be any of the element of $\{b_1, b_2, \ldots, b_n\}$, and $n-1$ choices available for $f(a_2)$ since it can be any of $\{b_1, b_2, \ldots, b_n\} \setminus f(a_1)$. So, in general, there are $n-(i-1)$ choices for $f(a_i)$ since it could be any of $\{b_1, b_2, \ldots, b_n\} \setminus \{f(a_1) \cup f(a_2) \cup \cdots \cup f(a_{i-1})\}$.

Now it is easy to verify the claim by induction. $\square$

Exercise 1.

(a) Make a list of all injective maps

$$f : \{1, 2, 3\} \to \{1, 2, 3, 4\}.$$

Show that none is injective.

(b) How many injective maps

$$f : \{1, \cdots, 8\} \to \{1, \cdots, 10\}$$

are there?

Solution.

$\boxed{\text{(a)}}$ For simplicity, let us ONLY in this Exercise use the nortation $h : \{a, b\} \mapsto \{b, a\}$ to mean $h(a) = b$ and $h(b) = a$.

Note 6.3 tells us that there are $4 \cdot 3 = 12$ injective functions. The list of them is given by

$$\begin{aligned}
f_1 : \{1,2,3\} &\mapsto \{1,2,3\}\\
f_2 : \{1,2,3\} &\mapsto \{1,2,4\}\\
f_3 : \{1,2,3\} &\mapsto \{1,3,2\}\\
f_4 : \{1,2,3\} &\mapsto \{1,3,4\}\\
f_5 : \{1,2,3\} &\mapsto \{2,1,3\}\\
f_6 : \{1,2,3\} &\mapsto \{2,1,4\}\\
f_7 : \{1,2,3\} &\mapsto \{2,3,1\}\\
f_8 : \{1,2,3\} &\mapsto \{2,3,4\}\\
f_9 : \{1,2,3\} &\mapsto \{3,1,2\}\\
f_{10} : \{1,2,3\} &\mapsto \{3,1,4\}\\
f_{11} : \{1,2,3\} &\mapsto \{3,2,1\}\\
f_{12} : \{1,2,3\} &\mapsto \{3,2,4\}\,.
\end{aligned}$$

It is obvious that none of them is bijective.

$\boxed{\text{(b)}}$ Note 6.3 shows there are $10\cdot9\cdot8\cdot7\cdot6\cdot5\cdot4\cdot3$ injective functions. □

Exercise 2. Show that if B is not finite and $B \subset A$, then A is not finite.

Solution. The statement is equivalent to the following:

$$\begin{aligned}
&\neg^{\exists} n \in \mathbb{Z}_+ \left[\mathrm{Bij}(B, S_n) \neq \emptyset\right] \wedge B \subset A \Rightarrow \neg^{\exists} m \in \mathbb{Z}_+ \left[\mathrm{Bij}(A, S_m) \neq \emptyset\right]\\
\equiv\ & B \subset A \Rightarrow \left[\neg^{\exists} n \in \mathbb{Z}_+ \left[\mathrm{Bij}(B, S_n) \neq \emptyset\right] \Rightarrow \neg^{\exists} m \in \mathbb{Z}_+ \left[\mathrm{Bij}(A, S_m) \neq \emptyset\right]\right]\\
\equiv\ & B \subset A \Rightarrow \left[^{\exists} m \in \mathbb{Z}_+ \left[\mathrm{Bij}(A, S_m) \neq \emptyset\right] \Rightarrow {}^{\exists} n \in \mathbb{Z}_+ \left[\mathrm{Bij}(B, S_n) \neq \emptyset\right]\right].
\end{aligned}$$

The last one is what Corollary 6.6 justifies. □

Exercise 3. Let X be the set two-element set $\{0, 1\}$. Find a bijective correspondence between X^ω and a proper subset of itself.

Solution. A function

$$f : X^\omega \ni (x_i)_{i\in\mathbb{Z}_+} \mapsto (0, x_1, x_2, \cdots) \in \{0\} \times X \times X \cdots$$

is bijective since it admits the inverse given by

$$g : \{0\} \times X \times X \cdots \ni (0, x_1, x_2, \cdots) \mapsto (x_1, x_2, \cdots) \in X^\omega.$$

□

Exercise 4.

(a) Show that A has a largest element.

(b) Show that A has the order type of a section of the positive integers.

Solution. Let A be a nonempty finite simply ordered set.

$\boxed{\text{(a)}}$ Let X be a subset of $\mathbb{Z}_+$ such that A has a largest element if there holds $\operatorname{card} A \in X$. It is clear that $1 \in X$. Assuming $n \in X$, consider the case $\operatorname{card} A = n+1$. Let $a \in A$. By assumption, there exists a largest element m of $A \setminus \{a\}$. Then, we have $\max A = \max\{a, m\}$, that is, A has a largest element, which implies $n+1 \in X$. Thus, induction establishes our claim.

$\boxed{\text{(b)}}$ Let Y be a subset of $\mathbb{Z}_+$ such that if $\operatorname{card} A \in Y$, then there exists an order isomorphism from S_n to A for some $n \in \mathbb{Z}_+$. Remember S_n is the section of n. We prove that Y is inductive. Obviously we have $1 \in Y$. Suppose $n \in Y$ and $\operatorname{card} A = n+1$. Let m be a largest element of A, and let $f : S_{n+1} \to A \setminus \{m\}$ be an order isomorphism, whose existence is guaranteed by the assumption here. Then, define a function $g : S_{n+2} \to A$ by setting

$$g(x) := \begin{cases} f(x) & : x \in S_{n+1} \\ m & : x = n+1. \end{cases}$$

It turns out that g is an order isomorphism by construction. Thus, $n+1 \in Y$. $\square$

Exercise 5. If $A \times B$ is finite, does it follow that A and B are finite?

Solution. No. Observe $\mathbb{Z}_+ \times \emptyset = \emptyset$ as a counterexample. $\square$

Exercise 6.

(a) Let $A = \{1, \cdots, n\}$. Show there is a bijection of $\mathcal{P}(A)$ with the cartesian product X^n, where X is a two-element set $\{0, 1\}$.

(b) Show that if A is finite, then $\mathcal{P}(A)$ is finite.

Solution.

$\boxed{\text{(a)}}$ Firstly, for $B \in \mathcal{P}(A)$, define a function $x^B : A \to \{0, 1\}$, called *a characteristic function of* B, by setting

$$x_i^B := \begin{cases} 1 & : i \in B \\ 0 & : i \in A \setminus B. \end{cases}$$

Then, each of the function f and g given by

$$f : \mathcal{P}(A) \ni B \mapsto (x_i^B)_{i \in A} \in X^n,$$
$$g : X^n \ni (x_i)_{i \in A} \mapsto \{i \in A \mid x_i = 1\} \in \mathcal{P}(A)$$

is easily checked to be the inverse of the other. Thus, both are bijective.

$\boxed{\text{(b)}}$ Corollary 6.8 implies that if A is finite, then so is A^n, which proves, by (a), $\mathcal{P}(A)$ is also finite.

$\square$

Exercise 7. If A and B are finite, show that the set of all function $f : A \to B$ is finite.

Solution. Let A and B be finite sets. Note that, in general, $\text{Func}(A, B)$ and $B^{\text{card}A}$ are in bijective correspondence. Thus, the claim immediately follows from Corollary 6.8.

Or you might prove the result based on the fact $\text{Func}(A, B) \subset \mathcal{P}(A \times B)$. $\square$

7 Countable and Uncountable Sets

Let us introduce new notations "$\sim$" and "$\hookrightarrow$".

Because we study cardinality through bijection, we are repeatedly in need to write "there exists a bijection from... to...". Let us introduce a notation to omit this and thus simplify our description: let X be a set; define a relation on $\mathcal{P}(X)$ by setting $A \sim B$ if A and B are in bijective correspondence. It is easy to check that this is well-defined and is an equivalence relation. This observation invites us to the idea of classifying sets based on their cardinality. This idea is developed as we proceed; see §7 Exercise 6, §9 Exercise 7, §10 Exercise 11. Note that $A \sim B$ if and only if $\text{Bij}(A, B) \neq \emptyset$.

Although the concept of cardinality is defined in terms of of bijection, it is quite often hard to verify the existence of a desired bijective function between sets of concern. That is why we have investigated some alternative conditions that allow us to say something about cardinality without explicit mention of bijection. One of them is the one given in terms of the existence of injection, such as §6 Corollary 6.7, §Theorem 7.1. However, it turns out that, as Schroeder-Bernstein theorem (Exercise 6) implies, injection is indeed a very good substituting tool to explore with. (Moreover, several concepts related to cardinality are formulated in terms of injection rather than bijection; see §9 Exercise 7, §10 Exercise). So, it is natural to have a notation that expresses the existence of injection; we write $A \hookrightarrow B$ if there exists an injection from A to B, that is, if $\text{Inj}(A, B) \neq \emptyset$.

Proposition 7.1 (Cardinality of the set of functions).
Let A, B, C, etc. be sets. We claim that we have

(a) $\mathcal{P}(A) \sim \text{Func}(A, \{0, 1\})$.

(b) $\mathrm{Func}(A \times B, C) \sim \mathrm{Func}(A, \mathrm{Func}(B, C))$.

(c) $A \sim A' \wedge B \sim B' \Rightarrow \mathrm{Func}(A, B) \sim \mathrm{Func}(A', B')$.

(d) $A \sim A' \Rightarrow \mathcal{P}(A) \sim \mathcal{P}(A')$.

Proof. Consider (a). Let τ be a function given by

$$\tau : \mathcal{P}(A) \ni B \mapsto \in 1_B \in \mathrm{Func}(A, \{0, 1\}),$$

where 1_B is the characteristic function of B given by

$$1_B(x) := \begin{cases} 1 & : x \in B \\ 0 & : x \in A \setminus B. \end{cases}$$

τ is bijective since it admits the inverse τ' defined as follows:

$$\tau' : \mathrm{Func}(A, \{0, 1\}) \ni \chi \mapsto \{x \in A \mid \chi(x) = 1\} \in \mathcal{P}(A).$$

We show (b). Consider a function ϕ given by

$$\phi : \mathrm{Func}(A \times B, C) \times A \ni (f, a) \mapsto f(a, \cdot) \in \mathrm{Func}(B, C),$$

and define a function Φ by setting

$$\Phi : \mathrm{Func}(A \times B, C) \ni f \mapsto \phi(f, \cdot) \in \mathrm{Func}(A, \mathrm{Func}(B, C)).$$

It is straightforward to check that Φ is bijective.

(c) is easy to prove; let $\varphi \in \mathrm{Bij}(A, A')$ and $\psi \in \mathrm{Bij}(B, B')$. A function given by

$$\mathrm{Func}(A, B) \ni g \mapsto \psi \circ g \circ \varphi^{-1} \in \mathrm{Func}(A', B')$$

is bijective since it admits the inverse

$$\mathrm{Func}(A', B') \ni h \mapsto \psi^{-1} \circ h \circ \varphi \in \mathrm{Func}(A, B).$$

(d) follows as $\mathcal{P}(A) \sim \mathrm{Func}(A, \{0, 1\}) \sim \mathrm{Func}(A', \{0, 1\}) \sim \mathcal{P}(A')$.

Exercise 1. Show that $\mathbb{Q}$ is countably infinite.

Solution. Note $\mathbb{Q} = \mathbb{Q}_- \cup \{0\} \cup \mathbb{Q}_+$ by definition, where $\mathbb{Q}_- := \{q \mid -q \in \mathbb{Q}_+\}$ which is countable. Apply Theorem 7.5. □

Exercise 2. Show that the maps f and g of Example 1 and 2 are bijections.

Solution. We list these functions here for convenience as follows: The function $f : \mathbb{Z} \to \mathbb{Z}_+$ in Example 1 is defined by

$$f(n) = \begin{cases} 2n & : n > 0 \\ -2n+1 & : n \le 0. \end{cases}$$

Let $A := \{(x, y) \in \mathbb{Z}_+ \times \mathbb{Z}_+ \mid y \le x\}$. The functions $f : \mathbb{Z}_+ \times \mathbb{Z}_+ \to A$ and $g : A \to \mathbb{Z}_+$ in Example 2 are defined by

$$f(x, y) = (x + y - 1, y),$$

and

$$g(x, y) = \frac{1}{2}(x - 1)x + y.$$

We first consider f in Example 1. Define a function $h : \mathbb{Z}_+ \to \mathbb{Z}$ by setting

$$h(n) := \begin{cases} m & : {}^{\exists}m \in \mathbb{Z}_+ \, [n = 2m] \\ -m & : {}^{\exists}m \in \mathbb{Z}_+ \cup \{0\} \, [n = 2m+1] \, . \end{cases}$$

It is straightforward to check that h is well-defined and is the inverse of f.

We proceed to Example 2. The function

$$\varphi : A \ni (x, y) \mapsto (x - y + 1, y) \in \mathbb{Z}_+ \times \mathbb{Z}_+$$

is the inverse of f. So, both f and φ are bijective.

Next, we show how g is constructed. Let $s : \mathbb{Z}_+ \cup \{0\} \to \mathbb{Z}_+ \cup \{0\}$ be a function such that $s(0) := 0$ and $s(x) := \sum_{k=0}^{x} k$ for $x \ge 1$. Observe that there holds, for $x \ge 1$,

$$s(x - 1) = \frac{1}{2}x(x - 1),$$

and that, given $n \in \mathbb{Z}_+$, the set

$$X := \{x \in \mathbb{Z}_+ \mid n \le s(x)\}$$

is a nonempty subset of $\mathbb{Z}_+$. So, we can take the smallest element x^* of X with the following property:

$$s(x^* - 1) < n \le s(x^*).$$

Since we have $s(x) - s(x - 1) = x$ in general, there exists $y \in \mathbb{Z}_+$ with $1 \le y \le x$ such that

$$n = s(x^* - 1) + y.$$

Uniqueness of x^* and y is obvious by construction. Hence, we have shown that

$$^{\forall}n \in \mathbb{Z}_+{}^{\exists 1}(x, y) \in A \left[n = \frac{1}{2}x(x-1) + y \right],$$

which implies g is bijective. □

Exercise 3. Let X be the two-element set $\{0, 1\}$. Show there is a bijective correspondence between the set $\mathcal{P}(\mathbb{Z}_+)$ and the cartesian product X^ω.

Solution. Noting the fact $X^\omega = \mathrm{Func}(\mathbb{Z}_+, X)$, it suffices to show that, for a set A, we have $\mathcal{P}(A) \sim \mathrm{Func}(A, \{0, 1\})$. But, we have already proved this. See Proposition 7.1. □

Exercise 4.

(a) A real number x is said to be ***algebraic***(over the rationals) if it satisfies some polynomial equation of positive degree

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0$$

with rational coefficients a_i. Assuming that each polynomial equation has only finitely many roots, show that the set of algebraic numbers is countable.

(b) A real number is said to be ***transcendental*** if it is not algebraic. Assuming the reals are uncountable, show that the transcendental numbers are uncountable.

Solution.

$\boxed{\text{(a)}}$ Without loss of generality, we may assume each polynomial equation has integer coefficients. Let E represent the set of polynomial equations over integers, and, given $f \in E$, let $\deg(f)$ be the degree of f. Let E_n be the set consisting of elements of E whose degree is n. We define a function $H : E \to \mathbb{Z}_+$ by setting, for $E \ni f(x) = a_0 + a_1 x + \cdots + a_n x^n$,

$$H(f) := n + \sum_{k=0}^{\deg(f)} |a_k|.$$

Observe that $H(f) \geq 2$ for all $f \in E$, and that the set defined as $E_n(h) := \{f \in E_n \mid H(f) = h\}$ is finite for every $n, h \in \mathbb{Z}_+$, and that $E_n = \bigcup_{h \in \mathbb{Z}_+} E_n(h)$ for every $n \in \mathbb{Z}_+$, and that $E = \bigcup_{n \in \mathbb{Z}_+} E_n$.

Let $g : E \to \mathcal{P}(\mathbb{R})$ be a function that maps every $f \in E$ to the set consisting of its roots. Each $g(E_n(h))$ is finite since each $E_n(h)$ is finite and each of $f \in E_n(h)$ has at most finitely many roots. Thus, the set of algebraic numbers

$$g(E) = g\left(\bigcup_{n,h\in\mathbb{Z}_+} E_n(h)\right) = \bigcup_{n,h\in\mathbb{Z}_+} g(E_n(h))$$

is countable.

$\boxed{\text{(b)}}$ We prove that the set of transcendental number has the same cardinality as that of real numbers. Let A be the set of algebraic numbers. In order to establish the existence of transcendental numbers, we claim $\mathbb{R} \setminus A \neq \emptyset$. In fact, we have in general $\mathbb{R} \setminus A = \emptyset \Leftrightarrow \mathbb{R} \subset A$, and RHS turns out to be false since a countable set cannot include an uncountable set.

Let θ be a transcendental number. It is obvious that $n\theta$ is also transcendental number for all $n \in \mathbb{Z}_+$. Consider a partition of $\mathbb{R}$ given by

$$\mathbb{R} = A \cup T_1 \cup T_2,$$

where $T_1 := \{n\theta \mid n \in \mathbb{Z}_+\}$ and $T_2 := \mathbb{R} \setminus (A \cup T_1)$, both combined constitutes the set of transcendental numbers. It is easy to see that A and T_1 are countable, and hence $A \cup T_1 \sim T_1$, which implies $\mathbb{R} = (A \cup T_1) \cup T_2 \sim T_1 \cup T_2$. Thus, the proof is completed. $\square$

Exercise 5. Determine, for each of the following sets, whether or not it is countable. Justify your answers.

(a) The set A of all functions $f : \{0, 1\} \to \mathbb{Z}_+$.

(b) The set B_n of all functions $f : \{1, \cdots, n\} \to \mathbb{Z}_+$.

(c) The set $C = \bigcup_{n\in\mathbb{Z}_+} B_n$.

(d) The set D of all functions $f : \mathbb{Z}_+ \to \mathbb{Z}_+$.

(e) The set E of all functions $f : \mathbb{Z}_+ \to \{0, 1\}$.

(f) The set F of all functions $f : \mathbb{Z}_+ \to \{0, 1\}$. that are "eventually zero." [We say that f is ***eventually zero*** if there is a positive integer N such that $f(n) = 0$ for all $n \geq N$.]

(g) The set G of all functions $f : \mathbb{Z}_+ \to \mathbb{Z}_+$. that are eventually 1.

(h) The set H of all functions $f : \mathbb{Z}_+ \to \mathbb{Z}_+$. that are eventually constant.

(i) The set I of all two-element subsets of $\mathbb{Z}_+$.

(j) The set J of all finite subsets of $\mathbb{Z}_+$.

Solution. Remember the concept of cartesian product is defined in terms of functions.

$\boxed{\text{(a)}}$ $A = \mathrm{Func}(\{0,1\}, \mathbb{Z}_+) \sim \mathbb{Z}_+ \times \mathbb{Z}_+$ is countable by Theorem 7.6.

$\boxed{\text{(b)}}$ $B_n = \mathrm{Func}(\{1,2,\cdots,n\}, \mathbb{Z}_+) \sim \mathbb{Z}_+^n$ is countable by Theorem 7.6.

$\boxed{\text{(c)}}$ Theorem 7.5 yields that C is countable.

$\boxed{\text{(d)}}$ Observe that we have $D = \mathrm{Func}(\mathbb{Z}_+, \mathbb{Z}_+) \sim \mathbb{Z}_+^\omega$, and $\{0,1\}^\omega \sim \{1,2\}^\omega \subset \mathbb{Z}_+^\omega$. Then, Theorem 7.7 and contrapositive of Theorem 7.3 implies $\mathbb{Z}_+^\omega$ is uncountable.

$\boxed{\text{(e)}}$ $E = \mathrm{Func}(\mathbb{Z}_+, \{0,1\}) \sim \mathcal{P}(\mathbb{Z}_+)$ is uncountable by Theorem 7.8.

$\boxed{\text{(f)}}$ We provide a bit generalized proof for (f) here in order to illustrate the idea that works for (g) and (h).

Consider a function $N : F \to \mathbb{Z}_+$ that assigns, to every $f \in F$, the smallest element $N(f)$ of the set

$$\left\{n \in \mathbb{Z}_+ \mid {}^\forall m \in \mathbb{Z}_+ {}^\forall k \in \mathbb{Z}_+ \left[m \geq n \wedge k \geq n \Rightarrow f(m) = f(k)\right]\right\},$$

which the definition of F guarantees is nonempty. In other words, $N(f)$ designates exactly when f starts to be a constant. Setting $F_n := \{f \in F \mid N(f) = n\}$, we see that $F = \bigcup_{n \in \mathbb{Z}_+} F_n$, and that each F_n is countable since a function φ defined by

$$\varphi : F_n \ni f \mapsto f|_{S_{n+1}} \in \mathrm{Func}(S_{n+1}, \mathbb{Z}_+)$$

is injective, where $f|_{S_{n+1}}$ represents the restriction of f onto S_{n+1}. Thus, F is countable.

$\boxed{\text{(g)}}\boxed{\text{(h)}}$ G and H can be proved to be countable by trivial modifications of (f).

$\boxed{\text{(i)}}$ Let $I^* := \{(x_i)_{i=0,1} \in \mathbb{Z}_+ \times \mathbb{Z}_+ \mid x_0 < x_1\}$. It is obvious that $I \sim I^*$. We then deduce from the fact $I^* \hookrightarrow \mathbb{Z}_+ \times \mathbb{Z}_+$ that I is countable.

$\boxed{\text{(j)}}$ Let $J_n := \{N \in J \mid \mathrm{card} N = n\}$, and let

$$J_n^* := \left\{(x_i)_{i \in S_{n+1}} \in \mathbb{Z}_+^n \mid {}^\forall i \in S_n \left[x_i < x_{i+1}\right]\right\}.$$

As before, we deduce from the fact $J_n^* \hookrightarrow \mathrm{Func}(S_{n+1}, \mathbb{Z}_+) \sim \mathbb{Z}_+^n$ that J_n^* is countable, and so is J_n since $J_n^* \sim J_n$. Thus, $J = \bigcup_{n \in \mathbb{Z}_+} J_n$ is countable. □

Exercise 6. We say that two sets A and B ***have the same cardinality*** if there is a bijection of A with B.

(a) Show that if $B \subset A$ and if there is an injection

$$f : A \to B,$$

then A and B have the same cardinality.

(b) Theorem (Schroeder-Bernstein theorem). If there are injections $f : A \to C$ and $g : C \to A$, then A and C have the same cardinality.

Solution.

$\boxed{\text{(a)}}$ Let $B \subset A$ and $f \in \mathrm{Inj}(A, B)$. Define the sets A_n and B_n recursively by the formula

$$\begin{aligned} A_1 &:= A, \\ B_1 &:= B, \\ A_n &:= f(A_{n-1}), \\ B_n &:= f(B_{n-1}), \end{aligned}$$

for $n > 1$. We have $A_1 \supset B_1 \supset A_2 \supset B_2 \supset A_3 \supset \cdots$ by definition. Consider a function $h : A \to B$ such that $h(x) := f(x)$ if $x \in A_n \setminus B_n$ for some n, and $h(x) := x$ if otherwise.

Then, we claim h is bijective. We first show the injectivity. Let $x, y \in A$ with $x \neq y$. We may assume $x \in A_n \setminus B_n$ for some n and $y \notin A_n \setminus B_n$ for all $n \in \mathbb{Z}_+$ since otherwise injectivity trivially follows. But, in this case, $f \in \mathrm{Inj}(A, B)$ implies $h(x) = f(x) \in f(A_n) \setminus f(B_n) = A_{n+1} \setminus B_{n+1}$, and hence,

$$h(y) = y \neq h(x),$$

establishing $h \in \mathrm{Inj}(A, B)$.

We prove next that h is surjective. Let $b \in B(= B_1)$. Seeing the definition of h, we may assume $b \in A_n \setminus B_n$ for some $n \geq 2$. Then, $f \in \mathrm{Inj}(A, B)$ implies

$$b \in f(A_{n-1}) \setminus f(B_{n-1}) = f(A_{n-1} \setminus B_{n-1}),$$

which completes the proof.

$\boxed{\text{(b)}}$ Let $C_0 := f(A)$ and $C_1 := f \circ g(C) \subset C$. It is obvious that $A \sim C_0$ and $C \hookrightarrow C_1$. Then (a) yields $c \sim C_1$. On the other hand, it follows from $g(C) \subset A$ and $f(A) \subset C$ that $f \circ g(C) \subset f(A) \subset C$, that is, $C_1 \subset C_0 \subset C$, which gives $C \hookrightarrow C_0 \hookrightarrow$. Hence, (a) establishes $C_0 \sim C \sim C_1$. Thus, $A \sim C$. $\square$

Exercise 7. Show that the sets D and E of Exercise 5 have the same cardinality.

Solution. Note that $D = \mathrm{Func}(\mathbb{Z}_+, \mathbb{Z}_+) \sim \mathbb{Z}_+^\omega$, $E = \mathrm{Func}(\mathbb{Z}_+, \{0, 1\}) \sim \{0, 1\}^\omega$, and that $\{0, 1\}^\omega \sim \mathcal{P}(\mathbb{Z}_+)$.

It is clear that there holds $E \hookrightarrow D$. Conversely, we claim $D \hookrightarrow E$. Indeed, the fact $\mathbb{Z}_+ \subset \mathcal{P}(\mathbb{Z}_+)$ and Proposotion 7.1 allow us to have

$$\begin{aligned} \mathrm{Func}(\mathbb{Z}_+, \mathbb{Z}_+) &\hookrightarrow \mathrm{Func}(\mathbb{Z}_+, \mathcal{P}(\mathbb{Z}_+)) \\ &\sim \mathrm{Func}(\mathbb{Z}_+, \mathrm{Func}(\mathbb{Z}_+, \{0,1\})) \\ &\sim \mathrm{Func}(\mathbb{Z}_+ \times \mathbb{Z}_+, \{0,1\}) \\ &\sim \mathrm{Func}(\mathbb{Z}_+, \{0,1\}). \end{aligned}$$

Schroeder-Bernstein theorem then establishes $D \sim E$. □

Exercise 8. Let X denote the two-element set $\{0,1\}$; let $\mathcal{B}$ be the set of *countable* subsets of X^ω. Show that X^ω and $\mathcal{B}$ have the same cardinality.

Solution. A function $\{0,1\}^\omega \mapsto \{x\} \in \mathcal{B}$ is obviously injective. In order to claim by Schroeder-Bernstein theorem that $\mathcal{B} \sim \{0,1\}^\omega$, we show $\mathcal{B} \hookrightarrow \{0,1\}^\omega$.

Observe that countability allows us to write down each $B \in \mathcal{B}$ as $B = \{x^{(1)}, x^{(2)}, \cdots\}$ and that constructing a function that assigns, to $\{x^{(1)}, x^{(2)}, \cdots\}$,

$$\left(x_1^{(1)}, x_2^{(1)}, x_1^{(2)}, x_3^{(1)}, x_2^{(2)}, x_1^{(3)}, x_4^{(1)}, x_3^{(2)}, x_2^{(3)}, x_1^{(4)}, x_5^{(1)}, x_4^{(2)}, x_3^{(3)}, x_2^{(4)}, x_1^{(5)}, \cdots\right)$$

finishes the proof. We do that by considering a function $\varphi : \mathcal{B} \to \{0,1\}^\omega$ given by

$$\varphi\left(\left\{\left(x_i^{(1)}\right)_{i\in\mathbb{Z}_+}, \left(x_i^{(2)}\right)_{i\in\mathbb{Z}_+}, \cdots\right\}\right) = \left(x_{s(n-1)+n-i+1}^{(i-s(n-1))}\right)_{i\in\mathbb{Z}_+}$$

where $n \in \mathbb{Z}_+$ is uniquely determined by given i and the inequality

$$s(n-1) < i \leq s(n),$$

and $s(\cdot)$ is defined as $s(0) := 0$, and $s(n) := \sum_{j=1}^n j$ for $n \geq 1$. It is easy to verify that φ is injective.

See Exercise 2, which shares the same underlying idea as this one. Here we have considered the following way of lining up the elements of B:

$$\begin{aligned} &x_1^{(1)}, \\ &x_2^{(1)}, x_1^{(2)}, \\ &x_3^{(1)}, x_2^{(2)}, x_1^{(3)}, \\ &x_4^{(1)}, x_3^{(2)}, x_2^{(3)}, x_1^{(4)}, \\ &x_5^{(1)}, x_4^{(2)}, x_3^{(3)}, x_2^{(4)}, x_1^{(5)}, \cdots . \end{aligned}$$

Compare this with g there. □

Exercise 9.

(a) The formula

$$\begin{aligned} h(1) &= 1, \\ h(2) &= 2, \\ h(n) &= [h(n+1)]^2 - [h(n-1)]^2 \end{aligned}$$

for $n \geq 2$ is not one to which the principle of recursive definition applies. Show that nevertheless there does exist a function $h : \mathbb{Z}_+ \to \mathbb{R}$ satisfying this formula.

(b) Show that the formula of part (a) does not determine h uniquely.

(c) Show that there is no function $h : \mathbb{Z}_+ \to \mathbb{R}$ satisfying the formula

$$\begin{aligned} h(1) &= 1, \\ h(2) &= 2, \\ h(n) &= [h(n+1)]^2 + [h(n-1)]^2 . \end{aligned}$$

for $n \geq 2$.

Solution.

$\boxed{\text{(a)}}$ Let h satisfy the following formula:

$$\begin{aligned} h(1) &= 1, \\ h(2) &= 2, \\ h(n) &= \sqrt{h(n-1) + [h(n-2)]^2)}, \end{aligned}$$

for $n \geq 3$. Then, principle of recursive definition yields unique $h : \mathbb{Z}_+ \to \mathbb{R}_+$ that satisfies the formula above. It is clear that h also satisfies the given formula.

$\boxed{\text{(b)}}$ Consider the formula

$$\begin{aligned} f(1) &= 1, \\ f(2) &= 2, \\ f(3) &= -\sqrt{3}, \\ f(n) &= \sqrt{f(n-1) + [f(n-2)]^2)}, \\ f(n) &> 0, \end{aligned}$$

for all $n \geq 4$. This defines unique $f : \mathbb{Z}_+ \to \mathbb{R}$ also satisfying the formula at (a), but $f \neq h$, where h is the function we have construct in (a).

$\boxed{\text{(c)}}$ The given formula requires $h(3) \pm 1$, and $h(4)^2 = h(3) - 4$; the latter is impossible. $\square$

8 The Principle of Recursive Definition

Exercise 1. Let $(b_1, b_2, \cdots)$ be an infinite sequence of real numbers. The sum $\sum_{k=1}^{n} b_k$ is defined by induction as follows:

$$\begin{aligned} \sum_{k=1}^{1} b_k &= b_1, \\ \sum_{k=1}^{n} b_k &= \left(\sum_{k=1}^{n-1} b_k \right) + b_n \end{aligned}$$

for $n > 1$. Let A be the set of real numbers; choose ρ so that Theorem 8.4 applies to define the sum rigorously.

Solution. For $f \in \operatorname{Func}(\{1, \cdots, m\}, \mathbb{R})$, define $\rho(f) := f(m) + b_{m+1}$. Let b_1 be the initial value for h. $\square$

Exercise 2. Let $(b_1, b_2, \cdots)$ be an infinite sequence of real numbers. We define the product $\prod_{k=1}^{n} b_k$ by the equations

$$\begin{aligned} \prod_{k=1}^{1} b_k &= b_1, \\ \prod_{k=1}^{n} b_k &= \left(\prod_{k=1}^{n-1} b_k \right) \cdot b_n \end{aligned}$$

for $n > 1$. Use Theorem 8.4 to define this product rigorously.

Solution. For $f \in \operatorname{Func}(\{1, \cdots, m\}, \mathbb{R})$, define $\rho(f) := f(m) \cdot b_{m+1}$. Let b_1 be the initial value for h. Theorem 8.4 gives $h \in \operatorname{Func}(\mathbb{Z}_+, \mathbb{R})$ such that

$$\begin{aligned} h(1) &= b_1, \\ h(n) &= \prod_{k=1}^{n} b_k \end{aligned}$$

for $n \geq 2$. $\square$

Exercise 3. Obtain the definitions of a^n and $n!$ for $n \in \mathbb{Z}_+$ as special cases of Exercise 2.

Solution. For a^n, assume $b_n = a$ for all $n \in \mathbb{Z}_+$. For $n!$, assume $b_n = n$ for all $n \in \mathbb{Z}_+$. $\square$

Exercise 4. The *Fibonacci numbers* of number theory are defined recursively by the formula

$$\begin{aligned} \lambda_1 &= \lambda_2 = 1, \\ \lambda_n &= \lambda_{n-1} + \lambda_{n-2} \end{aligned}$$

for $n > 2$. Define them rigorously by use of Theorem 8.4.

Solution. For $f \in \mathrm{Func}(\{1, \cdots, m\}, \mathbb{R})$, define

$$\rho(f) := \begin{cases} 1 & : f \in \mathrm{Func}(S_2, \mathbb{R}) \\ f(m) + f(m-1) & : f \in \mathrm{Func}(S_m, \mathbb{R}) \wedge m > 2. \end{cases}$$

Let $a_0 := 1$ be the initial value for h. $\square$

Exercise 5. Show that there is a unique function $h : \mathbb{Z}_+ \to \mathbb{R}$ satisfying the formula

$$\begin{aligned} h(1) &= 3, \\ h(i) &= [h(i-1) + 1]^{1/2} \end{aligned}$$

for $i > 1$.

Solution. We only have to construct a good function ρ since theorem 8.4 then gives the required function h. To this end, for $f \in \mathrm{Func}(\{1, \cdots, m\}, \mathbb{R})$, set $\rho(f) := (f(m) + 1)^{1/2}$. $\square$

Exercise 6.

(a) Show that there is no function $h : \mathbb{Z}_+ \to \mathbb{R}$ satisfying the formula

$$\begin{aligned} h(1) &= 3, \\ h(i) &= [h(i-1) - 1]^{1/2} \end{aligned}$$

Explain why this example does not violate the principle of recursive definition.

(b) Consider the recursive formula

$$\begin{aligned} h(1) &= 3, \\ h(i) &= \begin{cases} [h(i-1) - 1]^{1/2} & : h(i-1) > 1 \\ 5 & : h(i-1) \leq 1 \end{cases} \end{aligned}$$

for $i > 1$.

Show that there exists a unique function $h : \mathbb{Z}_+ \to \mathbb{R}$ satisfying this formula.

Solution.
$\boxed{\text{(a)}}$ The given formula requires $h(1) = 3$, $h(2) = \sqrt{2}$, $h(3) = \left(\sqrt{2}-1\right)^{1/2}$, but $\left(\left(\sqrt{2}-1\right)^{1/2} - 1\right)^{1/2}$ does not exist in $\mathbb{R}$, which implies $h(4)$ cannot be defined.

Note that the corresponding ρ should be given as

$$\rho(f) := (f(m) - 1)^{1/2}$$

for $f \in \mathrm{Func}(\{1, \cdots, m\}, \mathbb{R})$, but ρ is not a function from $\bigcup_{m\in\mathbb{Z}_+} \mathrm{Func}(\{1, \cdots, m\}, \mathbb{R})$ to "$\mathbb{R}_+$."

This trouble is because of the absence of a good function ρ, not a violation of the principle of recursive definition, which gives a requested function h if the hypothesis, for instance, existence of ρ, is satisfied, and says nothing if otherwise. In fact, the principle of recursive definition can be expressed as

$$\begin{aligned} &a_0 \in A \wedge \rho \in \mathrm{Func}(\bigcup_{m\in\mathbb{Z}_+} \mathrm{Func}(\{1, \cdots, m\}, A), A) \\ \Rightarrow\ &{}^{\exists 1}h \in \mathrm{Func}(\mathbb{Z}_+, A) \left[h(1) = a_0 \wedge {}^{\forall}i \in \mathbb{Z}_+ \setminus \{1\} \left[h(i) = \rho\left(h|_{\{1,\cdots,i-1\}}\right)\right]\right] \end{aligned}$$

and the whole statement is still true even if ρ is absent and the hypothesis is false. (Remember that statement $A \Rightarrow B$ is true if A is false.)

$\boxed{\text{(b)}}$ For $f \in \mathrm{Func}(\{1, \cdots, m\}, \mathbb{R})$, define

$$\rho(f) := \begin{cases} (f(m) - 1)^{1/2} & : f(m) > 1 \\ 5 & : f(m) \le 1. \end{cases}$$

Then, $\rho(f) > 0$ for all f. The result follows from Theorem 8.4. □

Lemma 8.1. *Let A be a set, and let $a_0 \in A$. Suppose ρ is a function from $\bigcup_{m\in\mathbb{Z}_+} \mathrm{Func}(\{1, \cdots, m\}, A)$ to A. Then for all $n \in \mathbb{Z}_+$ there exists a unique function $h_n : \{1, \cdots, n\} \to A$ such that*

$$\begin{aligned} h_n(1) &= a_0 \\ h_n(i) &= \rho\left(h_n|_{\{1,\cdots,i-1\}}\right) \end{aligned}$$

for all $1 < i \le n$.

Proof. Let X be the set of all $n \in \mathbb{Z}_+$ for which the lemma holds. Suppose $\{1, \cdots, m\} \in X$, and let h_{n-1} satisfy (8.1) for all i in its domain. Since

$h_{n-1} \in \mathrm{Func}(\{1, \cdots, n-1\}, A)$, we can define a function $h_n : \{1, \cdots, n\} \to A$ via

$$h_n(i) := \begin{cases} h_{n-1}(i) & : i \in \{1, \cdots, n-1\} \\ \rho\left(h_n|_{\{1,\cdots,i-1\}}\right) & : i = n \end{cases}$$

It is clear that h_n satisfies (8.1).

To show the uniqueness of h_n, suppose $g : \{1, \cdots, n\} \to A$ also satisfies (8.1) for all i in its domain. Let Y be a subset of $\mathbb{Z}_+$ consisting of all m such that we have $g(i) = h_n(i)$ for all $i \in \{1, \cdots, n\} \cap \{1, \cdots, m\}$. Suppose $\{1, \cdots, k-1\} \in Y$. We may assume $k - 1 \le n$. We have $g|_{\{1,\cdots,k-1\}} = h_n|_{\{1,\cdots,k-1\}}$ by assumption, and hence

$$g(k) = \rho\left(g|_{\{1,\cdots,k-1\}}\right) = \rho\left(h_n|_{\{1,\cdots,k-1\}}\right) = h_n(k),$$

which yields $k \in Y$. Thus, strong induction principle gives $Y = \mathbb{Z}_+$, establishing the uniqueness of h_n. We the conclude that $n \in X$ and that $X = \mathbb{Z}_+$. □

Exercise 7. Prove Theorem 8.4.

Solution. Define a function $h : \mathbb{Z}_+ \to A$ by setting its rule to be the union of the function h_n, we have constructed in Lemma 8.1. h is proved to be a function by exactly the same argument as Theorem 8.3 since there we have exploited nothing of the property of the range of f_n.

We show that h satisfies

$$\begin{aligned} h(1) &= a_0 \\ h(i) &= \rho\left(h|_{\{1,\cdots,i-1\}}\right) \end{aligned} \tag{8.1}$$

for all $i > 1$. Let W be the set of all $i \in \mathbb{Z}_+$ for which (8.1) is true. It is easy to see $W = \mathbb{Z}_+$ by induction.

To prove the uniqueness of h, argue as we have done in Lemma 8.1. □

Exercise 8. Verify the following version of the principle of recursive definition: Let A be a set. Let ρ be a function assigning, to every function f mapping a section S_n of $\mathbb{Z}_+$ into A, a element $\rho(f)$ of A. Then there is a unique function $h : \mathbb{Z}_+ \to A$ such that $h(n) = \rho(h|_{S_n})$ for each $n \in \mathbb{Z}_+$.

Solution. Note that $S_1 = \emptyset$, and that $h|_{S_1}$ is well-defined as a unique empty function, denoted by $\emptyset$.(see NOTE **??**). Set $a_0 := \rho(\emptyset)$ and apply Theorem 8.4. □

9 Infinite Sets and the Axiom of Choice

Remark 9.1 (Equivalence of the axiom of choice and choice function)**.**

We have seen the axiom of choice imply Lemma 9.2, the existence of choice function. But the converse is also true. Indeed, if $\mathcal{A}$ is a collection of disjoint nonempty sets, and if c is a choice function for $\mathcal{A}$, then the set $C := \bigcup_{A\in\mathcal{A}} c(A)$ has the required property. So, we from now on identify the choice axiom and Lemma 9.2.

Exercise 1. Define a injective map $f : \mathbb{Z}_+ \to X^\omega$, where X is the two element set $\{0, 1\}$, without using the choice axiom.

Solution. Letting $\varphi \in \mathrm{Bij}(\mathcal{P}(\mathbb{Z}_+), \{0,1\}^\omega)$, define a function $f : \mathbb{Z}_+ \to X^\omega$ via

$$f(n) := \varphi \circ i(n),$$

where $i : \mathbb{Z}_+ \to \mathcal{P}(\mathbb{Z}_+)$ is given by $i(x) := \{x\}$, which is obviously injective. □

Exercise 2. Find if possible a choice function for each of the following collections, without using the choice axiom:

(a) The collection $\mathcal{A}$ of nonempty subsets of $\mathbb{Z}_+$.

(b) The collection $\mathcal{B}$ of nonempty subsets of $\mathbb{Z}$.

(c) The collection $\mathcal{C}$ of nonempty subsets of the rational numbers $\mathbb{Q}$.

(d) The collection $\mathcal{D}$ of nonempty subsets of X^ω, where $X = \{0, 1\}$.

Solution.

$\boxed{\text{(a)}}$ It is easy to check that $\bigcup_{A\in\mathcal{A}}(A, x_A)$ qualifies as the rule for a required choice function, where x_A denotes the smallest element of A.

$\boxed{\text{(b)}}$ Let $f : \mathbb{Z} \to \mathbb{Z}_+$ be a function defined at §7 Example 1, which is proved to be bijective at §7 Exercise 2. If we set c to be a choice function for $\mathcal{A}$ we have found in (a), $f^{-1} \circ c \circ f$ is a choice function for $\mathcal{B}$.

$\boxed{\text{(c)}}$ Note that $\mathbb{Q}$ is shown to be countable without the choice axiom.(See Exercise 4 if you need Theorem 7.5 to establish the countability.) Hence, there exists a bijective function $g : \mathbb{Q} \to \mathbb{Z}_+$. Then $g^{-1} \circ c \circ g$ is a requested choice function.

$\boxed{\text{(d)}}$ We cannot find it without the choice axiom, and we need some knowledge about well-ordering theorem to explain why. (so, you can skip here for the first reading.) As Exercise 8 shows, X^ω and $\mathbb{R}$ have the same cardinality, which implies, as we have seen in (b) and (c), that the existence of a choice function for $\mathcal{D}$ is equivalent to that for $\mathcal{P}(\mathbb{R})\setminus\emptyset$. However, Exercise 7 in Supplementary Exercise tells us that this is true if and only if the well-ordering theorem is valid on $\mathbb{R}$, which, as a well-known result in set theory says, is true only under the choice axiom. $\square$

Exercise 3. Suppose that A is a set and $\{f_n\}_{n\in\mathbb{Z}_+}$ is a given indexed family of injective functions

$$f_n : \{1, \cdots, n\} \to A.$$

Show that A is infinite. Can you define an injective function $f : \mathbb{Z}_+ \to A$ without using the choice axiom?

Solution. For every $n \in \mathbb{Z}_+$, let $A_n := f_n(\{1, \cdots, n\})$, and let a function $g_n : \{1, \cdots, n\} \to A_n$ be given by $g_n(i) := f_n(i)$. It is clear that every g_n is bijective. As we have seen in Exercise 2(a), there exists, without the choice axiom, a choice function, denoted by c, for $\mathcal{P}(\mathbb{Z}_+)\setminus\emptyset$, and then, the argument in Exercise 2(d) shows that $c_n := g_n \circ c \circ g_n^{-1}$ is a choice function for $\mathcal{P}(A_n)\setminus\emptyset$ for every n.

Now we are ready to construct the required function. Consider the following formula for $f : \mathbb{Z}_+ \to A$:

$$\begin{aligned} f(1) &= A_1 \\ f(n) &= c_n\left(A_n \setminus f(\{1, \cdots, n-1\})\right) \end{aligned} \tag{9.1}$$

for every $n > 1$. In order to show that (9.1) is what principle of recursive definition applies, we claim that the formula designates for f the unique element of A for every $n \in \mathbb{Z}_+$. We do this by induction. Assume our claim holds for up to $i-1$. We deduce from the fact A_n has exactly n elements while $f(\{1, \cdots, n-1\})$ does at most $n-1$ that $A_n \setminus f(\{1, \cdots, n-1\} \neq \emptyset$, of which choice function c_n extracts one element. This means that the claim holds for i. Then strong induction principle gives that our claim holds for all $n \in \mathbb{Z}_+$. Hence, (9.1) defines unique $f : \mathbb{Z}_+ \to A$ that satisfies the formula.

We at last insists that f is injective. In fact, if $n < m$, or equivalently $n \leq m-1$, then the fact that c_n is a choice function for $\mathcal{P}(A_n) \setminus \emptyset$ gives

$$f(m) \in A_m \setminus f\left(\{1, \cdots, n-1, n, n+1, \cdots, m-1\}\right),$$

which yields $f(n) \neq f(m)$. Thus, f is injective, and we complete the construction of f without the choice axiom.

Note that the implication of (1) $\Rightarrow$ (3) in Theorem 9.1 is valid without the choice axiom. Thus, the fact A is infinite can also be proved without it. $\square$

Exercise 4. There was a theorem in §7 whose proof involved an infinite number of arbitrary choices. Which one was it? Rewrite the proof so as to make explicit the use of the choice axiom.

Solution. That is where Theorem 7.5 says "Because each A_n is countable, we can choose, for each n, a surjective function $f_n : \mathbb{Z}_+ \to A$". Here the set A_n is indexed by the set J, which might be infinite. If it is infinite, we need the choice axiom (or a choice function) in order to extract f_n from each of nonempty set $\mathrm{Surj}(\mathbb{Z}_+, A_n)$. If otherwise, say, if J is finite, of course we do not need it since the finite axiom of choice suffices.

Now it is obvious how we should rewrite it. □

Exercise 5.

(a) Use the choice axiom to show that if $f : A \to B$ is surjective, then f has a right inverse $h : B \to A$.

(b) Show that if $f : A \to B$ is injective and A is not empty, then f has a left inverse. Is the axiom of choice needed?

Solution. This Exercise is the converse of §2 Exercise 5(a), which insists that, *if a function admits a right inverse, the function is surjective; if a function admits a left inverse, the function is injective.*

$\boxed{\text{(a)}}$ Consider the indexed family $\{f^{-1}(b)\}_{b\in B}$ of sets. Each $f^{-1}(b)$ is nonempty since f is surjective. The choice axiom allows us to have a choice function $c : B \to \bigcup_{b\in B} f^{-1}(b)$. Then we see that for every $b \in B$

$$f \circ c(b) = f(c(b)) = b$$

since $c(b) \in f^{-1}(b)$. Seeing c as a function from B to A, we conclude that f has a right inverse c.

$\boxed{\text{(b)}}$ The axiom choice is not needed this time. Indeed, if $f \in \mathrm{Inj}(A, B)$ and $A \neq \emptyset$, then a function $\varphi : A \to f(A)$ given by $\varphi(a) = f(a)$ is bijective, and hence we have

$$\varphi^{-1} \circ f(a) = a$$

for all $a \in A$. (Extend φ^{-1} onto B if necessary by setting arbitrary value outside $f(A)$.) □

Remark 9.2. Here are several technical remark about empty function.

In Exercise 5(b), we cannot drop the assumption that "A is not empty"; $A = \emptyset$ means that f and therefore φ are an empty function, which is born injective but not surjective whatever the range is. Hence, φ by no means has the inverse.

Also, note that in Exercise 5(a), the assumption f is surjective rejects the possibility of f being an empty function, and consequently guarantees that A and B are not empty. We use this fact implicitly in the above proof.

Exercise 6. Most of famous paradoxes of naive set theory are associated in some way or other with the concept of the "set of all sets." None of the rules we have given for forming sets allows us to consider such a set. And for good reason—the concept itself is self-contradictory. For suppose that $\mathcal{A}$ denotes the "set of all sets."

(a) Show that $\mathcal{P}(\mathcal{A}) \subset \mathcal{A}$; derive a contradiction.

(b) (*Russell's paradox.*) Let $\mathcal{B}$ be the subset of $\mathcal{A}$ consisting of all sets that are not element of themselves;

$$\mathcal{B} = \{A \mid A \in \mathcal{A} \wedge A \notin A\}$$

(Of course, there may be *no* set A such that $A \in A$; if such is the case, then $\mathcal{B} = \mathcal{A}$.) Is $\mathcal{B}$ an element of itself or not?

Solution.

$\boxed{\text{(a)}}$ If $a \in \mathcal{P}(A)$, then a itself is a set by the definition of $\mathcal{A}$. Hence, $a \in \mathcal{A}$. Thus, $\mathcal{P}(\mathcal{A}) \subset \mathcal{A}$. But $\mathcal{P}(A) \subset A$ leads $\mathcal{P}(A) \hookrightarrow A$, contradicting to Theorem 7.8.

$\boxed{\text{(b)}}$ BOTH !

If we assume $\mathcal{B} \notin \mathcal{B}$, then the definition of $\mathcal{B}$ leads to $\mathcal{B} \in \mathcal{B}$. On the other hand, if we suppose $\mathcal{B} \in \mathcal{B}$, then we are led to $\mathcal{B} \notin \mathcal{B}$. $\square$

Exercise 7. Let A and B be two nonempty sets. If there is an injection of B into A, but no injection of A into B, we say that A has ***greater cardinality*** than B.

(a) Conclude from Theorem 9.1 that every uncountable set has greater cardinality than $\mathbb{Z}_+$.

(b) Show that if A has greater cardinality than B, and if B has greater cardinality than C, then A has greater cardinality than C.

(c) Find a sequence $A_1, A_2, \cdots$ of infinite sets, such that for each $n \in \mathbb{Z}_+$, the set A_{n+1} has greater cardinality than A_n.

(d) Find a set that for every n has greater cardinality than A_n.

Solution.

$\boxed{\text{(a)}}$ Let A be a uncountable set, and B a countable set. Theorem 9.1 tells us that $B \hookrightarrow A$ while definition of uncountable set implies $\mathrm{Bij}(A, B) = \emptyset$. Thus, contrapositive of Schroeder-Bernstein theorem gives $\mathrm{Inj}(A, B) = \emptyset$, from which we conclude the result.

$\boxed{\text{(b)}}$ §2 Exercise 4 gives $C \hookrightarrow A$, and contrapositive of that Exercise also implies $\mathrm{Inj}(A, C) = \emptyset$. Thus, the result follows.

$\boxed{\text{(c)}}$ Define A_n by the formula below:

$$\begin{aligned} A_1 &= A, \\ A_n &= \mathcal{P}(A_{n-1}) \end{aligned}$$

for all $n > 1$. It is easy to check that principle of recursion definition applies for this formula, and hence A_n is unique defined.

A_{n+1} has greater cardinality than A_n since we see $A_n \hookrightarrow \mathcal{P}(A_n) = A_{n+1}$ while there holds $\mathrm{Inj}(\mathcal{P}(A_n), A_n) = \emptyset$ by Theorem 7.8.

$\boxed{\text{(d)}}$ Define $A^* := \mathcal{P}(\bigcup_{n \in \mathbb{Z}_+} A_n)$, and repeat the same argument. □

Exercise 8. Show that $\mathcal{P}(\mathbb{Z}_+)$ and $\mathbb{R}$ have the same cardinality.

Solution. It is easy to see, by considering decimal expansion for real numbers, that

$$[0, 1) \hookrightarrow \{0, 1, 2, \cdots, 8, 9\}^\omega \sim \{0, 1\}^\omega \hookrightarrow [0, 1). \tag{9.2}$$

For instance, using the fact that any member x of $[0, 1)$ can be uniquely written, if we limit the expression to those not having infinite 9's tail, as

$$x = 0.x_1 x_2 x_2 \cdots,$$

where $x_i \in \{0, 1, 2, \cdots, 8, 9\}$ for every $i \in \mathbb{Z}_+$, we can define a injection $\iota : [0, 1) \to \{0, 1, 2, \cdots, 8, 9\}^\omega$ via

$$\iota(x) := (x_i)_{i \in \mathbb{Z}_+}.$$

Since we can deduce from Schroeder-Bernstein theorem that all the sets at (9.2) have the same cardinality, it remains to show $[0, 1) \sim \mathbb{R}$, and $\{0, 1, 2, \cdots, 8, 9\}^\omega \sim \{0, 1\}^\omega \sim \mathcal{P}(\mathbb{Z}_+)$.

We first work on the former. We claim

$$\mathbb{R} \sim (-1, 1) \sim (0, 1) \sim [0, 1).$$

For the first equivalence, see §3 Exercise 10; for the second consider a function $f : (-1, 1) \to (0, 1)$ given by

$$f(x) := \frac{1}{2}x + 1;$$

for the third, verify that a function $g : [0, 1) \to (0, 1)$ given by

$$g(x) := \begin{cases} \frac{1}{2} & : x = 0 \\ \frac{x}{2} & : {}^{\exists}n \in \mathbb{Z}_+ \left[x = \frac{1}{2^n}\right] \\ x & : \text{otherwise} \end{cases}$$

is bijective.

For the latter, we have, as we have repeatedly seen, $\mathcal{P}(\mathbb{Z}_+) \sim \mathbb{Z}_+^\omega \sim \{0, 1\}^\omega$ and

$$\{0, 1\}^\omega \hookrightarrow \{0, 1, 2, \cdots, 8, 9\}^\omega \hookrightarrow \mathbb{Z}_+^\omega,$$

which yield, by Schroeder-Bernstein theorem, that

$$\{0, 1, 2, \cdots, 8, 9\}^\omega \sim \mathbb{Z}_+^\omega \sim \mathcal{P}(\mathbb{Z}_+).$$

Thus, we conclude $\mathbb{R} \sim \mathcal{P}(\mathbb{Z}_+)$. □

10 Well-Ordered Sets

Exercise 1. Show that every well-ordered set has the least upper bound property.

Solution. See §3 Exercise 13,14, where we have established that an ordered set has the least upper bound property if and only if it has the greatest lower bound property.

Let W be a well-ordered set, and let W_0 be its nonempty subset bounded below. Definition of well-ordered set gives a smallest element s of W_0. Since s is also an infimum of W_0, we conclude that W has the the greatest lower bound property. □

Exercise 2.

(a) Show that in a well-ordered set, every element except the largest (if exists) has an immediate successor.

(b) Find a set in which every element has an immediate successor that is not well-ordered.

Solution.

$\boxed{\text{(a)}}$ Let W be a well-ordered set, and let a be a non-largest element of W. Set $W_0 := \{b \in W \mid a < b\}$, which is nonempty by the choice of a. Then W_0 admits a smallest element s, for which we have $(a, s) = \emptyset$. Thus, s is the immediate successor of a.

$\boxed{\text{(b)}}$ $\mathbb{Z}$. □

Exercise 3. Both $\{1,2\} \times \mathbb{Z}_+$ and $\mathbb{Z}_+ \times \{1,2\}$ are well-ordered in the dictionary order. Do they have the same order type?

Solution. No, they have different order types. Note that every element of $\{1,2\} \times \mathbb{Z}_+$ admits the immediate successor while that of $\mathbb{Z}_+ \times \{1,2\}$ does not. (consider $(a,1) \in \mathbb{Z}_+ \times \{1,2\}$, for instance.) Our claim follows from this fact and Proposition 3.3.

See also §3 Exercise 12, where we discussed a similar question. $\square$

Exercise 4.

(a) Let $\mathbb{Z}_-$ denote the set of negative integers in the usual order. Show that a simply ordered set A fails to be well-ordered if and only if it contains a subset having the same order type as $\mathbb{Z}_-$.

(b) Show that A is simply ordered and every countable subset of A is well-ordered, then A is well-ordered.

Solution.

$\boxed{\text{(a)}}$ "if" part is obvious since in that case that subset does not admits a smallest element, and hence A is not well-ordered.

Conversely, suppose that A is not well-ordered. Let A_0 be a subset of A that does not admit a smallest element. Also, let $a \in A_0$ and let c be a choice function for A_0. Consider the following formula for $f : \mathbb{Z}_+ \to A_0$:

$$\begin{aligned} f(-1) &:= a, \\ f(-i) &:= c\left(S_{f(-i+1)}\right) \end{aligned}$$

for all $i \in \mathbb{Z}_+$, where S_α denotes the section of A_0 by $\alpha \in A_0$ given by $S_\alpha := \{\alpha' \in A_0 \mid \alpha' < \alpha\}$. It is easy to check that principle of recursive definition works for this formula, and that so derived f is injective and order-preserving. Thus, $f(\mathbb{Z}_-)$ is a subset of A_0 which has the same order type as $\mathbb{Z}_-$.

$\boxed{\text{(b)}}$ This is a direct consequence of contrapositive statement of (a). $\square$

Exercise 5. Show the well-ordering theorem implies the choice axiom.

Solution. Let $\mathcal{A}$ be a collection of nonempty sets. It suffices to show that there exists a choice function for $\mathcal{A}$. Let $A^* := \bigcup_{A \in \mathcal{A}}$. Apply to $\mathcal{A}$ the well-ordering theorem so that $\mathcal{A}$ is well-ordered. Then, a function $c : \mathcal{A} \to A^*$ that assigns, to each $A \in \mathcal{A}$, a smallest element $c(A)$ of A is a choice function for $\mathcal{A}$. $\square$

Exercise 6. Let S_Ω be the minimal uncountable well-ordered set.

(a) Show that S_Ω has no largest element.

(b) Show that for every $\alpha \in S_\Omega$, the set $\{x \mid \alpha < x\}$ is uncountable.

(c) Let X_0 be the subset of S_Ω consisting of all elements x such that x has no immediate predecessor. Show that X_0 is uncountable.

Solution.

$\boxed{\text{(a)}}$ It suffices to show that every elements of S_Ω has the immediate successor in S_Ω. It is obvious that $S_\Omega \cup \Omega$ has a unique largest element Ω. Let $a \in S_\Omega$. Exercise 2(a) implies a has its immediate successor b in $S_\Omega \cup \Omega$. Observe that $S_b = S_a \cup \{a\}$ is countable while S_Ω is uncountable, and hence $S_b \neq S_\Omega$. Thus, $b \neq \Omega$. (Remember, in general, for a simply ordered set, $S_\alpha = S_\beta$ if and only if $\alpha = \beta$.) This means $b \in S_\Omega$.

We might argue as follows: suppose, to the contrary, that m is a largest element of S_Ω. Then we are led to $S_\Omega = S_m \cup \{m\}$. But this is impossible since S_Ω is uncountable while $S_m \cup \{m\}$ is countable. Thus, S_Ω has no largest element.

$\boxed{\text{(b)}}$ Given $\alpha \in S_\Omega$, it follows from (a) that the set $\{x \mid \alpha < x\}$ is unbounded above in S_Ω. Thus, we deduce from Theorem 10.3 that $\{x \mid \alpha < x\}$ is uncountable.

$\boxed{\text{(c)}}$ By Theorem 10.3, it suffices to show that X_0 is unbounded above. Let $s : S_\Omega \to S_\Omega$ be a function that assigns, to each $a \in S_\Omega$, the immediate successor of a. It is easy to verify that s preserves order. Given $\alpha \in S_\Omega$, let $A := \{\alpha, s(\alpha), s^2(\alpha), \cdots\}$. Since A is countable, A admits a supremum $\sup A$ in S_Ω by Theorem 10.3 and Exercise 1. We claim $\sup A \in X_0$. In fact, for any $x \in S_\Omega$ with $x < \sup A$, there exists $y \in A$ such that $x < y < S_\Omega$, that is, $(x, \sup A) \neq \emptyset$. Hence, $\sup A \in X_0$. It follows from $\alpha < \sup A$ that X_0 is unbounded above. $\square$

Exercise 7 (The principle of transfinite induction)**.** Let J be a well-ordered set. A subset J_0 of J is said to be ***inductive*** if for every $\alpha \in J$,

$$(S_\alpha \subset J_0) \Rightarrow \alpha \in J_0.$$

Theorem. If J is a well-ordered set and J_0 is an inductive subset of J, then $J_0 = J$.

Solution. Let J be a well-ordered set, and let J_0 be a subset of J. Suppose $J \setminus J_0 \neq \emptyset$. Then, there exists a smallest element x of $J \setminus J_0$. We see that $S_x \subset J_0$ and $x \notin J_0$. Thus, J_0 is not inductive. $\square$

Exercise 8.

(a) Let $(A_1, <_1)$ and $(A_2, <_2)$ be disjoint well-ordered sets. Define an order relation on $A_1 \cup A_2$ by letting $a < b$ either if $a, b \in A_1$ and $a <_1 b$, or if $a, b \in A_2$ and $a <_2 b$, or if $a \in A_1$ and $b \in A_2$. Show that this is a well-ordering.

(b) Generalize (a) to an arbitrary family of disjoint well-ordered sets, indexed by a well-ordered set.

Solution.

$\boxed{\text{(a)}}$ We first verify that the given relation is actually an order relation. Noting that A_1 and A_2 are disjoint, comparability and nonreflexivity are obvious. To show transitivity, suppose $a < b$ and $b < c$. If $a, b, c \in A_i$ for some i, then we have $a <_{A_i} c$, and so $a < c$. If $a, b \in A_1$ and $c \in A_2$, then definition of $<$ gives $a < c$. The same argument works for the remaining cases. Hence, the given relation is an order relation.

We establish $(A_1 \cup A_2, <)$ is a well-ordered set. Let B is a nonempty subset of $A_1 \cup A_2$. If B is included in A_i for some i, then there exists a smallest element s of B in $(A_i, <_i)$. It is obvious that s is also a smallest element of B in $(A_1 \cup A_2, <)$. If, on the other hand, $B \cap A_1 \neq \emptyset$ and $B \cap A_2 \neq \emptyset$, then a smallest element of $B \cap A_1$ in $(A_i, <_i)$ turns out to be that of $B \cap A_1$ in $(A_1 \cup A_2, <)$ because every element of $B \cap A_1$ is smaller than that of $B \cap A_2$.

$\boxed{\text{(b)}}$ Let $(W_i, <_i)$ be a family of disjoint well-ordered sets, indexed by a well-ordered set I. Define $W := \bigcup_{i \in I} W_i$ and an order relation on W by setting $a < b$ either if $a, b \in W_i$ and $a <_i b$ for some i, or if $a \in W_i$ and $b \in W_j$ for $i < j$.

The same argument as (a) shows that $<$ is an order relation on W. Let B be a subset of W and let i^* be a smallest element of the set $\{i \in I \mid B \cap W_i \neq \emptyset\}$. It is then easy to show that a smallest element of $B \cap W_{i^*}$ is also that of $(W, <)$ by the way $<$ is introduced. □

Remark 10.1. In Exercise 8, we require $(A_1, <_1)$ and $(A_2, <_2)$ to be disjoint. But this is not an essential assumption:

Given two (not necessarily disjoint) well-ordered sets $(B, <_B)$ and $(C, <_C)$, define new disjoint well-ordered sets $(\{1\} \times B, <_{B'})$ and $(\{2\} \times C, <_{C'})$, where $<_{B'}$ is defined to be $(1, b_1) <_{B'} (1, b_2)$ if $b_1 <_B b_2$, and similarly for $<_{C'}$. It is obvious that $(B, <_B)$ and $(C, <_C)$ have the same order types as $(\{1\} \times B, <_{B'})$ and $(\{2\} \times C, <_{C'})$, respectively. Thus, we may assume two well-ordered sets are disjoint if necessary.

Exercise 9. Consider the subset A of $(\mathbb{Z}_+)^\omega$ consisting of all infinite sequence of positive integers $\boldsymbol{x} = (x_1, x_2, \cdots)$ that end in an infinite string of 1's. Give A the following order: $\boldsymbol{x} < \boldsymbol{y}$ if $x_n < y_n$ and $x_i = y_i$ for $i > n$.

(a) Show that for every n, there is a section of A that has the same order type as $(\mathbb{Z}_+)^n$ in the dictionary order.

(b) Show that A is well-ordered.

Solution. We first note that A is an ordered set with the given order. Given $n \in \mathbb{Z}_+$, let $I^{(n)} := \{\underbrace{1, 1, \cdots, 1}_{n}, 2, 1, 1, \cdots\}$. Let $S(\boldsymbol{x})$ denote the section of A by $\boldsymbol{x}$.

$\boxed{\text{(a)}}$ We claim that for every $n \in \mathbb{Z}_+$ there exists an order isomorphism from $S\left(I^{(n)}\right)$ to $\mathbb{Z}_+^n$. First observe

$$S\left(I^{(n)}\right) = \{\{x_1, x_2, \cdots, x_n, 1, 1, \cdots\} \mid x_i \in \mathbb{Z}_+\}$$

We prove our claim by induction on n. It is easy to check that a function given by

$$S\left(I^{(1)}\right) \ni \{x, 1, 1, \cdots\} \mapsto x \in \mathbb{Z}_+$$

is an order isomorphism. Assume the claim holds for $n-1$, and let f be an order isomorphism from $S\left(I^{(n-1)}\right)$ to $\mathbb{Z}_+^{n-1}$. Define a function $h : S\left(I^{(n)}\right) \to \mathbb{Z}_+ \times \mathbb{Z}_+^{n-1}$ via

$$h\left(\{x_1, x_2, \cdots, x_{n-1}, x_n, 1, 1, \cdots\}\right) \to \left(x_n, f(x_1, \cdots, x_{n-1}, 1, 1, \cdots)\right)$$

It is clear that h is injective. Also, It is straightforward to verify that h preserves order. Since there exists an order isomorphism from $\mathbb{Z}_+ \times \mathbb{Z}_+^{n-1}$ to $\mathbb{Z}_+^n$, there also exists an order isomorphism from $S\left(I^{(n)}\right)$ to $\mathbb{Z}_+^n$.

$\boxed{\text{(b)}}$ Let A_0 be a nonempty subset of A. Let $\boldsymbol{x} \in A_0$. We can write $\boldsymbol{x} = (x_1, x_2, \cdots, x_n, 1, 1, \cdots)$ for some $n \in \mathbb{Z}_+$. Observe $\boldsymbol{x} \in S\left(I^{(n)}\right)$ and so $A_0 \cap S\left(I^{(n)}\right) \neq \emptyset$. Then, $f\left(A_0 \cap S\left(I^{(n)}\right)\right)$ is a nonempty subset of $\mathbb{Z}_+^n$, where f is an order isomorphism from $S\left(I^{(n)}\right)$ to $\mathbb{Z}_+^n$. Since $\mathbb{Z}_+^n$ is well-ordered, $f\left(A_0 \cap S\left(I^{(n)}\right)\right)$ admits a smallest element m. So, $f^{-1}(m)$ is a smallest element of $A_0 \cap S\left(I^{(n)}\right)$. It is obvious that $f^{-1}(m)$ is also a smallest element of A_0. Thus, A is a well-ordered set. $\square$

Exercise 10 (pre-general principle of recursive definition)**.**
Theorem. Let J and C be well-ordered sets; assume that there is no surjective function from a section of J onto C. Then there exists a unique function $h : J \to C$ satisfying the equation

$$h(x) = \min\left[C \setminus h(S_x)\right] \tag{10.1}$$

for each $x \in J$, where S_x is the section of J by x.

Proof.

(a) If h and k map sections of J, or all of J, into C and satisfy (10.1) for all x in their respective domains, show that $h(x) = k(x)$ for all x in both domains.

(b) If there exists a function $h : S_\alpha \to C$ satisfying (10.1), show that there exists a function $k : S_\alpha \cup \{\alpha\} \to C$ satisfying (10.1).

(c) If $K \subset J$ and for all $\alpha \in K$ there exists a function $h_\alpha : S_\alpha \to C$ satisfying (10.1), show that there exists a function

$$k : \bigcup_{\alpha \in K} S_\alpha \to C$$

satisfying (10.1).

(d) Show by transfinite induction that every $\beta \in J$, there exists a function $h_\beta : S_\beta \to C$ satisfying (10.1).

(e) Prove the theorem.

Solution.

$\boxed{\text{(a)}}$ This is a simple consequence of the principle of transfinite induction. Let D_h and D_k be the domain of h and k, respectively. Suppose h and k satisfy (10.1) for all x in their respective domains. Let D be a subset of $D_h \cap D_k$ consisting of all x for which we have

$$h(x) = k(x).$$

Note that $h(S_{\min J}) = h(\emptyset) = \emptyset$ and so $h(\min J) = \min C = k(\min J)$ is well-defined. We deduce from the fact D is inductive that $D = D_h \cap D_k$.

$\boxed{\text{(b)}}$ Consider a function $k : S_\alpha \cup \{\alpha\} \to C$ given by

$$k(x) := \begin{cases} h(x) & : x \in S_\alpha \\ \min\left[C \setminus h(S_\alpha)\right] & : x = \alpha. \end{cases}$$

$\boxed{\text{(c)}}$ Let $x \in \bigcup_{\alpha \in K} S_\alpha$. There exists $\alpha \in K$ such that $x \in S_\alpha$. (a) implies that the value $h_\alpha(x)$ is uniquely determined independent of α. Thus, a function $k : \bigcup_{\alpha \in K} S_\alpha \to C$ given by

$$k(x) := h_\alpha(x)$$

is well-defined.

Note that $\bigcup_{\alpha\in K} S_\alpha$ is a subset of J, a well-ordered set, and hence it is also a well-ordered set. Let X be a subset of $\bigcup_{\alpha\in K} S_\alpha$ consisting of all x for which

$$k(x) = \min\left[C \setminus k(S_\alpha)\right]$$

holds. Let $x \in \bigcup_{\alpha\in K} S_\alpha$ and suppose $S_x \subset X$. Then, (b) gives $x \in X$, which means X is inductive. Thus, $X = \bigcup_{\alpha\in K} S_\alpha$ by the principle of transfinite induction.

$\boxed{\text{(d)}}$ Let Y be a subset of J consisting of all β for which there exists a function $h_\beta : S_\beta \to C$ satisfying (10.1). We only have to show that Y is inductive so that the principle of transfinite induction yields the result. Let $\beta \in J$ and suppose $S_\beta \subset Y$. If β has an immediate successor α, then we have $S_\beta = S_\alpha \cup \{\alpha\}$ and so (b) implies $\beta \in Y$. If not, $S_\beta = \bigcup_{\alpha<\beta} S_\alpha$, and hence (c) gives $\beta \in Y$. Thus, Y is inductive.

$\boxed{\text{(e)}}$ Uniqueness of h (if exists) follows from (a).

(d) and (c) imply that there exists a function

$$k : \bigcup_{\alpha\in J} S_\alpha \to C$$

satisfying (10.1). If J has a largest element m, then we have

$$\bigcup_{\alpha\in J} S_\alpha = S_m = J \setminus \{m\}$$

and so, (b) yields a function $h : J \to C$ satisfying (10.1). If not, then $J = \bigcup_{\alpha\in J} S_\alpha$ and we are done. □

Exercise 11. Let A and B be two sets. Using the well-ordering theorem, prove that either they have the same cardinality, or one has the greater cardinality greater than the other.

Solution. Apply, to A and B if necessary, the well-ordering theorem so that both A and B are well-ordered. Remember that, under the choice axiom, a consequence of the well-ordering theorem, $\mathrm{Surj}(A, B) \neq \emptyset$ if and only if $\mathrm{Inj}(B, A) \neq \emptyset$ by §9 Exercise 5 and §2 Exercise 5.

So, if $\mathrm{Inj}(A, B) = \emptyset$, or equivalently if $\mathrm{Surj}(B, A) = \emptyset$, then Exercise 10 gives $\mathrm{Inj}(A, B) \neq \emptyset$. This means that there cannot hold $\mathrm{Inj}(A, B) = \emptyset$ and $\mathrm{Inj}(B, A) = \emptyset$ at the same time, from which we conclude the result. □

11 The Maximum Principle

Exercise 1. If a and b are real numbers, define $a \prec b$ if $b - a$ is positive and rational. Show that this is a strict partial order on $\mathbb{R}$. What are the maximal simply ordered subsets?

Solution. For any $a \in \mathbb{R}$ we have $a - a = 0$, which gives $\neg a \prec a$. Suppose $a \prec b$ and $b \prec c$, or equivalently suppose $b - a \in \mathbb{Q}_+$ and $c - b \in \mathbb{Q}_+$. Then we have

$$c - a = (c - b) + (b - a) \in \mathbb{Q}_+,$$

which leads to $a \prec c$. Hence, $\prec$ is a strict partial order on $\mathbb{R}$.

A maximal simply ordered subset is the set

$$\left\{ b \in \mathbb{R} \mid {}^{\exists} q \in \mathbb{Q} \, [b = a + q] \right\},$$

where $a \in \mathbb{R}$ is arbitrary given. $\square$

Exercise 2.

(a) Let $\prec$ be a strict partial order on the set A. Define a relation on A by letting $a \preceq b$ if either $a \prec b$ or $a = b$. Show that this relation has the following properties, which are called the ***partial order axioms***:

(i) $a \preceq a$ for all $a \in A$.

(ii) $a \preceq b$ and $b \preceq a \Rightarrow a = b$.

(iii) $a \preceq b$ and $b \preceq c \Rightarrow a \preceq c$.

(b) Let P be a relation on A that satisfies properties (i)-(iii). Define a relation S on A by letting aSb if aPb and $a \neq b$. Show that S is a strict partial order on A.

Solution.

$\boxed{\text{(a)}}$ (i) is obvious by the fact $a = a$ for all $a \in A$. To verify (ii), observe that, exploiting transitivity of $\prec$, we have

$$a \prec b \wedge b \prec a \equiv a \prec a \wedge a \prec b \wedge b \prec a \equiv F,$$

and that

$$\begin{aligned} a \preceq b \wedge b \prec a &\equiv (a \prec b \vee a = b) \wedge (b \prec a \vee a = b) \\ &\equiv (a \prec b \wedge b \prec a) \vee a = b \\ &\equiv a = b. \end{aligned}$$

To see (iii), suppose $a \preceq b$ and $b \preceq c$. If both $a \prec b$ and $b \prec c$ hold, then transitivity of $\prec$ gives $a \prec c$. If not, it is easy to check $a \preceq c$.

$\boxed{\text{(b)}}$ Nonreflexivity is proved by seeing

$$\begin{aligned} {}^{\forall}a \in A\,[aPa] &\equiv {}^{\forall}a \in A\,[aPa \wedge a \neq a] \\ &\equiv {}^{\forall}a \in A\,[T \wedge F] \\ &\equiv F. \end{aligned}$$

To show transitivity, observe that (iii) on the partial order axioms gives

$$aSb \wedge bSc \Rightarrow aPc.$$

Moreover, (ii) on the partial order axioms gives

$$\begin{aligned} aSb \wedge bSc \Rightarrow a = c &\equiv aSb \wedge bSc \Rightarrow aSb \wedge bSa \wedge a = c \\ &\equiv aSb \wedge bSc \Rightarrow a = b \wedge a = c \\ &\equiv F, \end{aligned}$$

which implies

$$aSb \wedge bSc \Rightarrow a \neq c.$$

Thus, $aSb \wedge bSc \Rightarrow aSc$. □

Exercise 3. Let A be a strict partial order $\prec$; let $x \in A$. Suppose that we wish to find a maximal simply ordered subset B of A that contains x. One plausible way of attempting to define B is to let B equal the set of all those elements of A that are *comparable* with x;

$$B = \{y \in A \mid x \prec y \vee y \prec x\}.$$

But this will not always work. In which of Example 1 and 2 will this procedure succeed and in which will it not?

Solution. It fails in Example 1. Consider $A := \{\{1,2\}, \{3\}, \mathcal{P}(\mathbb{Z}_+)\}$, and $x := \mathcal{P}(\mathbb{Z}_+)$, in which case we find $B = \{\{1,2\}, \{3\}\}$ not simply ordered. □

Exercise 4. Given two points (x_0, y_0) and (x_1, y_1) of $\mathbb{R}^2$, define

$$(x_0, y_0) \prec (x_1, y_1)$$

if $x_0 < x_1$ and $y_0 \leq y_1$. Show that the curves $y = x^3$ and $y = 2$ are maximal simply ordered subsets of $\mathbb{R}^2$, and the curve $y = x^2$ is not. Find all maximal simply ordered subsets.

Solution. We begin with considering the last question for convenience. We claim that

(i) Let A be a subset of $\mathbb{R}^2$. A is simply ordered by $\prec$ if and only if it is the rule of a weakly increasing function, denoted by f_A, on a subset of $\mathbb{R}$.

(ii) Moreover, $(A, \prec)$ is a maximal simply ordered set if and only if the domain of f_A equals $\mathbb{R}$.

"if" part of (i) is easy to check. Consider "only if" part. Suppose $(x, y_0), (x, y_1) \in A$. Since these points are not comparable under $\prec$, a simple order, comparability of it implies $(x, y_0) = (x, y_1)$ and so, Hence, A is the rule of some function on a subset of $\mathbb{R}$. It follows from the definition of $\prec$ that so derived function is in fact weakly increasing. This completes the proof of (i).

We proceed to "if" part of (ii). Suppose the domain of f_A, denoted by D, coincides with $\mathbb{R}$. Then, for any $(a, b) \in \mathbb{R}^2 \setminus A$, there exists $(a, f_A(a)) \in A$. Clearly, (a, b) and $(a, f_A(a))$ are not comparable, which implies $A \cup \{(a, b)\}$ is no longer a simply ordered set. So, A is maximal.

We next show (contrapositive of) the converse. Suppose $D \subsetneq \mathbb{R}$, and let $p \in \mathbb{R} \setminus D$. It is easy to construct a weakly increasing function g on $D \cup \{p\}$ that extends f_A. Let G be the rule of g. (i) tells us that G is simply ordered, and $G \supsetneq A$, which means A is not a maximal simply ordered set. Thus, (ii) follows.

We deduce from (i) and (ii) that the set of all maximal simply ordered subsets is the set of the rules of all weakly increasing functions from $\mathbb{R}$ to $\mathbb{R}$.

It is now trivial to to verify the claim about $y = x^3$, $y = 2$, and $y = x^2$. □

Exercise 5. Show that Zorn's Lemma implies the following:
Lemma (Kuratowski). Let $\mathcal{A}$ be a collection of sets. Suppose that for every subcollection $\mathcal{B}$ of $\mathcal{A}$ that is simply ordered by proper inclusion, the union of the elements of $\mathcal{B}$ belongs to $\mathcal{A}$. Then $\mathcal{B}$ has an element that is properly contained in no other element of $\mathcal{A}$.

Solution. Proper inclusion induces a strict partial $\prec$ order on $\mathcal{A}$. Every simply ordered subcollection $\mathcal{B}$ of $\mathcal{A}$ admits an upper bound in $\mathcal{A}$ since the union of the elements of $\mathcal{B}$ belongs to $\mathcal{A}$ by assumption, which means $(\mathcal{A}, \prec)$ satisfies the hypothesis of Zorn's Lemma. Hence, $(\mathcal{A}, \prec)$ has a maximal element, which is property contained in no other element of $\mathcal{A}$ by definition. □

Exercise 6. A collection $\mathcal{A}$ of subsets of a set X is said to be of *finite type*, provided that a subset B of X belongs to $\mathcal{A}$ if and only if every finite subset of B belongs to $\mathcal{A}$. Show that Kuratowski lemma implies the following: Lemma (Tukey, 1940). Let $\mathcal{A}$ be a collection of sets. If $\mathcal{A}$ is of finite type, then $\mathcal{A}$ has an element that is properly contained in no other element of $\mathcal{A}$.

Solution. Suppose $\mathcal{A}$ be of finite type. Let $\mathcal{B}$ be a subcollection of $\mathcal{A}$ that is simply ordered by proper inclusion. Set $B^* := \bigcup_{B\in\mathcal{B}} B$. If F is a finite subset of B^*, then there exists $B \in \mathcal{B}$ such that $F \subset B$. Since $B \in \mathcal{A}$ and $\mathcal{A}$ is assumed to be of finite type, we deduce $F \in \mathcal{A}$, which leads to $B^* \in \mathcal{A}$. Thus, Kuratowski lemma implies Tukey lemma. □

Exercise 7. Show that Tukey lemma implies the Hausdorff maximum principle.

Solution. Let $(A, \prec)$ be a set A equipped with a strict partial order $\prec$, and let $\mathcal{A}$ be the collection of all subsets of A that are simply ordered by $\prec$. It suffices to show that $\mathcal{A}$ is of finite type thanks to Tukey lemma.

§3 Exercise 7 implies that if $B \in \mathcal{A}$, then every (finite) subset of B belongs to $\mathcal{A}$.

Conversely, let B be a subset of A of which every finite subset belongs to $\mathcal{A}$. We insist $B \in \mathcal{A}$. For, if x and y are two distinct points in B, then the fact $(\{x, y\}, \prec_{\{x,y\}})$ is a simply ordered set gives that B satisfies comparability, where $\prec_{\{x,y\}}$ is a restriction of $\prec$ onto $\{x, y\}$. Similarly, we deduce that B satisfies nonareflexivity and transitivity, and so B is simply ordered. This means $B \in \mathcal{A}$. Thus, $\mathcal{A}$ is of finite type. □

Exercise 8. A typical use of Zorn's lemma in algebra is the proof that every vector space has a basis. Recall that if A is a subset of the vector space V, we say a vector belongs to the *span* of A if it equals a finite linear combination of elements of A. The set A is *independent* if the only finite linear combination of elements of A that equals zero vector is the trivial one having all coefficient zero. If A is independent and if every vector in V belongs to the span of A, then A is a *basis* of V.

(a) If A is independent and $v \in V$ does not belong to the span of A, show $A \cup \{v\}$ is independent.

(b) Show the collection $\mathcal{I}$ of all independent sets in V has a maximum element.

(c) Show that V has a basis.

Solution.

$\boxed{\text{(a)}}$ Let A is independent and let $v \in V$. Suppose $A \cup \{v\}$ is not independent. There exist nontrivial coefficients $c_1, c_2, \cdots, c_n$ such that

$$c_1 a_1 + \cdots + c_{n-1} a_{n-1} + c_n v = 0. \tag{11.1}$$

We may assume all c_n are nonzero. It then follows from (11.1) that v belongs to the span of A, from which we deduce that $v \in V$ belongs to the span of A.

$\boxed{\text{(b)}}$ Introduce a strict partial order $\prec$ on $\mathcal{I}$ by setting $A \prec B$ if $A \subsetneq B$. Then, Hausdorff maximum principle yields a required maximal element.

$\boxed{\text{(c)}}$ Let $I \in \mathcal{I}$, and let v be a vector in V. Suppose v does not belong to the span of I. Then (a) tells us that $I \cup \{v\}$ is independent. This means I is not a maximal element of $\mathcal{I}$. Thus, every vector in V belongs to the span of some maximal element of $\mathcal{I}$. □

Supplementary Exercises: Well-Ordering

Exercise 1 (General principle of recursive definition)**.** Let J be a well-ordered set; let C be a set. Let $\mathcal{F}$ be the set of all functions mapping sections of J into C. Given a function $\rho : \mathcal{F} \to C$, there exists a unique function $h : J \to C$ such that $h(\alpha) = \rho(h|S_\alpha)$ for each $\alpha \in J$.

Solution. Copy the argument for §10 Exercise 10. □

Remark 11.1 (Uniqueness of an order-preserving function)**.** We establish here an useful result, which is a consequence of general principle of recursive definition (Exercise 1) and Exercise 2(a), as follows:

There exists at most one order-preserving function from a well-ordered set to another, whose image is the range of the function or a section of it.

For, if we suppose that there are two such functions, then Exercise 2(a) tells us that they both satisfy the formula indicated in Exercise 2(a)(ii), recursive formulas for themselves. Hence, general principle of recursive definition implies that these two function equal to one another.

Exercise 2.

(a) Let J and E be well-ordered sets; let $h : J \to E$. Show the following two statements are equivalent:

  (i) h is order preserving and its image is E or a section of E.

  (ii) $h(\alpha) = \min\,[E \setminus h(S_\alpha)]$ for all α.

(b) If E is a well-ordered set, show that no section of E has the order type of E, nor do two different sections of E have the same order type.

Solution.

$\boxed{\text{(a)}}$ We first show that (i) implies (ii). So, suppose (i) holds. Without loss of generality, we may assume that $h : J \to E$ is an order isomorphism. Then it follows that $h(S_\alpha) = S'_{h(\alpha)}$ for all $\alpha \in J$, where S'_y denotes the section of E by y, and that

$$\min\left[E \setminus h(S_\alpha)\right] = \min\left[E \setminus S'_{h(\alpha)}\right] = \sup S'_{h(\alpha)}$$

for all $\alpha \in J$, where we note that $E \setminus h(S_\alpha)$ is nonempty by the assumption of (i). Now the conclusion follows from the obvious fact

$$h(\alpha) = \sup S'_{h(\alpha)}$$

for all $\alpha \in J$.

To prove the converse, suppose (ii) holds. It follows from (ii) that $E \setminus h(S_\alpha)$ is the set of all upper bounds for $h(\alpha)$ for every $\alpha \in J$, from which we deduce that

$$S'_{h(\alpha)} = h(S_\alpha) \tag{11.2}$$

for every $\alpha \in J$. Moreover, 11.2 lets us see h preserve order. Indeed, if $\alpha < \beta$, or equivalently, if $\alpha \in S_\beta$, then $h(\alpha) \in h(S_\beta) = S'_{h(\beta)}$, that is, $h(\alpha) < h(\beta)$.

Assume $h(J) \subsetneq E$. We explicitly construct a section of E that equals $h(J)$.

If J has a largest element, denoted by m, then we have

$$h(J) = h(S_m) \cup \{h(m)\} = S'_{h(m)} \cup \{h(m)\} \subsetneq E.$$

This means $h(m)$ is not a largest element of E, and hence, $h(m)$ admits an immediate successor n in E by §10 Exercise 2. Then, $h(J) = S'_n$.

If not, neither does $h(J)$. Note that we can write $J = \bigcup_{\alpha \in J} S_\alpha$, and it follows from the fact

$$h(J) = \bigcup_{\alpha \in J} h(S_\alpha) = \bigcup_{\alpha \in J} S'_{h(\alpha)} \subsetneq E$$

that $h(J)$ is bounded above. Then §10 Exercise 1 tells us that $h(J)$ has a supremum s in E, more precisely, in $E \setminus h(J)$. It is straightforward to deduce from the property of supremum that $h(J) = S'_s$.

$\boxed{\text{(b)}}$ Let $\alpha \in E$. It is obvious that the inclusion function

$$i : S_\alpha \ni x \mapsto x \in E$$

preserves order, and its image $i(S_\alpha)$ equals S_α, a section of E, and that i is not surjective. We deduce from Note 11.1 that no order preserving function from S_α to E could be surjective, nor could be an order isomorphism. Thus, no section of E has the order type of E.

The second implication is a direct consequence of the first; Replace E with a section of E and use the innocent fact that *a section of a section of E is a section of E.* $\square$

Exercise 3. Let J and E be well-ordered sets; suppose there is an order preserving map $K : J \to E$. Using Exercises 1 and 2, show that J has the order type of E or a section of E.

Solution. Choose $e_0 := \min E$. Define a function $h : J \to E$ by the recursive formula

$$h(\alpha) := \begin{cases} \min\left[E \setminus h(S_\alpha)\right] & : h(S_\alpha) \neq E \\ e_0 & : h(S_\alpha) = E. \end{cases}$$

We show by transfinite induction that for every $\alpha \in J$

$$h(\alpha) \leq k(\alpha)$$

holds. Let X be a subset of J consisting of all α for which the above inequality holds. Let $\alpha \in J$ and suppose $S_\alpha \subset X$. We may assume $h(S_\alpha) \neq E$ thanks to the choice of e_0, and so $h(\alpha) = \min\left[E \setminus h(S_\alpha)\right]$. On the other hand, it follows that $k(\alpha)$ is greater than every element of $k(S_\alpha)$ and $h(S_\alpha)$ by assumption, which leads to $k(\alpha) \in E \setminus h(S_\alpha)$. Hence, $h(\alpha) \leq k(\alpha)$, completing the induction.

It follows from the just proved inequality that $h(S_\alpha) \subset S'_{k(\alpha)}$, and hence $h(S_\alpha) \neq E$ for all $\alpha \in J$, where $S'_{k(\alpha)}$ denotes the section of E by $k(\alpha)$. So, h satisfies

$$h(\alpha) = \min\left[E \setminus h(S_\alpha)\right]$$

for all $\alpha \in J$. Then Exercise 2(a) tells us that h is order preserving and its image is E or a section of E, from which we conclude the result. $\square$

Exercise 4. Use Exercise 1-3 to prove the following:

(a) If A and B are well-ordered sets, then exactly one of the following three conditions holds: A and B have the same order type, or A has the order type of a section of B, or B has the order type of a section of A.

(b) Suppose that A and B are well-ordered sets that are uncountable, such that every section of A and of B is countable. Show A and B have the same order type.

Solution.

$\boxed{\text{(a)}}$ We first show that at least one of the three conditions is the case. Let $(A, <_A)$ and $(B, <_B)$ be well-ordered sets, and let $(A \cup B, <)$ be the well-ordered set constructed in the way indicated in §10 Exercise 8. Let $S_x(A), S_x(B)$, and S_x denote the section of A, B, and $A \cup B$ by x, respectively. We may assume $A \neq \emptyset$ and $B \neq \emptyset$. Consider an inclusion function ι given by

$$\iota : A \ni x \mapsto x \in A \cup B.$$

It is obvious that ι preserves order and its image is A, a section of $A \cup B$ by a smallest element b_0 of B in $(A \cup B, <)$. Hence, Exercise 3 tells us that we have either $A \simeq A \cup B$ or

$$^{\exists 1} y_A \in A \cup B \, [A \simeq S_{y_A}] \, . \tag{11.3}$$

But there cannot hold $A \simeq A \cup B$ since ι is not a surjection, neither is an order isomorphism. Using Exercise 3, we similarly deduce that

$$B \simeq A \cup B \vee {}^{\exists 1} y_B \in A \cup B \, [B \simeq S_{y_B}] \tag{11.4}$$

but not both. Note that we cannot reject $B \simeq A \cup B$ this time, from which we deduce that

$$S_{f^{-1}(b_0)}(B) \simeq A, \tag{11.5}$$

where f is an order isomorphism from B to $A \cup B$.

Suppose $y_A \neq y_B$ and $y_B \in B$. If $y_B = b_0$, then we deduce from 11.4 that

$$B \simeq A.$$

If not, we have

$$S_{g^{-1}(b_0)}(B) \simeq A$$

where g is an order isomorphism from B to S_{y_B}. If $y_A \neq y_B$ and $y_B \in A$, then we have

$$B \simeq S_{y_B}(A).$$

At last, if $y_A = y_B$ happens to be the case, then (11.3),(11.4) yield

$$A \simeq B. \tag{11.6}$$

Thus, at least one of the three does hold. Note that every term in 11.5 - 11.6 is a section of $(A \cup B, <)$. Hence, Exercise 2(b) applied to $(A \cup B, <)$ establishes that none of two holds at the same time. Thus, we have proved that exactly one of the three holds.

$\boxed{\text{(b)}}$ Definition of uncountable set requires the absence of bijective correspondence of A and B with a countable set. Hence, Neither A nor B have the order type of a section of the other. Thus, we conclude from (a) that $A \simeq B$. □

Exercise 5. Let X be a set; let $\mathcal{A}$ be a collection of all pairs $(A,<)$, where A is a subset of X and $<$ is a well-ordering of A. Define

$$(A,<) \prec (A',<')$$

if $(A,<)$ *equals* a section of $(A',<')$.

(a) Show that $\prec$ is a strict partial order on $\mathcal{A}$.

(b) Let $\mathcal{B}$ be a subcollection of $\mathcal{A}$ that is simply ordered by $\prec$. Define B' to be the union of the sets B, for all $(B,<) \in \mathcal{B}$; and define $<'$ to be the union of the relations $<$, for all $(B,<) \in \mathcal{B}$. Show that $(B',<')$ is a well-ordered set.

Solution.
$\boxed{\text{(a)}}$ Nonreflexivity follows from Exercise 2(b). To see transitivity, suppose $(A,<_A) \prec (B,<_B)$ and $(B,<_B) \prec (C,<_C)$. By definition of $\prec$, there exist $b \in B$ and $c \in C$ such that

$$A \simeq S_b(B), \;\; B \simeq S_c(C).$$

It is easy by considering an order isomorphism to see

$$S_b(B) \simeq S_{c'}(C)$$

for some $c' \in C$. Then we have $A \simeq S_{c'}(C)$, from which we conclude $(A,<_A) \prec (C,<_C)$, as desired.

$\boxed{\text{(b)}}$ We first establish that $<'$ is a simple order on B'. Nonreflexivity is obviously satisfied. To see $<'$ satisfy comparability, let b_1 and b_2 be two distinct points in B'. There exist $(B_1,<_1),(B_2,<_2) \in \mathcal{B}$ with $(B_1,<_1) \prec (B_2,<_2)$ such that $b_1 \in B_1$ and $b_2 \in B_2$. Then $b_1, b_2 \in B_2$ and these points are comparable under $<_2$, and so are under $<'$. To prove transitivity, suppose that $b_1 <' b_2$ and $b_2 <' b_3$. We deduce, as we do to see comparability, that there exists $(B,<) \in \mathcal{B}$ such that $b_i \in B$ and $b_i < b_{i+1}$ for all i. Then transitivity of $(B,<)$ gives $b_1 < b_3$. Hence, $b_1 <' b_3$. Thus, $<'$ is a simple order on B'.

We show that $(B',<')$ is a well-ordered set. Let W be a nonempty subset of B'. The fact $\mathcal{B}$ is simply ordered by $\prec$ gives $(B,<) \in \mathcal{B}$ with $W \subset B$. Then W has a smallest element m in $(B,<)$, for which we have $m < b$ and hence $m <' b$ for all $b \in W \setminus \{m\}$. Thus, W has a smallest element m in $(B',<')$. $\square$

Exercise 6. Use Exercise 1 and 5 to prove the following:
Theorem. The maximum principle is equivalent to the well-ordering theorem.

Solution. We have already seen the well-ordering theorem imply the (Hausdorff) maximum principle. Here we establish the converse.

Let X be a set, and let $\mathcal{A}$ and $\prec$ be as at Exercise 5, say, let $\mathcal{A}$ be the collection of all pairs $(A, <)$, where A is a subset of X and $<$ is a well-ordering of A; let $(A, <) \prec (A', <')$ if $(A, <)$ equals a section of $(A', <')$. Note that $\mathcal{A}$ is nonempty since $(\emptyset, \emptyset)$ is vacuously well-ordered, and that if $(A, <) \in \mathcal{A}$ and $x \in X \setminus A$, then we can define a new well-ordered set $(A \cup \{x\}, <_{+x}) \in \mathcal{A}$, where $<_{+x}$ is defined to be equal to $<$ on A and $a <_{+x} x$ for all $a \in A$. Observe that Exercise 5 implies that $\mathcal{A}$ satisfies the hypothesis of Zorn's lemma, which, as we have seen in §11, is equivalent to the maximum principle. Hence, $\mathcal{A}$ has a maximal element. It is easy to verify that $A^* = X$. □

Exercise 7. Use Exercise 1-5 to prove the following:
Theorem. The choice axiom is equivalent to the well-ordering theorem.
Proof. Let X be a set; let c be a fixed choice function for the nonempty subsets of X. If T is a subset of X and $<$ is a relation on T, we say that $(T, <)$ is a ***tower*** in X if $<$ is a well-ordering of T and if for each $x \in T$,

$$x = c\left(X \setminus S_x(T)\right), \tag{11.7}$$

where $S_x(T)$ is the section of T by x.

(a) Let $(T_1, <_1)$ and $(T_2, <_2)$ be two towers in X. Show that either two ordered sets are the same, or one equals a section of the other.

(b) If $(T, <)$ is a tower in X and $T \neq X$, show there is a tower in X of which $(T, <)$ is a section.

(c) Let $\{(T_k, <_k) \mid k \in K\}$ be the collection of all tower in X. Let

$$T := \bigcup_{k \in K} T_k, \quad <:= \bigcup_{k \in K} (<_k).$$

Show that $(T, <)$ is a tower in X. Conclude that $T = X$.

Solution.

$\boxed{\text{(a)}}$ By Exercise 5, we may assume that there exists a function $h : T_1 \to T_2$ that is order-preserving and $h(T_1)$ equals either T_2 or a section of T_2. Exercise 2 tells us that h satisfies

$$h(x) = \min\left[X \setminus h(S_x(T_1))\right]$$

and

$$h\left(S_x(T_1)\right) = S_{h(x)}(T_2) \tag{11.8}$$

for all $x \in T_1$. We show by transfinite induction that h is an identity function on T_1, that is,

$$h(x) = x$$

for all $x \in T_1$. Let I be a subset of T_1 consisting of all x for which the above equality holds. Let $x \in T_1$ and suppose $S_x(T_1) \subset I$. We then deduce from (11.8) that

$$h(x) = c(X \setminus S_{h(x)}(T_2)) = c(X \setminus h(S_x(T_1))) = c(X \setminus S_x(T_1)) = x,$$

and hence $x \in I$, completing the induction. Thus, T_1 coincides with its image under h, namely, with either T_2 or a section of T_2.

$\boxed{\text{(b)}}$ Let $p := c(X \setminus T)$. As in Exercise 6, we can define a new well-ordered set $(T \cup \{p\}, <_{+p})$. Observe that there holds $S_p(T \cup \{p\}) = T$ and so

$$p = c(X \setminus T) = c(X \setminus S_p(T \cup \{p\})),$$

from which we conclude that $(T \cup \{p\}, <_{+p})$ is a tower, as required. (Note that $(\emptyset, \emptyset)$ is a trivial tower, and this procedure yields the existence of a non-trivial tower in X that contains arbitrary element of X.)

$\boxed{\text{(c)}}$ Define

$$(T_k, <_k) \prec (T_\ell, <_\ell)$$

if $(T_k, <_k)$ equals a section of $(T_\ell, <_\ell)$. (a) and Exercise 5 imply that $\prec$ is a simple order on $\{(T_k, <_k)\}$. Hence, Exercise 5 tells us that $(T, <)$ is a well-ordered set. It remains to show that $(T, <)$ satisfies (11.7) since (b) then gives $T = X$.

Let $x \in T$. There exists $k_1 \in K$ for which $x \in T_{k_1}$. Then there holds

$$x \in c(X \setminus S_x(T_{k_1})).$$

We claim

$$S_x(T_{k_1}) = S_x(T).$$

To this end, we only have to show that
If $x \in T$ and if $x \in T_k$ for some $k \in K$, then $S_x(T_k) = S_x(T_{k'})$ for all $k' \in K$ for which $x \in T_{k'}$,
since we then deduce that, under the assumption $x \in T_{k_1}$, there holds

$$t \in S_x(T) \equiv {}^\exists k \in K\, [t \in S_x(T_k)] \equiv t \in S_x(T_{k_1}).$$

To prove the italic statement, let $x \in T$ and $x \in T_k \cap T_{k'}$. (a) implies that we may assume that there exists $s \in T_{k'} \setminus T_k$ such that $T_k = S_s(T_{k'})$. This leads to

$$\begin{aligned} S_x(T_k) &= \{t \in T_k \mid t <_k x\} \\ &= \{t \in S_s(T_{k'}) \mid t <_{k'} x\} \\ &= \{t \in T_{k'} \mid t <_{k'} x\} \\ &= S_x(T_{k'}). \end{aligned}$$

We have shown so far that a choice function for $\mathcal{P}(X) \setminus \emptyset$ yields the set of non-trivial towers in X, which gives a tower $(X, <)$. In other words, we have seen the choice axiom induce a well-ordering (even with the tower property). Thus, the well-ordering theorem follows from the choice axiom. $\square$

Exercise 8. Using the Exercise 1-4, construct an uncountable well-ordered set, as follows. Let $\mathcal{A}$ be the collection of all pairs $(A, <)$, where A is a subset of $\mathbb{Z}_+$ and $<$ is a well-ordering of A. (We allow A to be empty.) Define $(A, <) \simeq (A', <')$ if $(A, <)$ and $(A', <')$ have the same order type. It is trivial to show this is an equivalence relation. Let $[(A, <)]$ denote the equivalence class of $(A, <)$; let E denote the collection of there classes. Define

$$[(A, <)] \ll [(A', <')] \tag{11.9}$$

if $(A, <)$ has the order type of a section of $(A', <')$.

(a) Show that the relation $\ll$ is well defined and is a simple order on E. Note that the equivalence class $[(\emptyset, \emptyset)]$ is the smallest element of E.

(b) Show that if $\alpha = [(A, <)]$ is an element of E, then $(A, <)$ has the same order type as the section $S_\alpha(E)$ of E by α.

(c) Conclude that E is well-ordered by $\ll$.

(d) Show that E is uncountable.

Solution.

$\boxed{\text{(a)}}$ It suffices to show that $\ll$ is well-defined since it then follows from Exercise 2 and 3 that $\ll$ is a simply ordered. Note that Exercise 4 tells us that no member of $[(A, <)]$ has the order type of a section of a member of $[(A, <)]$. Suppose $(A, <_A)$ has the order type of a section $S_b(B)$ of $(B, <_B)$, which has the same order type as $(C, < C)$. Then $(A, <_A)$ has the order type of a section $(C, < C)$. This means that the relation 11.9 is determined independent of the choice of $(A', <_{A'})$. On the other hand, if $(D, <_D)$ has the same order type as $(A, <_A)$ which has the order type of $(B, <_B)$, then it is obvious that we have $(D, <_D)$ also has the order type of a section of $(B, <_B)$.

Hence, the relation 11.9 is determined independent of the choice of $(A, <_A)$ and $(A, <_{A'})$. In other words, the relation $\ll$ is well-defined.

$\boxed{\text{(b)}}$ Define a function $f : A \to E$ via

$$f(x) := \left[(S_x(A), <_{S_x(A)})\right],$$

where $<_{S_x(A)}$ is the restriction of $<$ onto $S_x(A)$. f preserves order by Exercise 2 and 4, and so it suffices to show $f(A) = S_\alpha(E)$. Clearly, $f(A) \subset S_\alpha(E)$. Let $e \in S_\alpha(E)$. We have $e = [(A', <_{A'})]$ for some $(A', <_{A'})$ satisfying $[(A', <_{A'})] \ll \alpha$, or equivalently, for some $(A', <_{A'})$ having the order type of a section of $(A, <_A)$. This means $e \in f(A)$, and hence $S_\alpha(E) \subset f(A)$.

$\boxed{\text{(c)}}$ Let $\mathcal{E}$ be a nonempty subset of E, and let $\alpha \in \mathcal{E}$. We may assume $S_\alpha(E) \cap \mathcal{E} \neq \emptyset$, and it suffices to show that $S_\alpha(E) \cap \mathcal{E}$ has a smallest element. If $\alpha = [(A, <_A)]$, then (b) gives

$$(A, <_A) \simeq S_\alpha(E),$$

from which we deduce that $S_\alpha(E) \cap \mathcal{E}$ has the order type of $(A, <_A)$ or a section of $(A, <_A)$. Since every nonempty subset of $(A, <_A)$ admits a smallest element, so does $S_\alpha(E) \cap \mathcal{E}$.

$\boxed{\text{(d)}}$ We ignore the advice by Munkres to argue by contradiction, and instead show that no injective function from $\mathbb{Z}_+$ to E is surjective. Let $k : \mathbb{Z}_+ \to E$ be an injection. It is obvious that such injection does exist; consider, for instance, a function $\mathbb{Z}_+ \ni n \mapsto [(S_n, <)] \in E$, where $<$ is the restriction of the usual order of $\mathbb{Z}_+$. Define a simple order $\prec_k$ on $\mathbb{Z}_+$ by setting $x \prec_k y$ if $k(x) \ll k(y)$. It is easy to check that $\prec_k$ is well-defined and k is an order-preserving function from $(\mathbb{Z}_+, \prec_k)$ to $(E, \ll)$, from which we deduce that $\prec_k$ is a well-ordering of $\mathbb{Z}_+$. Then it turns out that k is not surjective. In fact, we deduce from the fact $(\mathbb{Z}_+, \prec_k)$ has the order type of a section of E that the image of k does not coincide with E. Hence, k is not surjective.

The fact $\mathrm{Inj}(\mathbb{Z}_+, E) \neq \emptyset$ and none of its member is surjective tells us that E is infinite but is not in bijective correspondence with $\mathbb{Z}_+$. Thus, E is uncountable. $\square$

References

[1] James Munkres. *Topology.* Prentice Hall, Inc, 2nd ed edition, 2000.